

2016

# Scalar Short-Term Memory

Tyler D. Bancroft

Wilfrid Laurier University, banc6110@mylaurier.ca

Follow this and additional works at: <http://scholars.wlu.ca/etd>

 Part of the [Computational Neuroscience Commons](#)

---

## Recommended Citation

Bancroft, Tyler D., "Scalar Short-Term Memory" (2016). *Theses and Dissertations (Comprehensive)*. 1825.  
<http://scholars.wlu.ca/etd/1825>

This Dissertation is brought to you for free and open access by Scholars Commons @ Laurier. It has been accepted for inclusion in Theses and Dissertations (Comprehensive) by an authorized administrator of Scholars Commons @ Laurier. For more information, please contact [scholarscommons@wlu.ca](mailto:scholarscommons@wlu.ca).

Scalar Short-Term Memory

by

Tyler D. Bancroft

Bachelor of Arts, Wilfrid Laurier University, 2010  
Master of Science, Wilfrid Laurier University, 2011

DISSERTATION

Submitted to the Department of Psychology

in partial fulfillment of the degree requirements for

Doctor of Philosophy in Psychology

Wilfrid Laurier University

© Tyler D. Bancroft 2016

### **Declaration of Co-Authorship and/or Previous Publication**

This dissertation is largely based on the following manuscripts, which were produced with the listed individuals as co-authors. Manuscripts that have been published or accepted in the below form or a substantially similar form are noted as such. Co-authors have contributed to the below manuscripts (and, by extension, the present dissertation) intellectually, in research, and in writing, however, I declare that the majority of all said manuscripts are my work. Individual author contributions are listed in Appendix B.

Chapter 2: Bancroft, T. D., Hockley, W. E., & Servos, P. (2014). Does stimulus complexity determine whether working and short-term memory storage relies on prefrontal cortex or sensory cortex? *Attention, Perception, & Psychophysics*, *76*, 1954-1961.

Chapter 3: Bancroft, T. D., Hogeveen, J., Servos, P., & Hockley, W. E. (2014). TMS-induced neural noise in sensory cortex interferes with short-term memory storage in prefrontal cortex. *Frontiers in Computational Neuroscience*, *8*, 23.

Chapter 4: Bancroft, T. D., Hockley, W. E., & Servos, P. (2014). A shared short-term memory system for stimulus frequency and duration. *Canadian Journal of Experimental Psychology*, *68*, 236-241.

Chapter 5: Bancroft, T. D., Jones, J. A., Ensor, T. M., Hockley, W. E., & Servos, P. (in press). Inhibition, interference, and intrusion in short-term memory. *Memory & Cognition*.

### **Abstract**

The location of the brain's working and short-term memory (WM/STM) "system" is unclear. The existence of a dedicated WM/STM system is itself under debate. Recently, it has been proposed that WM/STM storage relies not on a dedicated system in prefrontal cortex, but rather that it is an emergent function of interaction between attentional and representational systems (e.g., sensory cortex) in the brain. However, mnemonic representations of very simple stimuli have repeatedly been shown to exist in frontal cortex. In this manuscript, I use computational and behavioural methods to demonstrate similarities between the representations of different types of very simple stimuli in memory, and argue that this is evidence that present theory must be extended to deal with current experimental results.

### **Acknowledgments**

There are many people to whom I owe thanks, and this page doesn't have room to list them all; names not listed should not be construed as evidence of ingratitude. I would be remiss, however, if I did not thank my committee (Britt Anderson, Bill Hockley, Jeff Jones, Diano Marrone, and Philip Servos), my research collaborators (Jeremy Hogeveen and Tyler Ensor), administrators whose help has been invaluable over my years at Laurier (Dawn Pavich, Ann Reid, and Rita Sharkey), and Trevor Saunderson from the Department of Mathematics, who has kindly provided me with mathematical assistance for many years. I would also like to thank Kay, Heidi, and Joe for their much-appreciated contributions to my sanity.

I am indebted to Justin Harris for providing us with the experimental means from Harris et al. (2002), used in Chapter 3. I also appreciate the suggestions provided by the various anonymous reviewers who reviewed the published manuscripts contained in this dissertation.

### **Statement of Originality**

This dissertation is based upon original research undertaken by the author. Contributions by co-authors and collaborators to the manuscripts comprising this dissertation are stated in the Declaration of Co-Authorship and/or Previous Publication, and in Appendix B.

## Table of Contents

Chapter 1: Introduction	10
Chapter 2: Does stimulus complexity determine whether working and short-term memory storage relies on prefrontal cortex or sensory cortex?	15
Chapter 3: TMS-induced neural noise in sensory cortex interferes with short-term memory storage in prefrontal cortex	34
Chapter 4: A shared short-term memory system for stimulus frequency and duration	53
Chapter 5: Overwriting and intrusion in short-term memory	70
Chapter 6: General Discussion	91
References	97
Appendix A: Copyright information for previously published manuscripts	114
Appendix B: Collaborator contributions to manuscripts	115

**List of Tables**

Table 1: Simulation parameters	p. 43
Table 2: Simulation parameters	p. 60
Table 3: Mean proportions of correct responses from Experiment 1	p. 80
Table 4: Mean proportions of correct responses from Experiment 2	p. 85



## List of Figures

Figure 1: Diagram of a C/M/D triplet	p. 41
Figure 2: Simulated and empirical results of TMS to ipsilateral SI	p. 45
Figure 3: Simulated and empirical results of TMS to contralateral SI	p. 46
Figure 4: Values of $r_M$ and $r_D$ during duration tasks, with $w_{MM} = 1$	p. 61
Figure 5: Values of $r_M$ and $r_D$ during duration tasks, with $w_{MM}$ variable	p. 64

**List of Appendices**

Appendix A: Copyright information for published manuscripts	p. 114
Appendix B: Collaborator contributions to manuscripts	p. 115

## Chapter 1: Introduction

Working memory and short-term memory (WM/STM) has traditionally been one of the main areas of research in cognitive psychology, and experiments treating WM/STM as separate from long-term memory were reported in the early 1900s (Thorndike, 1910). Over the past century, a variety of theoretical treatments of WM/STM have emerged, ranging from WM/STM being a single system, lying between perception and long-term memory (e.g., Atkinson & Shiffrin, 1968), to it being a collection of cognitive processes, dedicated to memory storage and maintenance (e.g., Vallar & Baddeley, 1984; Repovš & Baddeley, 2006), to WM/STM emerging from the coordinated activity of various neural systems that have evolved to serve non-mnemonic purposes (Postle, 2006). Despite decades of research devoted to the study of working and short-term memory, we have yet to develop a coherent theoretical picture of the neural processes and cognitive processes that underlie working and short-term memory.

The multi-components (or the multicomponents, or the multiple-components) model of working memory has a long pedigree stretching back to the 1970s, and has arguably exerted more influence on theory than any other model. It contains modality-specific buffers, coupled with a central executive system. In earlier instantiations, the model contained a buffer for visual and spatial information (the visuospatial sketchpad) and a buffer for auditory information (the phonological loop); more recently, a third buffer (the episodic buffer) has been added (e.g.,

Baddeley, 2000, 2001, 2012; Baddeley & Hitch, 2010; Repovs & Baddeley, 2006). Sustained activity in prefrontal cortex was proposed as the neural instantiation of a working memory system (see Postle, 2006, for a discussion).

Postle (2006), however, criticized this “standard model” on the grounds of parsimony and explanatory power. He noted that humans’ ability to store information in short-term memory is not limited to the domains included in the multi-components model: For example, short-term memory for both tactile (e.g., Burton & Sinclair, 2000; Harris, Harris, & Diamond, 2001; Harris, Miniussi, Harris, & Diamond, 2002; Romo & Salinas, 2003) and olfactory (e.g., Dade, Zatorre, Evans, & Jones-Gottman, 2001) information exists. Further, Postle pointed out that cognitive and neural dissociations exist between mental representations of information within the domains that are represented in the model (for example, a dissociation between manipulable versus nonmanipulable objects; Mecklinger et al., 2002). Postle also noted that frontal damage does not eradicate WM/STM, although it may create deficits under certain conditions (such as distraction).

As an alternative, he proposed working memory was an emergent property of the brain, resulting from the conjunction of attentional systems and representational systems, such as sensory cortex. In this view, WM/STM is essentially the maintenance of activated representations in neural systems already capable of representing a certain type of information.

For example, visual information would be maintained in visual cortex, somatosensory information in somatosensory cortex, and so forth.

This view is supported by an increasing number of experimental studies across a variety of domains, ranging from the use of fMRI decoding methods to determine the contents of visual short-term memory from activity in visual cortex, (e.g., Christophel, Hebart, & Haynes, 2012), to single-cell recording from non-human primates (e.g., Zhou & Fuster, 1996). However, the literature also contains a compelling body of evidence in favour of memory encoding in prefrontal cortex, from research on a form of short-term memory known as scalar short-term memory.

Scalar STM tasks are those where the salient property of a to-be-remembered stimulus can be represented as a scalar (i.e., unidimensional value). Scalar short-term memory has also been referred to as parametric working memory. However, this is potentially inaccurate in that such tasks do not necessarily recruit the executive/attentional resources that may be the signature of working memory (e.g., Engle, 2002), and may include categorical, not just parametric stimuli (Freedman, Riesenhuber, Poggio, & Miller, 2001), although this interpretation of the literature is still very much open to question (e.g., Sarma, Masse, Wang, & Freedman, 2016). Examples of stimulus types used in scalar STM include stimulus frequency (auditory, vibrotactile, or visual), stimulus amplitude, stimulus duration, and the numerosity of a set of stimuli (e.g., Spitzer &

Blankenburg, 2012; Spitzer, Fleck, & Blankenburg, 2014; Spitzer, Gloel, Schmidt, & Blankenburg, 2014; Spitzer, Wacker, & Blankenburg, 2010). Converging evidence from single-cell recording, EEG/MEG, and behavioural paradigms has suggested that scalar STM relies on shared, modality-independent neural systems and codes, and that the mnemonic representations of those stimuli are maintained in the frontal lobe, not in sensory cortex.

The aim of this dissertation is to use experimental and computational methods to test the hypothesis that there are similarities between the mental and neural representations of information in different forms of scalar STM, thereby evaluating the argument that these different scalar STM tasks recruit a shared neural system, most likely located in frontal cortex.

Four manuscripts are included below. The first, “Does stimulus complexity determine whether working and short-term memory storage relies on prefrontal cortex or sensory cortex?”, serves as an introduction to recent findings from the neuroscience literature and argues that simple (i.e., scalar) stimuli appear to share a modality-independent representation in prefrontal cortex, while more complex stimuli rely on task-relevant cortex, such as sensory cortex.

The second, “TMS-induced neural noise in sensory cortex interferes with short-term memory storage in prefrontal cortex”, is a modeling study that reconciles two views of scalar STM storage. Harris et al. (2002) reported that transcranial magnetic stimulation of primary somatosensory cortex during a scalar STM task (for vibrotactile stimuli) negatively impacted

performance, and suggested that somatosensory cortex acted as a (temporary) storage substrate for scalar STM; this view, however, is inconsistent with single-cell recording data that suggests that primary somatosensory cortex does not maintain a stimulus representation after stimulus offset (e.g., Romo & Salinas, 2003). We were able to replicate Harris et al.'s results by assuming that somatosensory cortex did not act as a storage system, but rather that TMS induced activity in somatosensory cortex that then fed-forward into the memory storage system in prefrontal cortex, degrading the stored representation of the remembered stimulus.

The third manuscript, "A shared memory system for stimulus duration and stimulus frequency" demonstrates that the model can also retain duration information, and uses the model to replicate the subjective shortening effect, a classic finding in the short-term memory literature.

The fourth manuscript, "Overwriting and intrusion in short-term memory" extends previous experimental findings of interference in vibrotactile STM to the auditory domain.

Finally, in the general discussion, the findings of the studies included in this dissertation are integrated and I discuss some implications for WM/STM theory.

## **Chapter 2: Does stimulus complexity determine whether working and short-term memory storage relies on prefrontal cortex or sensory cortex?**

Tyler D. Bancroft, William E. Hockley, and Philip Servos

### **Abstract**

Traditionally, working and short-term memory (WM/STM) have been believed to rely on storage systems located in prefrontal cortex. However, recent experimental and theoretical efforts have suggested that in many cases, sensory or other task-relevant cortex is the actual storage substrate for WM/STM. What factors determine whether a given WM/STM task relies on PFC or sensory cortex? In the present paper, we outline recent experimental findings and suggest that the dimensionality or complexity of the to-be-remembered property or properties of a stimulus can be a determining factor.

### **Introduction**

The working and short-term memory (WM/STM) literature<sup>1</sup> is in flux, and the development of theory has not kept up with the often confusing complexity of the experimental literature. Traditional working memory theory generally assumes that WM/STM storage relies on dedicated neural systems, often located in prefrontal cortex (PFC) (for example, D'Esposito et al., 1995). In contrast, Postle (2006) recently proposed that PFC does not contain dedicated storage systems, but rather that WM/STM storage relies on neural systems in task-relevant cortex (for example, sensory cortex) that contain pre-existing representations of a given type of stimulus. Evidence in support of both of these positions can be found in the literature, raising an important question: Why does WM/STM storage sometimes rely on neural systems in PFC, and other times rely on neural systems in sensory (or other task-relevant) cortex? In the present



paper, we will review some of the most interesting and relevant findings, and propose that stimulus complexity or dimensionality is one of the determining factors.

Recently, Postle (2006) developed a new model of working memory, the emergent-property model. While the classic multiple-components model (see Baddeley, 2010, for a recent overview) suggests that working memory relies on dedicated systems, the emergent-property model suggests that working memory emerges from the interaction of task-relevant sensory and cognitive systems that also (sometimes primarily) serve other purposes. The emergent-property model has the capacity to fit a wide variety of experimental results, including findings from neuroimaging and neuropsychological research. For example, it proposes that information in WM is maintained in task-relevant cortex, with the support of attentional/executive systems in PFC and parietal cortex. Postle explicitly states that PFC is not involved in WM storage. This model is capable of explaining a wider swath of the WM literature than the multiple-components model, in part because it is firmly rooted in the neuropsychological and neuroscientific literatures. However, it has a clear analogue in the psychological literature, in the work of Engle et al. (1999). Postle's division of the neural systems underlying working memory into storage substrate (e.g., task-relevant sensory or cognitive cortex) and executive system(s) is consistent with Engle et al.'s (1999) analysis of the correlates of working memory performance. In Engle et al.'s study, subjects performed a variety of tasks, including eleven WM/STM tasks (including classic tasks such as reading span and operation span), two intelligence tests (Raven's Matrices and Cattell's Culture-Free Test), and the verbal and quantitative sections of the SAT. Engle et al. performed a set of factor analyses to examine the relationship between working memory, short-term memory, and intelligence, and

found that both working memory and short-term memory were correlated with a common memory construct, but only working memory was correlated with fluid intelligence. These results suggest that executive processes may mediate the relationship between working memory and fluid intelligence.

While the emergent-property model represents a substantial step forward in our understanding of WM/STM, it has become increasingly evident that the model has certain deficiencies, and recent experimental findings have falsified certain core principles of the model (for example, that PFC is never involved in memory storage). The application of decoding methods to data has found evidence for information maintenance in regions other than task-relevant sensory or cognitive cortex, such as spatial attention systems in parietal cortex (Christophel et al., 2012, fMRI), and (contrary to Postle's explicit statement otherwise) in prefrontal cortex (Freedman et al., 2001, single-cell recording; Spitzer, Wacker, & Blankenburg, 2010, Spitzer & Blankenburg, 2011, 2012, and Spitzer et al., 2014, all EEG). In essence, we are presently without a single theory or model that can account for all major aspects of WM/STM storage, including the crucial question of where the storage system(s) can be found in the brain. We will outline some of the most relevant recent research, and suggest that the complexity or dimensionality of a stimulus is one factor that helps determine whether storage processes in PFC or task-relevant/sensory cortex are recruited to store a the representation of a stimulus.

### **Defining complexity**

How do we define the complexity of a stimulus? Readers should note that we are generally referring to the salient portion or property of a stimulus, rather than the whole stimulus. To fully describe an auditory pure tone, we would need to give the pitch (frequency) and

amplitude, but auditory STM tasks are often based on comparing only the frequencies of two stimuli. In this sense, the stimulus is multidimensional, but the salient property of the stimulus can be represented as a scalar value - just the frequency. Consider a task where subjects are presented with an image of a cat or a dog, followed by a delay period, followed by a second image of a cat or a dog, and asked to decide whether the two images belong to the same category (i.e., are they both cats or both dogs) or different categories. (For research along this line, see Freedman et al., 2001, discussed below). The stimuli (images) are obviously quite complex, but the salient aspect of the stimuli (category) is simple and scalar.

To quantify this notion, the storage of information on a computer is perhaps an appropriate analogy. To store the frequency of a simple auditory stimulus on a computer requires a single variable (in other words, it is a scalar, or unidimensional quantity). To store the frequency of a complex auditory stimulus created by superimposing multiple sine waves requires multiple variables (one for each sine wave), making it a multidimensional quantity. Storing a visual image in memory may require a variable or variables for each pixel in the image being viewed (for example, as is the case when an image is stored on a computer in a bitmap format). However, it should be remembered that we are generally referring to the salient aspects of a stimulus, as determined by task requirements. If subjects are presented with an image and asked to remember if it contains green trees or red tomatoes, it is possible (perhaps probable) that subjects will retain a label such as “green” or “red”, rather than a representation of the entire image. Note also that the above is not intended to be a rigorous metric for stimulus complexity or dimensionality, and the development of such a measure would be of benefit.

### **Some WM/STM tasks rely on PFC for information storage**

Research into the neural correlates of scalar short-term memory (in other words, memory for unidimensional stimulus properties, such as the pitch of an auditory pure tone, or the duration of a stimulus) has produced strong evidence for information storage in PFC, rather than in task-relevant sensory cortex. Extensive single-cell research in non-human primates has revealed frontal coding for stimulus frequency (for both vibrotactile and auditory pure tones) during the delay period of memory tasks (see Romo & Salinas, 2003, for a review; also see Romo et al., 1999, and Lemus, Hernández, & Romo, 2009b). In fact, the neural coding of these stimuli is surprisingly simple: Neuronal firing rates tend to be monotonic (positive or negative) functions of stimulus frequency (e.g., Romo et al., 1999). This simple neural code has facilitated several innovative EEG studies by Spitzer and colleagues (Spitzer, Wacker, & Blankenburg, 2010; Spitzer & Blankenburg, 2011, 2012; Spitzer, et al., 2014), who developed a method to decode the representation of a stored stimulus from beta-band EEG activity in the frontal cortex during the delay period of memory tasks. This method has allowed decoding of memory for vibrotactile frequency, auditory pure tone frequency, and visual flicker frequency (Spitzer et al., 2010; Spitzer & Blankenburg, 2012), as well as stimulus amplitude and duration (Spitzer et al., 2014). Source localization suggested that prefrontal (possibly inferior frontal, see Spitzer et al., 2010) cortex was responsible for storage. Sensory cortex, on the other hand, was not shown to maintain stimulus representations during the delay period. It should be noted, however, that frequency information has been shown to be represented in primary somatosensory cortex, and Spitzer et al.'s EEG findings are generally consistent with single-cell work done by the group of Romo et al. (see Romo & Salinas, 2003, for a review). Further, Zhou and Fuster (1996) have demonstrated maintenance of other types of stimulus information in primary somatosensory

cortex, suggesting that PFC storage is not due to an inability to maintain information in somatosensory cortex. The lack of representation outside of PFC poses a challenge to arguments that delay-period activity in PFC does not code stimulus information, but rather always represents attentional activity.

Other types of simple information have also been shown to be encoded in PFC: For example, Sakurai et al. (2004) found frontal neurons that represented stimulus duration during the delay period of a memory-for-duration task. Categorical information has also been shown to be encoded in PFC. Freedman et al. (2001) found neurons that represented the category of novel forms of stimuli (computationally morphed images of cats and dogs) in PFC during the delay period of memory tasks. While these stimuli are, obviously, far more complex than the stimuli we discuss above, task demands (in this case, match-to-category) do not require storage of stimulus information beyond category, which can be represented as a scalar value (dog = 1, cat = 2, etc.). More recently, Lee et al. (2013) have demonstrated prefrontal encoding for stimulus category using fMRI decoding methods in humans. In general, these findings directly contradict one of the main tenets of the emergent-property model: PFC is not involved in memory storage.

**Some WM/STM tasks rely on task-relevant (e.g., sensory or cognitive) cortex for storage**

In contrast to these results, however, are a group of recent imaging studies that have used various decoding methods (see Tong & Pratte, 2012) to examine the contents of memory, and which have found memory storage in task-relevant (often sensory) cortex, and not in prefrontal cortex. These decoding methods, which include multivoxel pattern analysis and classification methods based on linear support vector machines, allow a relatively direct examination of the contents of memory.

Serences et al. (2009) used multivoxel pattern analysis methods to identify visual stimulus information (orientation or colour) in primary visual cortex during the delay period; further, this activity was similar to that observed during sensory processing. Christophel et al. (2012) applied decoding methods to activity found during an fMRI study of the neural correlates of WM/STM for abstract coloured stimuli. During the maintenance period of the task, stimulus information was shown to be represented in early visual cortex, but not in prefrontal cortex. Intriguingly, they also identified stimulus information in posterior parietal cortex; as this area has retinotopic attentional maps, this may be related to attentional support of activity in visual cortex (Constantinidis & Steinmetz, 1996; Sereno et al., 2001). Similar findings have been reported by other researchers (e.g., Emrich et al. (2013) reported similar findings for visual STM over different memory loads).

Using a task involving STM for moving visual stimuli, Riggall and Postle (2012) found storage of stimulus information in posterior visual regions when applying multivoxel pattern analysis to fMRI data. In contrast, their analyses revealed that prefrontal activity reflected task instructions, not memory storage. (Linden et al. (2012) also reported a failure to find stimulus-specific activity in prefrontal cortex.) Perhaps most startlingly, Serences et al. (2009) and Riggall and Postle (2012) were able to identify stimulus information in cortical regions that did not display elevated delay-period activity, suggesting that standard analyses of BOLD signals, even load-dependent signals, may not be sufficient to identify regions involved in WM/STM storage. Harrison and Tong (2009) also reported similar findings, and were able to decode the orientation of a grating held in memory with approximately 80% accuracy, based on the application of a linear classifier to fMRI scans of early visual cortex - even though overall

levels of sustained activity were low. Ester et al. (2013) used similar methods to link the contents of visual cortex to subject performance.

Linke et al. (2011) applied fMRI decoding methods to sensory cortex during an auditory STM (for frequency) task. While frequency-specific increased activity was identified during stimulus encoding, it was absent during the maintenance period. Further, during maintenance, activity was actually suppressed below baseline, and this suppression was frequency-specific, suggesting that it may be actively inhibited during maintenance, potentially to protect the contents of memory in neural systems downstream.

Support for memory maintenance in sensory cortex also comes from other methodologies, such as single-cell recording. Zhou and Fuster (1996) reported object-selective delay-period activity in primary somatosensory neurons during a haptic memory task. Hayden and Gallant (2013) found stimulus-specific activity in V4 neurons during a delayed match-to-sample task. In addition, recent TMS research has supported a role for language cortex in verbal WM storage: Acheson et al. (2011) observed verbal WM deficits when applying TMS to regions of temporal cortex involved in language production. Note that we are focusing on activity-related memory in sensory cortex, not “structural” memory (i.e., plasticity; Menning et al., 2000; Weinberger, 2004).

In an intriguing study that poses a potential problem for our argument, Harris et al. (2002) reported negative effects on vibrotactile short-term memory when TMS was applied over contralateral primary somatosensory cortex shortly (300 or 600 ms) after target stimulus offset, but not late (900 or 1200 ms) after stimulus offset, which could potentially be interpreted as evidence for STM storage in SI. Harris et al. suggested that the difference between this finding

and Romo et al.'s (1999; Romo & Salinas, 2003) findings on PFC storage could be due to the lack of training Harris et al.'s subjects received compared to Romo et al.'s non-human subjects. This seems unlikely, given Spitzer et al.'s (2010; Spitzer & Blankenburg, 2011, 2012) findings of PFC storage in untrained human subjects. We have recently provided an alternate view: TMS over early sensory cortex increases neural "noise", which then feeds forward into storage systems in PFC; computational modeling of this hypothesis has provided results consistent with Harris et al.'s (2002) results (Bancroft, Hogeveen, Hockley, & Servos, under review).

Serences et al. (2009) reported finding information storage in early visual cortex when the salient stimulus property was the orientation of a Gabor stimulus. As orientation (for example, in degrees from some baseline) can be represented as a scalar quantity, we might expect to find PFC storage, rather than sensory storage. However, it is possible that task parameters made it difficult for subjects to encode a scalar representation of the salient stimulus property - the stimulus was presented for 1 second, but flickered on and off at a rate of 5 Hz. A similar study by Ester et al. (2013) also flickered their grating on and off. It is possible that subjects were unable to extract a reliable scalar representation of stimulus orientation, and as such, the stimulus representation in visual cortex was maintained. It is also possible that information maintenance in visual cortex reflects the retention of spatial information - the visual displays used by Serences et al. (2009) and Ester et al. (2013) include information other than orientation (for example, colour, stimulus size, etc.). As was reported by Christophel et al. (2012), areas in parietal cortex involved in visual attention also demonstrated information maintenance. Finally, neither Serences et al. (2009) nor Ester et al. (2013) reported whole-brain analyses, but rather focused on visual cortex; as such, it is possible that a scalar representation



did exist in PFC, but was not identified during analysis. While these studies pose a potential problem for our argument, we would argue that instead, they are evidence for the influence of task parameters and demands on neural coding.

The notion that sensory (or other task-relevant) cortex can be responsible for WM/STM maintenance is well-supported by recent research. However, it is clearly not the only possible storage substrate.

### **Discussion**

How can we account for the prefrontal encoding of representations of some types of stimuli (scalar stimuli; novel forms of stimuli, etc.), but encoding in task-relevant cortex of many other forms of stimuli? Consider: Information about visual flicker rate is stored in prefrontal cortex (Spitzer & Blankenburg, 2012), but a representation of abstract, “swirled” stimuli is stored in early visual cortex (Christophel et al., 2012), and working memory for faces and bodies appears to rely on higher visual areas (Linden et al., 2012). In the tactile modality, vibrotactile frequency information is found in prefrontal cortex (Romo & Salinas, 2003; Spitzer et al., 2010), but information about surface texture or the orientation of striations on an object appears to be maintained in primary somatosensory cortex (Zhou & Fuster, 1996). Similar results can be found in other sensory modalities (Postle, 2006; Spitzer & Blankenburg, 2012; Spitzer et al., 2014).

Can such a disparate set of results be explained by fractionation of WM/STM into multiple systems, specialized for a specific type of stimulus or task? While this approach has been applied to memory in the past (e.g., Wilson, O Scalaidhe, & Goldman-Rakic, 1993), Postle (2006) points out that the increasing body of literature on memory for different types of stimuli suggests that we would need to fractionate WM/STM into potentially hundreds of different

memory systems - hardly an elegant approach. It is much more parsimonious to argue that WM/STM relies on task-relevant sensory or cognitive cortex for storage, in combination with domain-general executive/attentional systems to help maintain information in task-relevant cortex, as well as a flexible, adaptable system in PFC to represent simple or novel types of stimuli.

There is a pattern in these findings, and we suggest that storage location can be determined (at least partially) by the complexity or dimensionality of stimuli. Simple, low-dimensional stimuli (in other words, those that are easily represented on a neural level) tend to be represented in PFC, while more complex stimuli tend to be represented in task-relevant cortex. The more complex a stimulus is, the more likely it is to require specialized neural circuitry to be processed and stored - neural circuitry such as is found in relevant sensory or language cortex. In contrast, simple stimuli (such as scalar representations of stimulus frequency) are less likely to require the specialized hardware of sensory cortex, and can be found in PFC.

This raises an important question. If information can be stored in task-relevant cortex, why have the ability to store information in PFC? Why not rely solely on sensory cortex, language cortex, et cetera? (At the very least, for information that can already be represented in those regions, if not for novel forms of stimuli.) Co-opting sensory (or other task-relevant) cortex for memory storage can interfere with the concurrent processing of sensory stimuli, and may also leave the contents of memory particularly vulnerable to interference. Indeed, working memory span tasks (one of the most common classes of WM task; see Conway et al., 2005, for a review) rely on a concurrent processing task during memory maintenance to increase difficulty and force the recruitment of executive/attentional resources. A recent study by Anderson et al. (2013; also

see Ester et al., 2013) reported that the accuracy of population tuning curves in visual cortex was affected by the size of a set of visual stimuli to be attended; it appears likely that similar effects could be found for neural populations when stimuli are being held in memory, rather than simply attended. Such an effect was found by Konstantinou and Lavie (2013), who found decreased perceptual sensitivity when increasing visual STM load, but not when increasing demand on executive processes, independent of STM load. The ability to offload certain types of information onto PFC can assist with minimizing the impact of memory maintenance on concurrent tasks, as well as helping to protect the contents of memory from interference caused by concurrent processing.

The neural organization of working and short-term memory appears to be even more fluid and dynamic than Postle (2006) suggested. We suggest that the neural systems recruited for memory maintenance may vary based not only on the sensory modality of stimuli, but also on the complexity of stimuli, and on the task-relevant aspects of stimuli. This differs from both the emergent-property model (Postle, 2006), and the traditional multiple-components model (Baddeley, 2010). For example, the emergent-property model is explicitly based on the notion that PFC does not serve a mnemonic function, but rather is involved in WM/STM through executive and attentional functions. However, this is inconsistent with the growing body of recent experimental evidence that demonstrates WM/STM storage in frontal cortex, and sometimes *only* in frontal cortex (e.g., Freedman et al., 2001; Sakurai et al., 2004; Lemus, Hernández, & Romo, 2009b; Spitzer et al., 2010; Spitzer & Blankenburg, 2011, 2012). In contrast, the multiple-components model suggests that WM/STM relies on dedicated working memory systems, but the neuroscience literature strongly argues against this view.

The neuroscientific research we have described in the present paper provides strong support for recent psychological theory on WM/STM. Recent psychological approaches to WM/STM have suggested domain-specific storage and rehearsal processes, and domain-general executive and attentional processes (Engle et al., 1999; Conway et al., 2005). The notion of domain-specific storage and rehearsal processes makes perfect sense, in light of the role often played by sensory (or other task-relevant) cortex in WM/STM storage. Further, the reliance of both memory and sensory/cognitive processing on the same cortical systems provides a neural basis for the inverse relationship between memory storage and concurrent processing (Barrouillet, Portrat, & Camos, 2011). At the same time, recent research findings (e.g., Christophel et al., 2012) have blurred the line between memory storage processes and attentional/executive processes. As we improve our understanding of how and why stimulus information is encoded in attentional/executive systems, psychological theory will likely have to adapt.

Further, the growing evidence that WM/STM relies on storage in sensory/task-relevant cortex casts serious doubt on the generality of models that suggest that WM is actually the activated contents of long-term memory (Ruchkin et al., 2003; Cowan, 1999). If the same sensory and cognitive systems that process stimuli are also responsible for their storage in WM/STM, what role is there for long-term memory? In addition, various studies (e.g., Freedman et al., 2001; Spitzer et al., 2010; Spitzer & Blankenburg, 2011, 2012; Christophel et al., 2012) have addressed WM/STM for stimulus types that we would not expect to be represented in long-term memory, as well as demonstrating that they are encoded in regions of PFC or sensory cortex. It is possible, of course, that long-term memory is involved in WM/STM for stimulus

types that are represented in long-term memory. However, given that WM/STM can be demonstrated for many types of stimuli that would not have pre-existing representations in long-term memory, models that equate WM with activated LTM are incompatible with a non-trivial portion of the literature. Indeed, Fuster (2003) differentiated between a dynamic basis for cognitive functions (the coordinated patterns of activity of cortical neurons) and a structural basis for cognitive functions (information stored in LTM about prior experiences with tasks, stimuli, etc.). The present discussion does not rule out a role for systems involved in long-term memory - if it is more efficient or simpler for WM/STM to store a “pointer” to relevant information in long-term memory, rather than storing the information itself, then doing so is consistent with current frameworks, such as some form of Postle’s (2006) emergent-property model. However, when dealing with novel or generic stimuli (i.e., vibrational frequency), we would not necessarily expect to find representations in LTM.

This raises the interesting possibility that some extant models of WM/STM could be considered as special cases of a more general model. Recent treatments of the multiple-components model have identified model components with task-relevant neural systems; the neural systems so identified are often the systems we might expect to be involved in a task based on the emergent-properties model (for example, attributing the phonological output buffer to Broca’s area) (Baddeley, 2010).

The present state of the theoretical literature raises some intriguing unanswered questions, and building a complete understanding of the neural systems underlying WM/STM will require the integration of information from all levels of analysis, from single-cell electrophysiology to functional neuroimaging. Recent studies (such as those of Spitzer et al.,

2010, and Spitzer and Blankenburg, 2012, and computational studies such as Bancroft et al., 2013) have made significant progress in doing so, but enormous future efforts will be required. The recent application of new methods of analyzing neural data (e.g., Christophel et al., 2012) also represents a tremendous advance in our understanding of the neural correlates of WM/STM, and it seems likely that new and important findings will come out of further use of advanced analysis methods.

We do not suggest that stimulus complexity is the sole determinant of storage substrate, nor do we suggest that PFC only stores simple stimuli, and sensory/task-relevant cortex only stores complex stimuli. Indeed, there have been reports that some tasks may recruit both sensory and prefrontal cortex to maintain different aspects of stimuli; Raabe et al. (2013) found coding for stimulus location in both the visual pathway and the frontal eye fields during a memory task that involved retention of spatial information. Analyzing local field potentials and spiking, Siegel et al. (2009) found phase-dependent PFC encoding for stimuli when multiple visual stimuli were retained. In addition, parietal encoding of stimulus information (in addition to encoding in visual cortex) has been reported by Christophel et al. (2012); this possibly represents the activity of attentional systems that support maintenance in sensory cortex. However, there seems to be a fairly clear differentiation between prefrontal encoding for simple (and occasionally novel) stimuli, and sensory encoding for complex stimuli for which pre-existing representations can be found in sensory and task-relevant cortex.

The ability of PFC to adapt to represent novel forms of information explains why subjects are capable of WM/STM for stimulus types that are novel or have limited ecological validity (e.g., vibrotactile stimuli, artificially-generated animal images). Given the increasing literature

on non-traditional WM/STM tasks, a flexible, adaptive storage system will probably become a required component of cognitive models of WM/STM. A growing body of research has suggested that PFC can adapt to represent and store novel forms of information (see Duncan, 2001, and Duncan & Miller, 2002, for reviews). Research into adaptive coding has demonstrated that prefrontal neurons can adapt to encode various types of information even after relatively short training periods. For example, Freedman et al. (2001) recorded from monkey PFC neurons during a delayed match-to-category task. Subjects were presented with two images, separated by a delay, that were created by morphing images of cats and dogs together. As such, each image was either more “cat-like” or more “dog-like”. Monkeys were trained to report whether the images were of the same category or different categories. After training, neurons in lateral PFC adapted to represent the categorization of the target image during the delay period, and that roughly equal numbers of neurons represented a cat-like or a dog-like categorization. As the monkeys were naive to images of cats and dogs, it appears unlikely that they had a preexisting code in PFC for representing cats and dogs. When the task was changed so that subjects had to fit stimuli into one of three different categories, PFC neurons adapted to encode the new categorical possibilities. In a recent study by Lee et al. (2013), subjects were presented with visual images and had to maintain the image for either a visual comparison task or a category comparison task. Using fMRI decoding methods, they identified stimulus representations in sensory cortex when subjects were performing the visual comparison task, but in PFC when subjects were performing the category comparison task. As well as being consistent with other work on PFC encoding of category information (e.g., Freedman et al., 2001), it serves as a clear demonstration that the salient (i.e., task-relevant) aspect(s) of a stimulus determines the storage

system recruited.

Duncan and Miller (2002) proposed an adaptive coding model of prefrontal cortex, in which prefrontal neurons are programmable or adaptable based on behavioural goals, and can serve mnemonic, attentional, and/or executive functions. While Duncan and Miller point out that this is only a framework for future research, the existence of a prefrontal system that can adapt to encode different types of information is consistent with the current experimental literature. Research into adaptive coding has answered the question of how subjects can store novel or unfamiliar forms of information for which there may not be pre-existing neural representations, a property of PFC that may or may not interact with the seeming differentiation between memory storage in PFC and memory storage in task-relevant cortex.

The complexity and breadth of the literature has made it increasingly difficult to construct models of WM/STM based purely on neuroscientific or purely on behavioural data, and any new model must be able to fit both sides of the literature. It is time for a fundamental change in our understanding of working and short-term memory. Rather than relying on dedicated neural systems, it relies on the coordinated activity of systems that have primary purposes other than memory. Rather than being able to point to a single cortical region as the storage substrate of memory, maintenance relies on the functional integration of anatomically-distinct systems, and the system(s) recruited for memory storage depend as much on the complexity of a stimulus as it does the sensory modality. Cortical systems can even rapidly adapt to represent novel stimuli.

One of the most important (even shocking) implications of the research outlined above (e.g., Serences et al., 2009; Riggall & Postle, 2012) is that even the interpretation of



neuroimaging data must be reconsidered, in light of results showing memory maintenance in regions that do not display increased delay-period activity. Rather, decoding methods may be necessary to accurately interpret neuroimaging data. Indeed, various recent fMRI studies (including some of those discussed in this paper) show inconsistency between activation found using standard analysis techniques, and information found using multivariate techniques. By extension, this suggests that theoretical treatments of working/short-term memory based on neuroimaging data should also be reconsidered. This reassessment of the literature, of course, depends on the validity of relatively novel decoding methods for fMRI and EEG/ERP data. It is possible that these methods do not fully identify stimulus representations stored in the brain, and that the studies reviewed above are only revealing part of the picture. However, given the demonstrated ability of these methods, and converging findings from various methodologies (fMRI, EEG/MEG, cellular recordings, etc.), we feel justified in suggesting that the WM/STM literature requires reassessment with an open mind.

Advances in empirical research are regularly revealing new and interesting aspects of working and short-term memory, and theoretical treatments of WM/STM have been slow to adapt. Concerted effort towards incorporating recent experimental results into theory is needed, and doing so offers enormous and exciting potential for our understanding of the relationship between cognitive processes and neural systems. In particular, it would be beneficial to develop or select a rigorous metric for stimulus complexity or dimensionality and formally assess the relationship between the complexity of a stimulus and the neural systems recruited for storage. One future direction may involve extending the work of Lee et al. (2013). Lee et al. found PFC storage of stimulus information when categorical information was retained, and sensory

encoding when visual object information was retained. Critically, when visual information was retained, subjects were asked to decide if a visual fragment belonged to the original object. This raises an interesting question: If we varied the number of relevant visual features, would the location of storage change? In other words, if there were few relevant visual features, would we see PFC storage rather than sensory storage (as we find when all features of the object are potentially relevant)?

The advent of techniques for decoding the contents of memory using techniques such as fMRI and EEG has rendered the already-complex literature on working and short-term memory even more convoluted. Application of those techniques, however, has helped reveal an apparent role for information complexity in determining the neural systems that are recruited for memory storage.

### **Acknowledgements**

We thank Britt Anderson and several anonymous reviewers for their insightful comments on this manuscript. The present research was supported by grants from the Natural Sciences and Engineering Research Council to WEH and PS, and an Ontario Graduate Scholarship to TDB.

### **Footnotes**

1. Note that we believe that working memory and short-term memory are necessarily coupled phenomena. It is increasingly accepted that working memory and short-term memory share a common memory system, and that working memory tasks and short-term memory tasks primarily differ in the degree to which they recruit executive and attentional systems. See Engle et al. (1999) for an example.

### **Chapter 3: TMS-induced neural noise in sensory cortex interferes with short-term memory storage in prefrontal cortex**

Tyler D. Bancroft, Jeremy Hogeveen, Philip Servos, and William E. Hockley

#### **Abstract**

In a previous study, Harris, Miniussi, Harris, and Diamond (2002) found disruption of vibrotactile short-term memory after applying single-pulse transcranial magnetic stimulation to primary somatosensory cortex (SI) early in the maintenance period, and suggested that this demonstrated a role for SI in vibrotactile memory storage. While such a role is compatible with recent suggestions that sensory cortex is the storage substrate for working memory, it stands in contrast to a relatively large body of evidence from human EEG and single-cell recording in primates that instead points to prefrontal cortex as the storage substrate for vibrotactile memory. In the present study, we use computational methods to demonstrate how Harris et al.'s results can be reproduced by TMS-induced activity in sensory cortex and subsequent feedforward interference with memory traces stored in prefrontal cortex, thereby reconciling discordant findings in the tactile memory literature.

#### **Introduction**

Vibrotactile short-term memory (often referred to as vibrotactile working memory (VWM)) is a powerful paradigm for studying the behavioural and neural correlates of working and short-term memory (Bancroft, Hockley, & Servos, 2011). VWM tasks usually involve presenting subjects with two vibrational stimuli delivered to the hand (the target and the probe), separated by an unfilled delay period, and instruct subjects to report whether the two stimuli are of same or different frequencies, or whether the probe is of a higher or lower frequency than the

target. Notably, the salient stimulus feature (vibrational frequency) can be represented as a scalar value, and the firing rates of neurons encoding vibrotactile stimuli tend to be monotonic functions of stimulus frequency (Romo, Brody, Hernández, & Lemus, 1999; Romo & Salinas, 2003). This makes vibrotactile memory a useful paradigm for integrating research results across various research methodologies, and recent studies have taken advantage of this property by demonstrating that it is possible to decode the stimulus frequency held in memory from beta-band EEG activity in frontal cortex (Spitzer, Wacker, & Blankenburg, 2010; Spitzer & Blankenburg, 2011, 2012). Intriguingly, recent research has suggested that vibrotactile memory may be one of a family of scalar short-term memory tasks, including auditory memory for pure tones and memory for the frequency of visual flicker (Spitzer & Blankenburg, 2012), as well as stimulus amplitude and duration (Spitzer, Gloel, Schmidt, & Blankenburg, 2014), that appear to share a similar, supramodal neural code in both sensory cortex and higher cortical regions, and rely on the same region of prefrontal cortex as a storage substrate.

An intriguing study, however, poses a challenge to this interpretation of results. Harris et al. (2002) presented subjects with two vibrotactile stimuli, separated by an unfilled delay period, and asked them to compare the stimuli. During the delay period, they applied single-pulse transcranial magnetic stimulation (TMS) to primary somatosensory cortex (SI). This study employed a ‘virtual lesion’ design, in which TMS-induced changes in behaviour suggest a causal relationship between peri-stimulation neural activity and task-related perceptual and cognitive functions (Robertson, Théoret, & Pascual-Leone, 2003). Harris et al. (2002) found a significant decrease in performance when the TMS pulse was applied to contralateral SI (relative to ipsilateral SI) 300 or 600 ms into a 1500 ms delay period, but not when it was applied 900 or

1200 ms into the delay period. (Note that while the decrease in performance in response to the 900 ms onset TMS pulse did not reach statistical significance ( $p = .16$ ), a trend is visible.) In contrast, TMS to ipsilateral SI did not significantly reduce performance. Harris et al. suggested that contralateral SI acts as a memory storage system for vibrotactile working memory. Such a notion is consistent with a previous single-cell recording study that reports SI encoding of complex tactile stimuli (Zhou & Fuster, 1996).

However, this notion conflicts with recent findings from human EEG studies and single-cell recording in non-human primates. Various studies by the research group of Romo et al. have suggested that regions in prefrontal cortex are the storage substrate used during VWM tasks and that no representation of the stored stimulus persists across the delay period in SI (see Romo & Salinas, 2003, for a review), and recent EEG studies by Spitzer and colleagues have reported being able to decode the frequency of a stored vibrational stimulus from prefrontal beta-band activity during the delay period of VWM (and other scalar STM) tasks (Spitzer & Blankenburg, 2011, 2012; Spitzer et al., 2010, 2014). The apparent incompatibility of these findings and those of Harris et al. (2002) raises questions about the scalar memory interpretation of results from VWM research, and also about whether the neural systems underlying VWM differ between humans and non-human primates.

The location of VWM storage has important implications for working and short-term memory theory, and the factors that determine storage location are unresolved. Postle (2006) suggested that stimuli tend to be stored in relevant regions of cortex that have pre-existing representations of that type of stimulus, such as sensory cortex; in order to account for recent experimental findings (including those around vibrotactile memory), we have recently suggested

that less complex stimuli with simple neural codes instead tend to be stored in prefrontal cortex (Bancroft, Hockley, & Servos, 2014). As this theoretical framework is partly based on research showing prefrontal storage of scalar stimuli, reconciling Harris et al.'s (2002) results with other findings (i.e., Romo & Salinas, 2003; Spitzer et al., 2010) has theoretical importance.

We offer an alternative interpretation of Harris et al.'s (2002) findings. According to the former view, the application of TMS suppressed neural activity within SI during the delay period, and the consequent impact on VWM performance can be interpreted as evidence that SI is involved in VWM storage. However, beyond local changes in cortical activity, TMS can induce distal effects at brain regions receiving feedforward inputs from the targeted brain region (e.g., Paus, Jech, Thompson, Comeau, Peters, & Evans, 1997). Rather than SI being a storage medium for vibrotactile memory, we suggest that the application of TMS induces or increases activity in sensory cortex (both in SI and in secondary somatosensory cortex (SII), via feedforward connections), and that this activity then interferes with VWM storage in PFC.

It has been established that TMS can induce neural activity when applied to some areas of sensory cortex, including somatosensory cortex (Sugishita & Takayama, 1993; Ray, Meador, Epstein, Loring, & Day, 1998; Steward, Walsh, & Rothwell, 2001; Ptito et al., 2008). As well, recent behavioural and computational studies have suggested that when irrelevant vibrotactile stimuli are presented during the maintenance period of a VWM task, they reduce performance by being encoded into memory (Bancroft & Servos, 2011; Bancroft, Servos, & Hockley, 2011; Bancroft, Hockley, & Servos, 2012, 2013). As there is a direct mapping between induced activity in primary somatosensory cortex and the frequency of the stimulus perception created by that induced activity (e.g., Romo et al., 1998), it follows that increased activity in SI due to TMS

could have similar effects to irrelevant somatosensory stimuli.

Perhaps most compellingly, somatosensory memory studies that have used TMS to increase activity in the middle frontal gyrus (a region of prefrontal cortex known to inhibit activity in primary somatosensory cortex) have reported decreased response times when TMS was applied early (300 ms onset) but not late (1200 ms onset) in the delay period, suggesting a decrease in interference (Hannula et al., 2010; also see Savolainen et al., 2011). Given that these TMS manipulations, known to suppress activity in SI, have been shown to *improve*, not reduce, performance on tactile memory tasks, it raises an interesting question: Is Harris et al.'s (2002) manipulation suppressing activity in SI, or is it producing excitatory or facilitatory effects that impact storage systems further downstream?

In the present study, we adapted a computational model of prefrontal cortex (Miller & Wang, 2006) in order to demonstrate that Harris et al.'s (2002) results can be produced by TMS-induced activity in sensory cortex, resulting in interference with information stored in prefrontal cortex. As pointed out by Miller and Wang, feeding noise into an integrator causes a decrease in performance proportional to the duration of noise. In the present study, the accumulation of noise in PFC leads to an inverse relationship between task performance and the delay between TMS offset and probe onset.

### **Model**

The model used in the present study was originally developed by Miller and Wang (2006) as a model of prefrontal neurons involved in vibrotactile working memory tasks. We have previously adapted it to model the interfering effects of distractor stimuli on vibrotactile working memory (Bancroft, Hockley, & Servos, 2013). It is a rate model, based on the interaction of

pairs of populations of prefrontal neurons. While the Miller and Wang model operates at a relatively high level of abstraction, it captures the fundamentally subtractive nature of the stimulus comparison process (Romo & Salinas, 2003), and has proven capable of fitting a variety of experimental data (e.g., Bancroft, Hockley, & Servos, 2013). In addition, the model can be fit to data with relatively few free parameters, which is beneficial when fitting a dataset with relatively few data points (such as the Harris et al. data we consider in this paper).

In this model, comparison ( $C$ ) populations receive input from sensory cortex and have excitatory outputs to populations of memory ( $M$ ) neurons. Memory populations have excitatory self-connections (allowing persistent activity in the absence of external input), and inhibitory connections to  $C$  populations. The equations governing the behaviour of the network are as follows:

$$\begin{aligned} dr_C/dt &= (1/\tau)(-r_C + w_{MC}r_M + w_{IC}I) \\ dr_M/dt &= (1/\tau)(-r_M + w_{MM}r_M + w_{CM}r_C) \end{aligned}$$

where  $r$  is the firing rate of a population,  $\tau$  is a time constant,  $w_{AB}$  represents the strength of a connection from a population  $A$  to another population  $B$ , and  $I$  is the input received from sensory cortex. The addition of  $w_{IC}$  to the model is intended as a potential scaling factor to allow presentation of stimulus frequencies outside of biologically-realistic firing rates (for example, auditory stimuli in the kHz range).

Note that if  $w_{MM}$  is set to 1, the  $M$  population becomes a perfect (i.e., lossless) integrator, and the equation governing behaviour of  $M$  populations can be reduced to:

$$dr_M/dt = (1/\tau)(w_{CM}r_C)$$

We have used this reduced equation in the present study.



Upon presentation of a target stimulus, a  $C$  population transmits the stimulus frequency to an  $M$  population. The  $M$  population then inhibits activity in the  $C$  population, driving the  $C$  firing rate back to baseline. The self-connection allows the  $M$  population to maintain its firing rate in the absence of external stimulation. Upon presentation of the probe stimulus to the  $C$  population, the combination of inhibitory input from the  $M$  population and excitatory input from sensory cortex results in the  $C$  population calculating some function of  $f_{target} - f_{probe}$ , consistent with experimental findings (Romo & Salinas, 2003), and also consistent with decision-making mechanisms used in abstract mathematical models of VWM (Bancroft, Hockley, & Servos, 2012). Note that experimental findings have reported finding neurons in sensory cortex that have firing rates that are positive monotonic functions of stimulus frequency, as well as neurons that have negative monotonic functions of stimulus frequency (Romo & Salinas, 2003). This plays an important role in the functioning of the model.  $C$  populations that receive positive monotonic input (we refer to these as  $C_+$  populations) will fire above baseline when the probe stimulus is a higher frequency than the target stimulus, while populations that receive negative monotonic input ( $C_-$  populations) will act as detectors for lower-frequency probes.

We have also added decision ( $D$ ) populations to the model to facilitate decision-making. The  $D$  populations receive excitatory output from  $C$  populations during the presentation of probe stimuli:

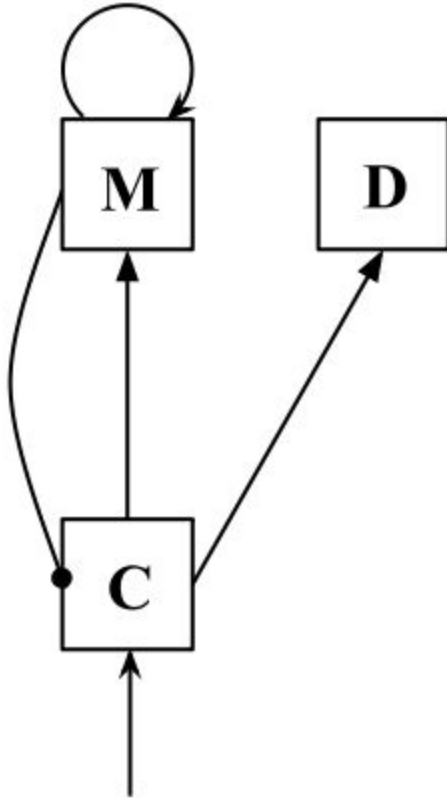
$$dr_D/dt = (1/\tau)(w_{CD}r_C)$$

During target presentation and the delay period,  $w_{CD}$  is set to 0, and only assumes a non-zero value upon presentation of the probe stimulus. During probe presentation, the  $D$  populations act as perfect integrators of the activity of the relevant  $C$  population; this allows a direct comparison

between the total activities of the  $C_+$  and  $C_-$  populations (and therefore the probe-higher and probe-lower detectors).

In the present study, we simulated two triplets of  $C/M/D$  populations (see Fig. 1), one receiving positive monotonic input (with subscript +), the other receiving negative monotonic input (with subscript -). The triplets were not connected to each other. To determine a simulated response, we compared the activity of the  $D_+$  and  $D_-$  populations shortly after probe offset. If activity in the  $D_+$  population exceeded that in the  $D_-$  population, it follows that overall activity in the  $C_+$  population exceeded that in the  $C_-$  population across the probe presentation period, and we recorded a probe-higher response. If activity in the  $D_-$  population exceeded that of the  $D_+$  population, we recorded a probe-lower response, and if activity in the two populations was equal, a response was randomly chosen.

**Figure 1: Diagram of a  $C/M/D$  triplet**



During the delay period, the model received constant input, with input values drawn from an exponential distribution with the distribution parameter  $\lambda$ , inversely proportional to the mean and variance of the distribution. This noisy input represents ongoing, baseline activity in sensory regions. Critically, we modeled the application of TMS to sensory cortex by allowing  $\lambda$  to vary as a free parameter. If TMS increases activity in sensory cortex, we would expect the magnitude of the noise to increase (and therefore the value of  $\lambda$  to decrease). Further, allowing values of  $\lambda$  to vary separately for ipsilateral and contralateral stimulation allows us to test for differing effects of inhibition depending on laterality - if ipsilateral SI is more greatly inhibited than

contralateral SI, we would expect a smaller magnitude of interference (and therefore a greater value for  $\lambda$ ). The exponential distribution was chosen for this study as it has one parameter that determines both the mean and the variance of the distribution.

### Simulation Methodology

In the present study, input to PFC was of two types. During target and probe presentation,  $C_+$  populations received input equal to  $w_{IC}f$ , and  $C_-$  populations received input equal to  $w_{IC}(40-f)$ , where  $f$  was the frequency (in Hz) of the stimulus, and  $w_{IC}$  was the strength of the connection from sensory cortex to prefrontal cortex. Consistent with previous work (Bancroft, Hockley, & Servos, 2013), stimulus frequency ( $f$ ) was drawn from a Gaussian distribution with a mean equal to that of the presented stimulus, and standard deviation ( $\sigma$ ) allowed to vary as a free parameter, in order to account for inaccuracy in the neural signal introduced during neural transmission and processing. Firing rates ( $r_{population}$ ) were not allowed to decrease below zero. Other parameter values are presented in Table 1.

**Table 1: Simulation parameters**

Parameter	Value
$f_{target}$	20 Hz
$f_{probe, higher}$	22 Hz
$f_{probe, lower}$	18 Hz
Stimulus duration	1000 ms
Delay period duration	1500 ms
$\tau$	10

$w_{IC}$	0.4
$w_{CM}$	0.4
$w_{MC}$	-0.4
$w_{CD}$	0 (during target presentation/delay periods); 0.5 (during probe presentation)
$r_{minimum}$	0

During the delay period,  $C$  populations received noisy input drawn from an exponential distribution at each integration timestep, with the distribution parameter  $\lambda$  set as a free parameter. The parameter  $\lambda$  determines the mean ( $1/\lambda$ ) and variance ( $1/\lambda^2$ ) of an exponential distribution.

Harris et al. (2002, Exp. 2) presented subjects with two 1000 ms vibrotactile stimuli (the target and probe), separated by a 1500 ms delay period. TMS was applied to either ipsilateral or contralateral primary somatosensory cortex, at an onset of either 300, 600, 900, or 1200 ms into the delay period. The target and probe stimuli differed by  $\pm 2$  Hz, and subjects were required to report whether the probe was of a higher or lower frequency than the target.

To simulate the effects of TMS,  $\lambda$  was allowed to assume two values during the delay period: The initial value ( $\lambda_{baseline}$ ), and a new value upon the application of TMS ( $\lambda_{TMS}$ ). Pilot studies were performed to estimate approximate parameter ranges (based on minimizing error between experimental and simulated results), after which the  $\sigma$  parameter was allowed to vary freely within the range [1.00, 3.00], with a stepsize of 0.5;  $\lambda_{baseline}$  was fixed at 0.5, and  $\lambda_{TMS}$  (*ipsilateral*) and  $\lambda_{TMS}$  (*contralateral*) were varied across the range of [0.5, 0.025], taking possible values of 0.5, 0.375, 0.25, 0.125, 0.1, 0.075, 0.05, and 0.025. Two thousand trials were simulated for each combination of onset time and free parameter values. Parameter fit was assessed by minimizing

the sum of squared error ( $SS$ ) between the experimental results from Harris et al. (2002) (rounded to four places) and simulated results, and the selected parameter values were those that minimized total  $SS$  across both ipsilateral and contralateral TMS conditions. (Note that parameter selection was constrained by requiring the value of  $\sigma$  to be the same for both ipsilateral and contralateral stimulation conditions.)

To improve the model fit, a second round of simulations was performed based on the best-fitting parameters from the first round of simulations ( $\sigma = 2.00$ ,  $\lambda_{TMS (ipsilateral)} = 0.375$ , and  $\lambda_{TMS (contralateral)} = 0.125$ ). The value of  $\sigma$  was set to 2.00, and  $\lambda_{TMS (ipsilateral)}$  and  $\lambda_{TMS (contralateral)}$  varied within the ranges [0.425, 0.325] and [0.175, 0.075], respectively, with a stepsize of 0.25.

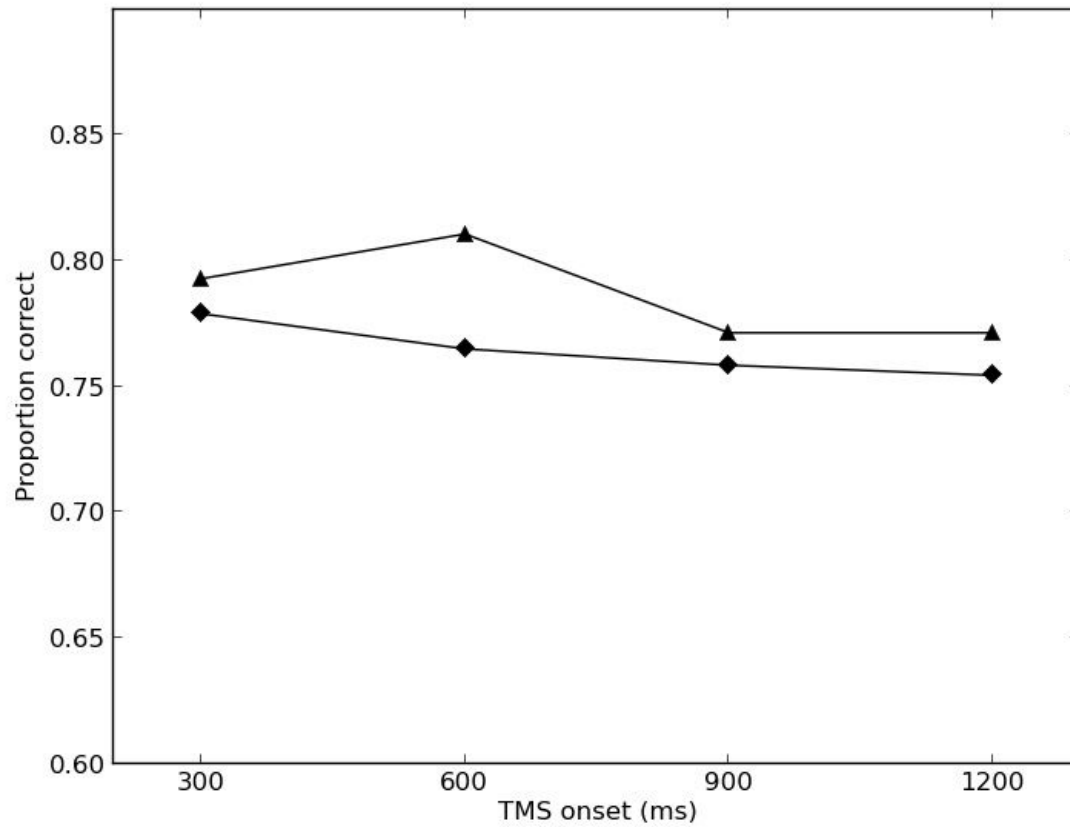
Simulations were performed with code written in Python, with the NumPy and standard Python *random* libraries (specifically, *random.expovariate* for the generation of noisy input). Integration was performed using a 4th-order Runge-Kutta, with an integration stepsize of 0.5.

### Results and Discussion

The results of the final round of simulations are presented in Figures 2 and 3. The best-fitting parameter values were found to be  $\sigma = 2.0$ ,  $\lambda_{TMS (ipsilateral)} = 0.350$ , and  $\lambda_{TMS (contralateral)} = 0.150$ . The sum of squared error ( $SS$ ) for the best fit was found to be 0.00446 (0.00273 for the ipsilateral condition, and 0.00173 for the contralateral condition), and the variance explained by the model ( $r^2$ ) was calculated to be 0.780.

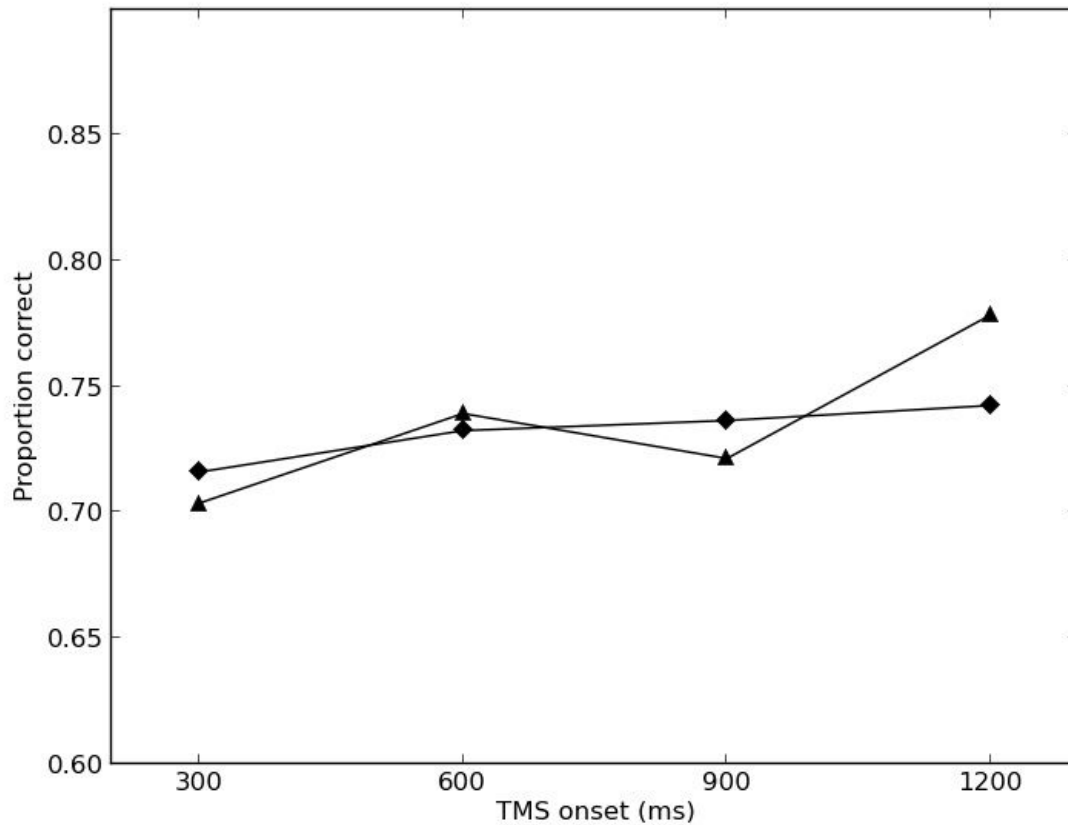
#### Figure 2: Simulated and empirical results of TMS to ipsilateral SI

**Triangles denote results from Harris et al. (2002), Exp. 2; diamonds denote simulated results.**



**Figure 3: Simulated and empirical results of TMS to contralateral SI**

**Triangles denote results from Harris et al. (2002), Exp. 2; diamonds denote simulated results.**



Model performance was largely robust against changes in parameter values, with maximum overall *SS* of 0.0962 in the final round of simulations (0.00485 for the ipsilateral condition, and 0.0914 for the contralateral condition).

The results of the present simulation suggest that Harris et al.'s (2002) results can be replicated by assuming that TMS increases activity in sensory cortex, which then interferes with the contents of memory, held in PFC. This interpretation is consistent with the single-cell electrophysiology and EEG literatures (e.g., Romo & Salinas, 2003; Spitzer et al., 2010, 2014; Spitzer & Blankenburg, 2011, 2012), and requires no need to suggest that SI is involved in



vibrotactile memory storage.

One crucial part of Harris et al.'s argument was that TMS to SI ipsilateral to the hand receiving vibrotactile stimulation did not produce effects on task performance. They suggested that if VWM storage relied (at least in part) on areas further downstream, such as SII (which possesses bilateral receptive fields), TMS to ipsilateral cortex would produce similar effects to TMS to contralateral cortex. However, recent EEG and MEG studies of tactile memory have reported greater alpha-band activity over ipsilateral SI than over contralateral SI (Haegens, Osipova, Oostenveld, & Jensen; 2010; Spitzer & Blankenburg, 2011; Haegens, Luther, & Jensen, 2012). Further, when irrelevant stimuli were expected to be presented to the opposite hand, pre-stimulus alpha power in cortex varied with laterality (Haegens et al., 2012). As alpha-band activity is believed to be linked to inhibitory activity (e.g., Rihs et al., 2007; Haegens, Nácher, Luna, Romo, & Jensen, 2011), Haegens et al. (2012) suggested that activity in ipsilateral SI could be suppressed in order to inhibit the processing of irrelevant sensory input. In this case, the failure to find effects of ipsilateral TMS does not necessarily reflect a reliance on contralateral SI for VWM storage, but rather may reflect differences in endogenous inhibitory activity between ipsilateral and contralateral sensory cortex.

The results of the present study have an impact reaching beyond the vibrotactile working memory literature. Postle (2006) introduced the emergent-property model of working memory, which suggests that working memory does not rely on a specialized neural system, but rather the interaction between neural systems that primarily serve other sensory, cognitive, or action-related functions. Indeed, Postle explicitly argues that PFC is not involved in the storage of information. For example, task-relevant sensory cortex has been suggested as the storage

medium for working memory, and recent neuroimaging studies that have applied novel methods for decoding the contents of sensory cortex have reported finding stimulus information in early visual cortex during the maintenance period of visual memory tasks (e.g., Serences, Ester, Vogel, & Awh, 2009; Christophel, Hebart, & Haynes, 2012). Other, similar results exist.

However, there is an increasing body of evidence that PFC is the storage substrate for simple stimuli and novel stimuli (e.g., Freedman, Riesenhuber, Poggio, & Miller, 2001; Bancroft et al., 2014). Perhaps most persuasive are recent studies that have reported decoding the contents of short-term or working memory from prefrontal beta-band activity, regardless of whether the stored aspect of the stimulus was delivered as a tactile vibration, auditory tone, visual flicker (Spitzer et al., 2010; Spitzer & Blankenburg, 2012), or stimulus intensity or duration (Spitzer et al., 2014). While the emergent-property model is compelling, in that it is simple, parsimonious and able to explain a wide variety of results from the literature, when combined with previous findings, the results of the present study suggest that vibrotactile working memory research may require an expansion of the emergent-property model. We have recently suggested that the complexity of a stimulus is at least a partial factor in determining what neural storage systems are recruited (Bancroft et al., 2014).

We acknowledge that the timecourse of the effects of TMS to SI are not well-understood. Indeed, the effects of TMS to SI are not well-understood in general. Harris et al. (2002) selected a target in SI by using TMS to identify the region of greatest tactile extinction, which could be interpreted as evidence of an *inhibitory*, rather than excitatory effect of TMS. However, other research has found excitatory or facilitatory effects of TMS over sensory and motor cortex (Gerwig, Kastrup, Meyer, & Niehaus, 2003; Ragert, Franzkowiak, Schwenkries, Tegenthoff, &

Dinse, 2008; Ragert, Becker, Tegenthoff, Pleger, & Dinse, 2004), and even combined excitatory and inhibitory effects (Oliveri et al., 2000; Moliadze, Zhao, Eysel, & Funke, 2003; Strafella, Vanderwerf, & Sadikot, 2004). Indeed, even inhibitory effects on neurons in a stimulated region can produce increased neural activity or excitability, due to a reduction in the activity of inhibitory interneurons. Further, Amassian et al. (2002) suggest that single-pulse TMS can excite a large number of neurons in sensory cortex without the effects reaching consciousness.

Effects are also likely to depend heavily on cortical structure and connectivity. Identical stimulation parameters can produce excitation or inhibition in different cortical regions (Paus, 2005), and there is growing evidence that the effects of TMS over a cortical region are state-dependent, with effects possibly depending on pre-existing activity in the region (Harris et al., 2008; Pasley, Allen, & Freeman, 2009; Abrahamyan, Clifford, Arabzadeh, & Harris, 2011). Recently, a number of authors (including Harris) have suggested that TMS resulting in what appears to be inhibitory behavioural effects can actually be due to increased neural excitability resulting in an unfavourable signal-to-noise ratio (Silvanto and Muggleton, 2008; Miniussi, Harris, & Ruzzoli, 2013).

Whether TMS-induced behavioural results are driven by cortical inhibition, an unfavourable neuronal signal-to-noise ratio, or both to some extent, the present work highlights another critical issue: the local vs. remote interpretation of the neural intervention. Combined TMS/fMRI studies have shown that, even at relatively low intensities, TMS modulates haemodynamic activity in both the targeted brain region and distant cortical and subcortical regions (e.g. Bohning et al., 1999; Bestmann, Baudewig, Siebner, Rotherwell, & Frahm, 2005). Though the distinction between local or remote effects of TMS may be inconsequential in some

settings, in the case of Harris and colleagues' (2002) findings, interpreting the effects of TMS as related to SI inhibition or PFC signal-to-noise modulation produce fundamentally different insights for neurocognitive models of STM storage. In such a case, it might be useful to, wherever possible, choose multiple stimulation sites (e.g. SI and PFC) and timings (e.g. early vs. late in the delay period at both sites) in order to design experiments that use TMS to conclusively elucidate the where and when of a given cognitive task in the brain.

It is likely that relatively limited activity in SI can produce effects downstream, given the feedforward nature of output connections from SI (Romo & Salinas, 2003). When discussing Harris et al.'s (2002) results, Romo and Salinas (2003) suggested that TMS was likely to produce localized effects in cortex for approximately 200 ms after application. This may be a conservative estimate; others have reported that the initial phase of increased neural activity can persist for approximately 500 ms after TMS application (Moliadze et al., 2003; Silvanto and Muggleton, 2008). However, given that the effects of activity in SII produced by vibrotactile stimuli can persist for several hundred milliseconds after stimulus offset (Romo & Salinas, 2003), as well as possible feedback loops within the vibrotactile memory system (e.g., Auksztulewicz, Spitzer, & Blankenburg, 2012), it appears plausible that effects of a TMS pulse in SI could produce longer-lasting effects downstream, and could produce substantial interference with VWM storage. Experimental tests of this hypothesis (possibly involving a combined TMS/ERP paradigm) should prove fruitful.

In the present study, we have suggested a way to integrate the intriguing TMS results of Harris et al. (2002) with the EEG and single-cell literatures. The present study also poses a challenge to the emergent-property model of working memory, and suggests a manner in which

that model may be extended. Finally, and perhaps most intriguingly, the results of the current study are consistent with growing evidence that short-term and working memory may rely on different neural storage substrates, based on the salient property of stimuli that is being maintained in memory (for example, we have shown that simple tactile stimuli are stored in PFC, but complex tactile stimuli may be stored in sensory cortex; Zhou & Fuster, 1996).

### **Acknowledgements**

We are indebted to Justin Harris for providing us with the means from Experiment 2 of Harris et al. (2002), as well as for granting permission to publish that data. We thank Britt Anderson, Diano Marrone, and two anonymous reviewers for their helpful comments. We also thank the Wilfrid Laurier University Library for their generous support for publication fees.

## **Chapter 4: A shared short-term memory system for stimulus duration and stimulus frequency**

Tyler D. Bancroft, William E. Hockley, and Philip Servos

### **Abstract**

Recent research has suggested the existence of a modality-independent memory system that is responsible for the storage of simple, scalar stimulus attributes, such as the frequency of an auditory pure tone or the duration of a stimulus. In the present study, we modify an existing computational model of short-term memory for stimulus frequency to allow it to perform short-term memory tasks for both stimulus frequency and stimulus duration, supporting the notion of a common scalar short-term memory system. We further demonstrate the utility of the model by showing that it can reproduce the subjective shortening effect - a classic finding in the psychophysical literature.

### **Introduction**

As yet, there is no general consensus on the cognitive and neural basis of working and short-term memory (WM/STM) in humans. Rather than a dedicated working memory “system”, recent reviews of the literature have suggested WM/STM storage may, in fact, rely on neural systems that also perform sensory processing (Pasternak & Greenlee, 2005; Postle, 2006). And, indeed, recent neuroimaging studies using advanced analysis methods would seem to support this notion (e.g., Christophel, Hebart, & Haynes, 2012; Lee, Kravitz, & Baker, 2013).

However, there is a class of WM/STM task that appears not to rely on task-relevant sensory cortex, but instead relies on a modality-independent storage system in frontal (prefrontal and/or premotor) cortex: Memory for scalar stimulus attributes (Bancroft, Hockley, & Servos, 2014). In these tasks, subjects are asked to remember a stimulus property that can be represented as a scalar (i.e., unidimensional) value: For example, the frequency of an auditory stimulus (Spitzer & Blankenburg, 2012), the frequency of a flickering LED (Spitzer & Blankenburg, 2012); the frequency of a vibration delivered to the hand (Bancroft & Servos, 2011; Bancroft, Hockley, & Servos, 2012; Bancroft, Servos, & Hockley, 2011; Brody, Hernández, Zainos, & Romo, 2003; Harris, Harris, & Diamond, 2001; Harris, Miniussi, Harris, & Diamond, 2002; Romo & Salinas, 2003; Romo, Brody, Hernández, & Lemus, 1999; Spitzer & Blankenburg, 2011; Spitzer, Wacker, & Blankenburg, 2010), the duration or intensity of a stimulus (Spitzer, Gloel, Schmidt, & Blankenburg, 2014) or the number of brief sensory pulses presented (Spitzer, Fleck, & Blankenburg, 2014). These tasks have been referred to as *parametric working memory*; we refer to them as *scalar short-term memory*, as it is not yet clear whether they are true working memory tasks, in the sense that WM tasks recruit both short-term memory storage systems and also executive/attentional resources (e.g., Engle, Tuholski, Laughlin, & Conway, 1999). As such, we prefer the less theoretically-loaded term. Further, there is evidence that non-parametric, categorical stimuli may rely on a similar storage system (e.g., Freedman, Riesenhuber, Poggio, &

Miller, 2001).

What common property of these tasks appears to underlie these similarities in neural representation? The answer lies in the ability to represent the salient stimulus property as a scalar value. Single-cell recording work in monkeys during STM for vibrotactile frequency has identified neurons in prefrontal and/or premotor cortex have firing rates that are monotonic functions of vibrotactile frequency, and further that many have firing rates that are linear or “soft” sigmoidal (i.e., can be treated as approximately linear) functions of frequency (e.g., Hernández, Zainos, & Romo, 2002; Romo et al., 1999; Brody et al., 2003). Similar findings have been reported for stimulus duration: Sakurai et al. (2004) reported categorical encoding of stimulus duration by PFC neurons, and Durstewitz (2004) suggested that gradually increasing or decreasing activity in PFC neurons may serve a timing function. The most compelling evidence for a common scalar STM system comes from a recent series of EEG studies by Spitzer and Blankenburg, which have been able to decode the contents of memory based on beta-band activity over frontal cortex. Spitzer, Wacker, and Blankenburg (2010) were able to extract the frequency of a vibrotactile stimulus from beta-band activity over frontal cortex. Spitzer and Blankenburg (2012) found a similar representation for auditory and visual frequency. Most recently, Spitzer, Gloel, et al. (2014) also found a similar representation for stimulus duration and intensity, and Spitzer et al. (2014) found a similar representation for the number of stimuli in



a series of brief sensory stimuli. Collectively, such scalar short-term memory tasks appear to share common neural and behavioural properties (Bancroft et al., 2014).

It seems appropriate, then, to apply a model of a different scalar STM task (memory for stimulus frequency) to memory for stimulus duration. Miller and Wang (2006) developed such a model. Although it does not fully reproduce the complex, heterogenous patterns of neuronal activity in PFC (e.g., Jun et al., 2010), it does incorporate the subtractive decision-making process that lies at the heart of scalar STM (e.g., Romo & Salinas, 2003; Jun et al., 2010), and we have successfully used it to reproduce experimental results on capacity, the effects of TMS, and interference in STM for vibrotactile frequency (Bancroft, Hockley, & Servos, 2013; Bancroft, Hogeveen, Hockley, & Servos, 2014). Further, ongoing research in our group has found that relatively minor changes to model parameters can substantially improve the model's ability to reproduce experimentally-observed neuronal activity, such as time-dependent coding of stimulus information during the delay period (Bancroft, Hockley, & Servos, unpublished data). In the present paper, we demonstrate how the model can be used to perform duration comparison tasks, and also produce a classic effect from the duration comparison literature. We simulated a duration comparison task in which subjects are presented with a to-be-remembered stimulus (the target), followed by a delay period, followed by a second stimulus (the probe), and asked to compare the durations of the stimuli.

## Model

We adapted Miller and Wang's (2006) model for use in STM for duration. The model simulates the firing rates of populations of frontal neurons, containing a comparison ( $C$ ) population and a memory ( $M$ ) population. We have previously also added a decision ( $D$ ) population (Bancroft et al., 2013; 2014). The  $C$  population receives excitatory external input and has an excitatory output to the  $M$  population; the  $M$  population has an inhibitory output to the  $C$  population and an excitatory self-connection. The  $C$  population has an excitatory output to the  $D$  population that is activated upon probe stimulus presentation and inactive before this. During target presentation, the activity of the  $C$  population increases and drives activity in the  $M$  population. Upon target offset, leak and inhibitory input from the  $M$  population will drive  $C$  activity to baseline, while the self-connection will maintain activity in the  $M$  population during the delay period. On probe presentation, the combination of excitatory input and inhibition from the  $M$  population will result in the  $C$  population calculating a subtraction between the *frequencies* of the input and stored stimuli. In the past, we have then made the same/different or higher/lower frequency comparison by examining the activity of the  $D$  population(s). In the case of duration comparison, although we hold frequency constant (as is common in studies of subjective shortening; e.g., Wearden & Ferrara, 1993), we can still model a longer/same/shorter decision based on  $D$  activity.

Experimental findings have suggested that many neurons involved in scalar STM have firing rates that are linear or approximately linear functions of the salient stimulus property, such as frequency (e.g., Romo et al., 1999). The equations governing firing rates are as follows:

$$dr_C/dt = (1/\tau)(-r_C + w_{MC}r_M + w_{IC}r_I)$$

$$dr_M/dt = (1/\tau)(-r_M + w_{MM}r_M + w_{CM}r_C)$$

$$dr_D/dt = (1/\tau)(w_{CD}r_C)$$

where  $w_{AB}$  is a connection from population  $A$  to population  $B$ ,  $r_A$  is the firing rate of population  $A$ ,  $\tau$  is a time constant, and where  $r_{\text{initial}} = 0$  for all populations.

Note that the  $M$  population can be instantiated as a perfect integrator if the strength of the self-connection ( $w_{MM}$ ) is set to 1 (Miller & Wang, 2006). In this case, the firing rate of the  $M$  population changes according to the following equation:

$$dr_M/dt = (1/\tau)(w_{CM}r_C)$$

However, if the value is not fine-tuned to 1 and  $w_{MM} < 1$ , the  $M$  population will “leak” information, and we can recast the equation in the following form:

$$dr_M/dt = (1/\tau)(r_M(w_{MM} - 1) + w_{CM}r_C)$$

This appears to be consistent with observed activity: Overall, PFC maintains a representation of the stimulus across the delay period, but not all individual neurons maintain the representation across the entire period. As such, the representation in PFC may decay (e.g., Romo et al., 1999;

Barak, Tsodyks, & Romo, 2010). The difference in activity between instantiating the  $M$  population as a perfect integrator or a leaky integrator will become relevant later in this study.

In the past, we have set  $w_{CM}$  and  $w_{MC}$  to have equal, relatively large amplitudes (although given the inhibitory nature of the  $M$  to  $C$  condition, opposite signs). In this case,  $r_M$  and  $r_C$  quickly reach stable values, with  $r_C$  being driven to zero (due to increasing inhibition from the  $M$  population), and  $dr_M/dt$  also reaching zero (assuming  $M$  is a perfect integrator). This prevents the model from being applied to memory for duration tasks, as  $r_M$  and  $r_C$  may reach stable values before stimulus offset. If, however, we set  $|w_{MC}|$  to be a relatively small value, it takes longer for inhibitory output from the  $M$  population to drive activity in the  $C$  population to zero, due to the relatively stronger external input to  $C$ . As such, the  $M$  population will continue to integrate input from the  $C$  population, and  $r_M$  will index the duration of the stimulus.

Upon probe presentation, the connection from  $C$  to  $D$  will be activated (i.e., set so  $w_{CD} > 0$ ), and the  $D$  population will compute a function of the differences between the duration of the target and the probe. The longer the target stimulus, the higher the eventual value of  $r_M$ , and as such, the greater the inhibition of the  $C$  population. Therefore, the activity of the  $D$  population will be a negative function of the duration of the target stimulus, and a positive function of the duration of the probe stimulus. A decision rule could then be applied. For example: If, on probe stimulus offset,  $r_D > \theta_{\text{longer}}$ , the model makes a “longer” response (i.e., reports that the probe

stimulus is longer than the target stimulus). If  $r_D < \theta_{\text{shorter}}$  the model makes a “shorter” response.

If  $\theta_{\text{longer}} > r_D > \theta_{\text{shorter}}$  the model makes a “same” response.

In Simulation 1, we demonstrate that the model can perform duration comparison tasks.

In Simulation 2, we demonstrate how the model can provide a neurobiological basis for the *subjective shortening effect*, an effect commonly found in duration comparison tasks (e.g., Spetch & Wilkie, 1983).

All simulations were run using code written in Python, and using the Euler method with an integration timestep of 0.5 ms.

### Simulation 1

In Simulation 1, we demonstrate the ability of the model to perform duration comparison computations. The value of  $w_{\text{MM}}$  was set to 1, and all other parameter values are presented in Table 2.

**Table 2: Simulation parameters**

Parameter	Value
$\tau$	10
stimulus frequency (target and probe)	400 Hz
$w_{\text{CD}}$	0.1
$w_{\text{CM}}$	0.1
$w_{\text{IC}}$	0.01
$w_{\text{MC}}$	-0.1

First, we presented the model with three different target stimuli, in the form of auditory pure tones: A 500 ms stimulus, a 750 ms stimulus, and a 1000 ms stimulus. (Subjective shortening experiments have often used auditory stimuli, but the model is compatible with other types of periodic stimulus, such as vibrotactile stimuli.) Figure 4A displays  $r_M$  across the target presentation period and the delay period for these three stimuli. As expected, the value of  $r_M$  upon target stimulus offset is a function of stimulus duration.

**Figure 4A: Value of  $r_M$  during and after target presentation; Solid line indicates 500 ms target, dashed line indicates 750 ms target, dotted line indicates 1000 ms target.**

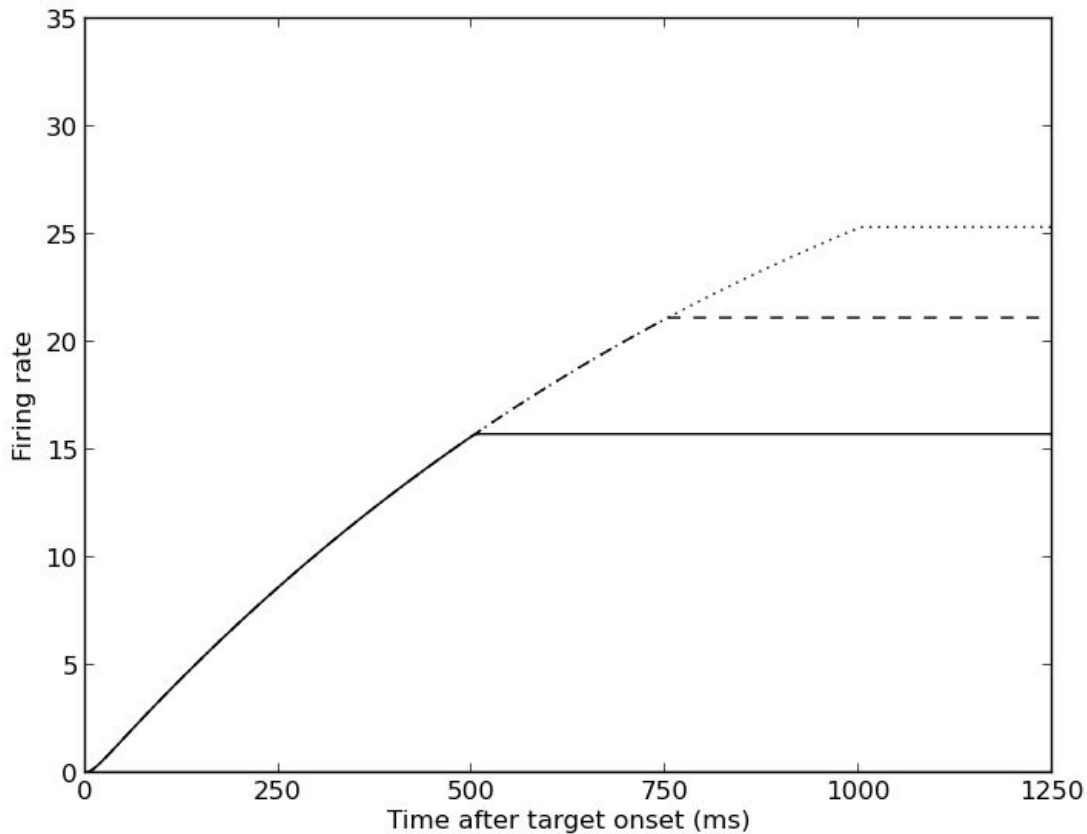
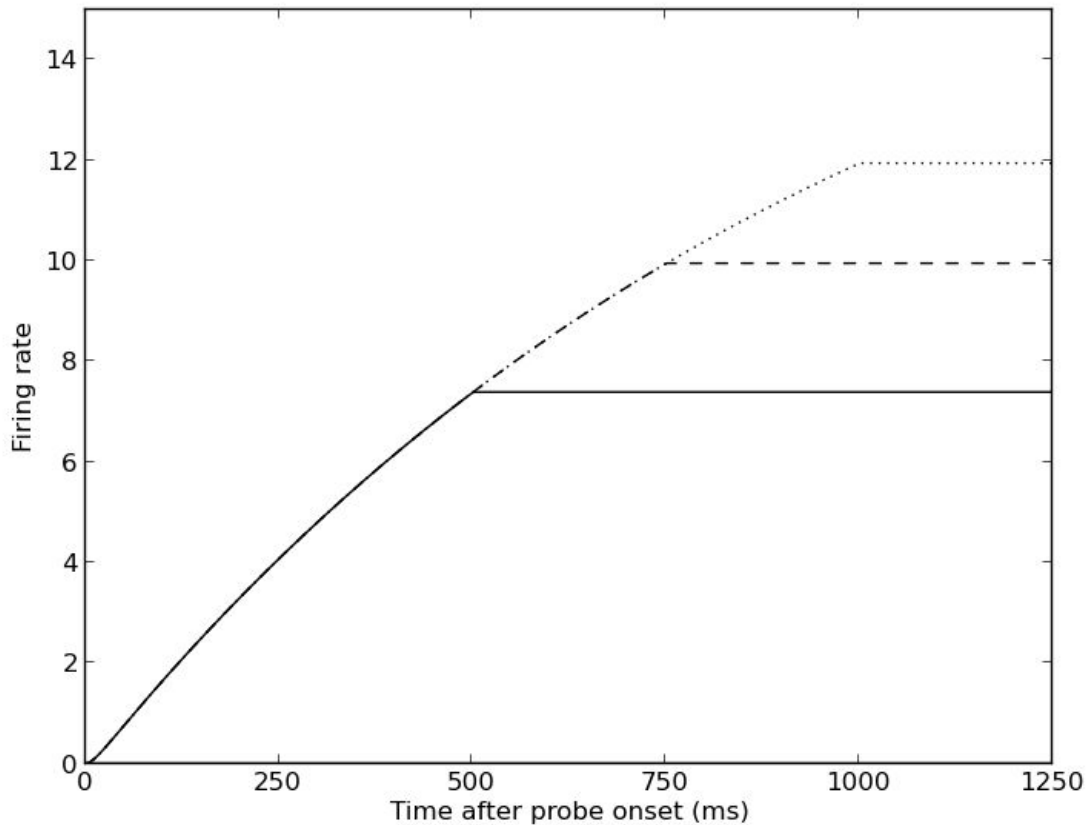


Figure 4B displays the activity in the  $D$  population ( $r_D$ ) during probe presentation in three different conditions: A 500 ms probe, a 750 ms probe, and a 1000 ms probe. The target duration was 750 ms in all three conditions. Again, as expected, the value of  $r_D$  upon probe stimulus offset is a function of the difference between probe and target durations, with greater values of  $r_D$  when the probe duration exceeds the target duration, and smaller values of  $r_D$  when the probe duration is less than the target duration.

**Figure 4B: Value of  $r_D$  during and after probe presentation, all target stimuli 750 ms; solid line denotes 500 ms probe, dashed line denotes 750 ms probe, dotted line denotes 1000 ms probe.**



It is clear from the above that the model is able to perform duration comparisons, with the activity of the  $D$  population ( $r_D$ ) serving as a measure of the difference in duration between target and probe stimuli.

### Simulation 2

A subjective shortening effect has repeatedly been found in experimental examinations of temporal comparison and memory for stimulus duration in both humans and non-humans (e.g., Spetch & Wilkie, 1983; Wearden & Ferrara, 1993; Wearden, Parry, & Stamp, 2002; Wearden,



Goodson, & Foran, 2007; Van Rooyen & Santi, 2009; Santi, Hoover, & Simmons, 2011). In the subjective shortening effect, the remembered duration of a stimulus decreases as a function of the delay between presentations of the target and probe stimuli. The longer the delay period, the more likely subjects are to report the probe as being longer than the target. In a classic study, Wearden and Ferrara (1993) presented subjects with auditory tones with a mean duration of 400 ms, separated by a delay ranging from one to ten seconds, and instructed subjects to report whether the target and probe stimuli were the same duration or different durations. Although overall performance was generally constant or declined only slightly, a significant effect of delay duration was found, with subjects reporting the target and probe as different more often at longer delays. In a second experiment, subjects reported the probe as being either longer, shorter, or the same duration as the target, and neither overall performance nor the proportion of correct responses in the probe-shorter condition were affected by the delay duration. However, increasing the duration of the delay period resulted in a decrease in the number of correct responses in the probe-same condition, and an increase in the number of correct responses in the probe-longer condition, consistent with a subjective shortening of the stored representation of the target stimulus.

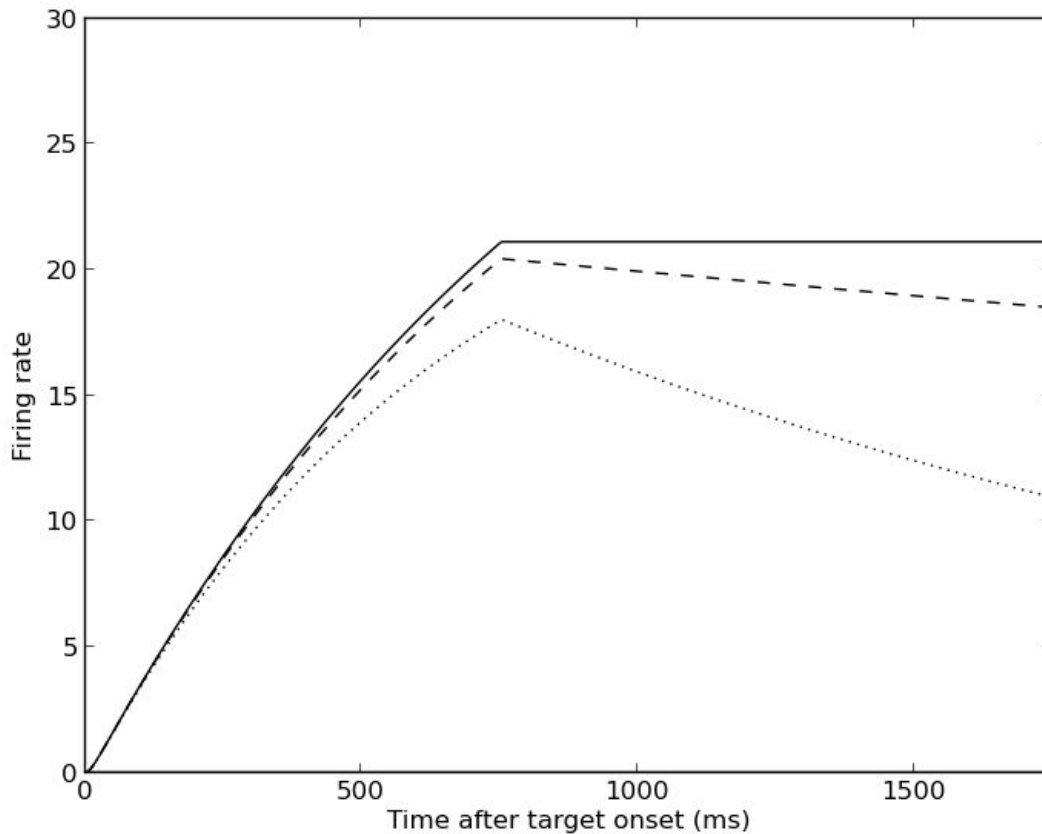
In Simulation 2, we demonstrate that when  $M$  is a leaky integrator (in other words, when  $w_{MM} < 1$ ), a subjective shortening effect is produced. Consider: If  $M$  is a leaky integrator,  $r_M$  will

decay over the delay period. As  $r_M$  is a function of the duration of the target stimulus, a decrease in  $r_M$  results in a subjective decrease in the duration of the target stimulus.

To demonstrate, we presented a 750 ms target stimulus to three versions of the model:

One in which  $w_{MM} = 1$ , one in which  $w_{MM} = 0.999$ , and one in which  $w_{MM} = 0.995$ . From Figure 5A, we can clearly see that when  $w_{MM} < 1$ ,  $r_M$  decreases over the delay period. As  $r_M$  is a function of the target stimulus duration, this decay in activity represents a decrease in the perceived duration of the target.

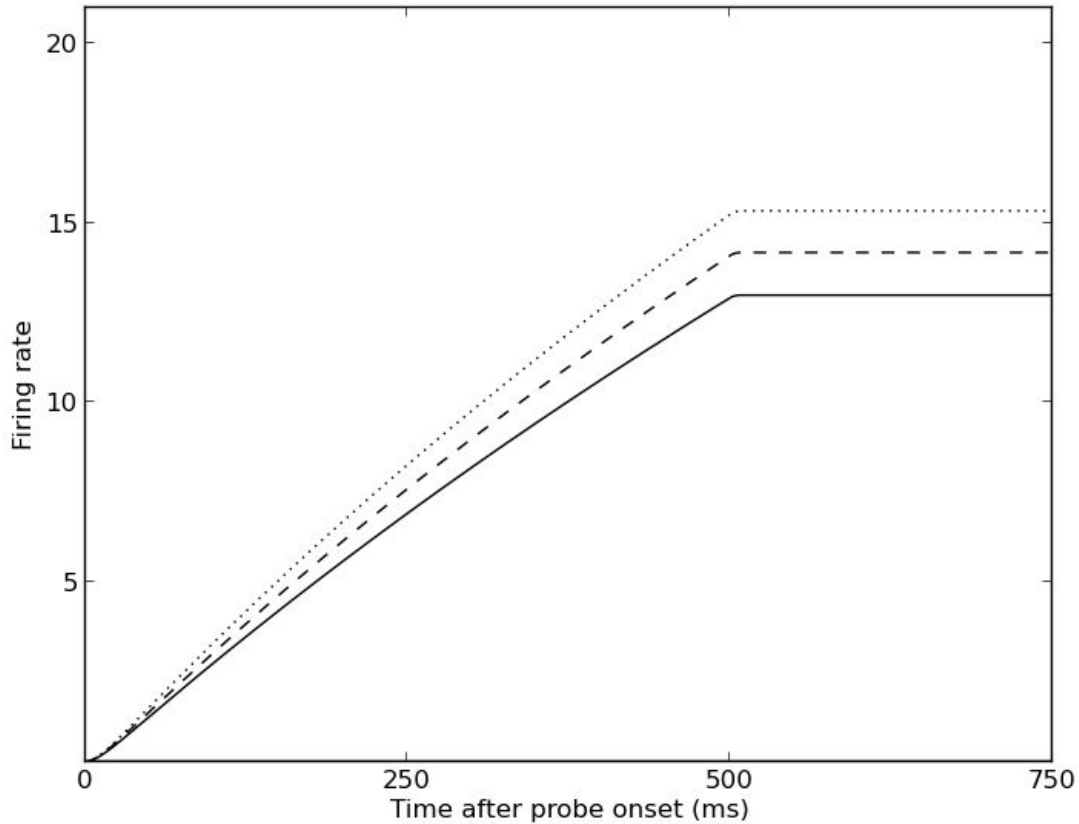
**Figure 5A: Value of  $r_M$ , 750 ms target; Solid line denotes  $w_{MM} = 1$ , dashed line denotes  $w_{MM} = .999$ , dotted line denotes  $w_{MM} = .995$ .**



To demonstrate the effects of this decay on decision-making, we presented the model with a 500 ms target, and a 500 ms probe, separated by a delay period of either 1000, 2000, or 4000 ms. The value of  $w_{MM}$  was set to 0.995. From Figure 5B, we can see that the activity of the  $D$  population ( $r_D$ ) was greatest at the longest delay period, and least at the shortest delay period. As  $r_D$  is a function of  $duration_{probe} - duration_{target}$ , we can see that increasing the delay period results in a decrease in the perceived duration of the target: In other words, a subjective shortening effect.

**Figure 5B: Value of  $r_D$ , 500 ms target and probe stimuli. Solid line denotes a 1000 ms delay**

period, dashed line denotes a 2000 ms delay period, dotted line denotes a 4000 ms delay period.



### General Discussion

The present study demonstrates that a single model can perform duration comparison tasks, as well as frequency comparison tasks (e.g., Bancroft et al., 2013, 2014). Further, not only can Miller and Wang's (2006) model be applied to STM for duration, a small change in the  $w_{MM}$  parameter can produce a subjective shortening effect.

One intriguing implication of the above work is that the representation in memory of

stimulus duration and non-duration stimulus properties are entwined. Experimental efforts in this area have been mixed. While Wearden, Parry, and Stamp (2002) failed to find an analogue to the subjective shortening effect for visual length, and Ogden and Jones (2011) also failed to find an analogue in memory for pitch and visual length, Khoshnoodi, Motiei-Langroudi, Omrani, Diamond, and Abbassian (2008) reported that vibrotactile stimulus frequency interacted with perceived stimulus duration. Although not all attempts to find an analogue to the subjective shortening effect in other scalar domains have been successful, Khoshnoodi et al.'s results are consistent with recent experimental efforts (Spitzer, Gloel, et al., 2014), and we have recently suggested that scalar stimuli share a common, modality-independent memory system (Bancroft, Hockley, & Servos, 2014). Further efforts to find similarities in both the cognitive processes and neural representations underlying different scalar STM tasks are likely to be rewarding.

One objection that could be made to the present model is that in the absence of external input, activity in the  $M$  populations will eventually dwindle to zero, causing catastrophic memory failure. We offer two counterarguments. First, the rate at which  $r_M$  decays is proportional to  $r_M$  (i.e.,  $dr_M/dt \propto -r_M$ ), and therefore the rate of decay decreases as the delay period progresses. As such, the greatest decay will occur early in the delay period. Sinclair and Burton (1996) found results consistent with this in a vibrotactile frequency STM task (although their results were partially obscured by order effects).

Second, it is not yet clear whether all neurons involved in duration memory tend to display decreasing activity. During the delay period of a vibrotactile STM task, many PFC neurons that encode vibrotactile frequency are early-selective (i.e., decaying) or persistent (i.e., non-decaying) (Romo et al., 1999). It may be the case that duration memory relies on both decaying and non-decaying subpopulations, in which case a partial memory representation would persist even at long durations. Testing this hypothesis will likely require single-cell recording efforts.

In the present paper, we have modeled duration comparison and subjective shortening. Perhaps most excitingly, we have used a model (Miller & Wang, 2006) designed for tasks requiring STM for stimulus frequency, and with relatively minor modifications. It should be noted that a shared system for scalar STM does not necessarily imply that the exact same neurons are involved in coding all types of scalar information. Some of the strongest evidence for such a shared system comes from the Spitzer et al. EEG studies discussed above, which obviously do not provide sufficient resolution to analyse activity on the level of individual neurons. The relatively small leak values ( $1-w_{MM}$ ) in the present study appear appropriate for application to other tasks; for example, STM for vibrotactile frequency has been shown to be stable for relatively short delay periods (Harris et al., 2002), but to decay at longer delay periods (Sinclair & Burton, 1996). This is consistent with memory for duration being one of a class of

scalar short-term memory tasks that share similar behavioural properties and neural systems, despite having different salient stimulus properties (such as frequency, intensity, or duration).

### **Acknowledgments**

We appreciate the comments provided by several anonymous reviewers. The present research was supported by grants from the Natural Sciences and Engineering Research Council to WEH and PS, and an Ontario Graduate Scholarship to TDB.

## **Chapter 5: Overwriting and intrusion in short-term memory**

Tyler D. Bancroft, Jeffery A. Jones, Tyler M. Ensor, William E. Hockley, and Philip Servos

### **Abstract**

Studies of interference in working and short-term memory suggest that irrelevant information may overwrite the contents of memory or intrude into memory. While some previous studies have reported greater interference when irrelevant information is similar to the contents of memory than when it is dissimilar, other studies have reported greater interference for dissimilar distractors than similar distractors. In the present study, we find the latter effect in a paradigm that uses auditory tones as stimuli. We suggest that the effects of distractor similarity to memory contents are mediated by the type of information held in memory, particularly the complexity or simplicity of information.

### **Introduction**

Working memory and short-term memory (WM/STM) are vulnerable to interference from irrelevant information. One method used to examine interference is to present subjects with a concurrent processing task during the memory task (e.g., Conway, Kane, & Engle, 2003; Conway et al., 2005; Klingberg, 1998; Pazzaglia, 1999). Oberauer, Farrell, Jarrold, Pasiecznik, and Greaves (2012) outline several possible mechanisms through which concurrent processing during WM/STM maintenance can interfere with memory. In one possible mechanism, increased demand on some cognitive resource that is general to both memory maintenance and concurrent processing results in decreased performance (resource-sharing accounts). In another possible mechanism, representations in WM/STM require constant active maintenance and concurrent processing interferes with this task (time-sharing accounts; see the model of



Barrouillet, Portrat, and Camos, 2011). In a third possible mechanism (similarity-based interference), interference is due to interference between the representations of items held in memory and representations of information processed during the concurrent task.

Evidence for the latter account comes from studies that have found an effect of item-distractor similarity on task performance. Notably, Oberauer, Farrell, et al. (2012) point out that such effects are not predicted by the first two accounts of interference described above, because these accounts predict that the magnitude of interference is determined by the amount of the shared resource required by the concurrent processing task, or the amount of time dedicated to the concurrent processing task (e.g., Barrouillet, Portrat, and Camos, 2011), and would apply equally to similar and dissimilar distractor items.

One mechanism that may underlie similarity-based interference is overwriting (e.g., Nairne, 1990; Oberauer & Kliegl, 2006). In overwriting accounts, item representations in memory are comprised of a set of features. When items in memory share features, they compete for the representations of those features in memory, and items can be degraded by “losing” the features to the representation of a different item. As such, distractors would be expected to produce greater interference when they are similar to items in memory, rather than dissimilar. For example, in a study by Lange and Oberauer (2005), subjects held a list of consonant-vowel-consonant trigrams in memory and were asked to read a second set of trigrams aloud during the retention period. Critically, a single trigram from the memory list was selected as a target, and the distractor list was designed to contain a distractor with the same first letter as the target, a distractor with the same second letter as the target, and a distractor with the same third letter as the target. As predicted by overwriting theory, target items were recalled less often

than the non-target items, which did not share features (i.e., letters) with the distractors.

In Lange and Oberauer's (2005) study, a trigram maintained in STM was susceptible to interference when a distractor had the potential to "capture" one of its features (i.e., when a distractor was similar to, rather than dissimilar from, the target trigram). In other words, distractor similarity impaired performance. In contrast, however, a number of studies have reported greater interference between dissimilar distractors and memory items than between similar distractors and memory items. For example, in a common vibrotactile STM task (e.g., Bancroft & Servos, 2011; Harris, Miniussi, Harris, & Diamond, 2002; Romo & Salinas, 2003), subjects are presented with a to-be-remembered vibration to the hand (the target), followed by an unfilled delay, followed by a second vibration (the probe), and are asked to report whether the target and probe are the same vibrational frequency or different vibrational frequencies (or, alternatively, whether the probe is a higher or lower frequency than the target). Bancroft and Servos (2011) presented a brief, irrelevant distractor stimulus during the delay period, with the frequency of the distractor either the same as the target or different from the target. Critically, when the target and probe were different frequencies, the distractor frequency could be the same as the target frequency, shifted toward the probe frequency, or shifted away from the probe frequency. For example, if the frequency of the target was 18 Hz, and the frequency of the probe was 22 Hz, the distractor could be 20 Hz (in the toward-shift condition), 16 Hz (in the away-shift condition), or 18 Hz (in the same-as-target condition).

Bancroft and Servos (2011) reported an interference effect, with poorer performance (i.e., a higher proportion of "same" responses) in the toward-shift condition than in the away-shift condition, which was interpreted as evidence that the distractor was encoded into memory, and

that the probe was being compared against some combination of the target and distractor (or, alternately, that the distractor was replacing the target in memory on some trials. Further, when the target and probe were the same frequency, there was no effect of distractor frequency, with similar performance whether the distractor was offset by 0 Hz, 2 Hz, or 4 Hz. This interference effect has been replicated experimentally (Bancroft, Servos, & Hockley, 2011), and computationally, using both a model of frontal cortex (Bancroft, Hockley, & Servos, 2013) and a diffusion model (Bancroft, Hockley, & Servos, 2012b). In the present paper, we will refer to interference that appears to be caused by the encoding of irrelevant information as intrusion-based interference.

Similar interference effects have been found in other modalities. In an early experiment examining short-term memory for simple stimuli, Guilford and Park (1931) employed a paradigm similar to that used by Bancroft et al. (Bancroft & Servos, 2011; Bancroft, Servos, & Hockley, 2011). All trials began with subjects lifting a 200 g target weight, which they compared to a probe that varied in weight from 185 to 215 g. There were three distractor conditions: a "heavier" condition, in which the distractor weighed 400 g; a "lighter" condition, in which the distractor weighed 100 g; and a control condition, in which the distractor weighed the same as the target (i.e., 200 g). Performance on the lighter-heavier discrimination task was poorer in the lighter- and heavier-distractor conditions relative to the control condition. Critically, there was evidence that the distractor affected subjects' memory of the target, with more "heavier" responses when the 100 g distractor was lifted between the target and probe, and more "lighter" responses when the 400 g distractor was lifted between the target and probe.

Mercer and McKeown (2010a) presented a distractor during the delay period of an

auditory STM task. Stimuli were complex tones comprised of four harmonics (always including the fundamental and 8th harmonic). Three types of distractors were used: those whose critical harmonics were shared by the target, those whose critical harmonics were shared by the probe, and novel distractors whose critical harmonics differed from both the target and probe.

Compared to performance for distractors including frequencies from the target, performance suffered when the distractor included novel frequencies or frequencies from the probe.

Critically, performance for novel distractors exceeded performance for distractors that were similar to the probe. Mercer and McKeown (2010b) reported similar results.

Deutsch (1970) demonstrated that same-different discrimination of pure tones was significantly impaired when six distractor tones were inserted between the target and probe compared to when six spoken digit distractors were inserted. In a subsequent experiment using four tone distractors rather than six, Deutsch (1972) reported that repeating the target among the distractors enhanced discrimination compared to presenting the probe among the distractors. In a recent replication of this experiment, Ries and DiGiovanni (2009) reported poorer discrimination for distractor sequences incorporating the probe relative to distractor sequences incorporating the target. Finally, Deutsch (1973) documented a pattern of results similar to those reported in the vibrotactile domain (Bancroft & Servos, 2011; Bancroft, Servos, & Hockley, 2011). This time, rather than repeating the target or presenting the probe among the distractors, Deutsch showed that including a distractor differing by a semitone from the target in the opposite direction from the probe (i.e., an away-shift distractor) in the distractor sequence yielded superior performance compared to a condition wherein one of the interpolated distractors was half a semitone different from the target, but in the direction of the probe (i.e., a toward-shift). Critically, and again

consistent with Bancroft and colleagues' vibrotactile work, although the direction of the shift affected performance on "different" trials (i.e., trials where  $f_{\text{target}} \neq f_{\text{probe}}$ ), performance on "same" trials (i.e.,  $f_{\text{target}} = f_{\text{probe}}$ ) was unaffected by the direction of the shift.

Magnussen, Greenlee, Asplund, and Dyrnes (1991) also found greater interference for dissimilar than for similar distractors. Subjects compared the spatial frequencies of two consecutively-presented visual gratings. When a distractor grating was presented during the delay period, the impact on performance increased when the spatial frequency of the distractor was dissimilar to the target.

Why does the similarity of the distractor to items in memory seem to produce different effects in different studies? The answer may lie in the nature of the stimuli used. Most of the above studies have a common attribute that may shed light on the relevant mechanism of interference: They are scalar short-term memory studies. A growing body of research suggests that short-term memory for scalar stimulus properties (properties of a stimulus that can be represented as a single value, such as auditory pure tone frequency, stimulus duration, or vibrational frequency) is supramodal, relies on a common storage system, and is represented with similar neural codes (e.g., Spitzer & Blankenburg, 2011, 2012; Spitzer, Fleck, & Blankenburg, 2014 see Bancroft, Hockley, and Servos, 2014b, for a recent review). In this storage system, the salient property of a stimulus (for example, the frequency of an auditory pure tone) is thought to be represented as a monotonic, often linear or approximately linear function of neuronal firing rates; this coding scheme has been demonstrated or implied for a variety of stimuli, including vibrotactile frequency, auditory frequency, visual flicker frequency, stimulus duration, stimulus amplitude, and the numerosity of a series of brief stimuli (e.g., Bancroft,

Hockley, & Servos, 2012b, 2013, 2014a; Jun et al., 2010; Lemus, Hernández, & Romo, 2009a, 2009b; Romo, Brody, Hernández, & Lemus, 1999; Romo & Salinas, 2003; Spitzer & Blankenburg, 2012; Spitzer, Fleck, & Blankenburg, 2014; Spitzer, Gloel, Schmidt, & Blankenburg, 2014; Spitzer, Wacker, & Blankenburg, 2010).

Although Deutsch's (1972, 1973) and Ries and DiGiovanni's (2009) experiments investigating short-term memory for pitch are qualitatively consistent with intrusion-based interference, aspects of these researchers' methodology limit the degree to which this interpretation is valid. First, in all of these studies, four distractors were interpolated between the target and probe. In Deutsch's (1972) and Ries and DiGiovanni's (2009) experiments, one of the conditions repeated the target among the distractors, a second included the probe among the distractors, and a third included neither the target nor the probe among the distractors. Deutsch (1973) took a similar approach, save that critical distractors were displaced by a semitone from the target. In all conditions, the nontarget and nonprobe distractors were drawn randomly from a predetermined frequency range. For an intrusion-based interference interpretation to be valid, we must make the assumption that, because the distractor frequencies were drawn randomly, the distractors would converge on a common frequency value across a sufficient number of trials, producing a neutral effect on performance. However, there is no way of determining whether or not this occurred in the finite number of trials used in Ries and DiGiovanni's and Deutsch's studies.

A second aspect of Ries and DiGiovanni's (2009) and Deutsch's (1972, 1973) methodology limits an intrusion-based interference interpretation. In Bancroft et al.'s (Bancroft & Servos, 2011; Bancroft, Servos, & Hockley, 2011) vibrotactile experiments, the single

distractor was considerably shorter in duration than the target and probe. In spite of this shorter duration, Bancroft et al. still found evidence that it intruded into memory. In contrast, the distractors in Ries and DiGiovanni's and Deutsch's experiments were the same duration as the targets and probes. Consequently, it is possible that, rather than demonstrating limited intrusion into memory, these experiments demonstrated complete overwriting of the target stimulus.

It should be noted that the task used in the present study (and in previous studies, such as those of Bancroft and Servos, 2011, and Bancroft, Servos, and Hockley, 2011) does not have an active concurrent processing task: Subjects do not actively engage in some active distractor task or activity (for example, performing mathematical operations, as in certain span tasks; see Conway et al., 2005, for an overview). Rather, the distractor stimulus is irrelevant to the task. Recent evidence suggests that task-irrelevant stimuli (such as distractors) may be inhibited during short-term memory maintenance, resulting in less interference with the contents of memory. Haegens, Osipova, Oostenveld, and Jensen (2010) reported increased alpha-band power over task-irrelevant sensory regions during maintenance compared to the pre-stimulus interval, and also early in the delay period on correct than incorrect trials, while Haegens, Luther, and Jensen (2012) reported that somatosensory alpha increased to suppress irrelevant information. Critically, increases in alpha activity are thought to be linked to inhibition of neuronal firing (Haegens, Nácher, Luna, Romo, & Jensen, 2011). Bisley, Zaksas, Droll, and Pasternak (2004) reported that activity in motion-encoding neurons in visual cortex was suppressed during the middle of the delay period, but not early or late in the delay. Most compellingly, Linke, Vicente-Grabovetsky, and Cusack (2011) applied a decoding method to fMRI data collected during an auditory STM experiment, and found inhibition of activity in

auditory cortex during maintenance - specifically, inhibition of activity around the frequency of stored stimuli. While it is theoretically possible that frequency-specific inhibition in sensory cortex plays some role in stimulus comparison or the decision-making process, Lemus et al. (2009a) reported that primary auditory cortex encoded stimulus information during stimulus presentation, but did not encode stimulus information during the memory or decision-making processes. Rather, memory and decision-making processes seem to rely on frontal systems (Lemus et al., 2009b).

While scalar STM tasks are of particular use in testing theories of interference, scalar STM also has significant theoretical implications for our understanding of the cognitive and neural structure of working and short-term memory. The recruitment of frontal cortex as the critical storage substrate for scalar STM contrasts with the recruitment of task-relevant sensory cortex as the storage substrate for STM for more complex forms of stimuli (Bancroft et al., 2014). While recent theoretical treatments of the neural basis of working and short-term memory suggest that memory storage relies on the task-relevant regions of cortex (Postle, 2006), and while recent fMRI studies of WM/STM for relatively complex stimuli that have applied decoding methods have supported this notion (e.g., Christophel, Hebart, & Jaynes, 2012), studies of scalar STM have found information storage in frontal cortex, rather than sensory cortex (Romo & Salinas, 2003; Spitzer & Blankenburg, 2012; Spitzer et al., 2014).

The assumption that scalar stimuli share a common, modality-independent representation implies that the intrusion-based interference effects we have previously identified in vibrotactile STM should also exist in other forms of scalar STM. In the present study, we extend previous vibrotactile interference findings to the auditory domain by adapting the vibrotactile task used by



Bancroft and Servos (2011) to use auditory pure tones as stimuli. Specifically, we aim to replicate their intrusion effect in the auditory domain - worse performance when the frequency of a distractor is shifted toward that of the probe (the toward-shift condition) than when it is further from the probe than target (the away-shift condition).

## Experiment 1

### Method

**Subjects.** Twelve undergraduate students from Wilfrid Laurier University participated for course credit. As pointed out in the statistical guidelines for Psychonomic Society journals, the optional stopping rule (e.g., testing for a significant effect as data is collected and stopping once it is found) has been shown to inflate the Type I error rate. We used an ad hoc stopping rule, in which we continued to collect data and re-test after a significant effect was found, which has been shown to reduce Type I error relative to the optional stopping rule (Yu, Sprenger, Thomas, & Dougherty, 2014). In Experiment 1, our effect of interest was the comparison between the different-away and different-toward distractor conditions, in which a significant effect was found at  $n = 3$ , and data collection and confirmatory testing was continued until  $n = 12$ .

**Apparatus and materials.** Auditory stimuli were generated using Audacity (version 2.0.4) and presented over Sony MDR-NC40 headphones (with the noise-cancelling function disabled), using SuperLab (version 4.5.1, San Pedro, California: Cedrus) running on a Windows XP machine. Target and probe stimuli were pure tones, 1000 ms in duration, and all target and probe stimuli had a frequency of either 510 or 515 Hz. Distractor stimuli were 250 ms in duration, and were 505, 510, 515, or 520 Hz. Different-frequency stimuli (both distractor and

probe) differed from the target stimulus by 5 Hz. Note that as previous efforts by our group have found little difference between toward-shift distractors with frequencies equal to the probe, and toward-shift distractors with frequencies between those of the target and probe (Bancroft & Servos, 2011; Bancroft, Servos, & Hockley, 2011), we opted for a relatively large shift in order to maximize statistical power.

**Procedure.** Subjects were presented with 240 trials. On each trial, subjects were presented with a 1000 ms target stimulus, followed by a 900 ms unfilled delay, a 250 ms distractor, a 350 ms delay, and a 1000 ms probe stimulus. Stimulus frequencies varied based on a 2 (probe type: same frequency as target vs. different frequency from target) X 2 (distractor frequency: same as target vs. different from target) within-subjects design, with 60 trials in each condition. Critically, on different-probe/different-distractor trials, the distractor frequency could be shifted away from the target frequency and toward the probe frequency (for example,  $f_{\text{target}} = 510 \text{ Hz}$ ,  $f_{\text{distractor}} = 515 \text{ Hz}$ ,  $f_{\text{probe}} = 515 \text{ Hz}$ ), or away from the target frequency and also away from the probe frequency (for example,  $f_{\text{target}} = 510 \text{ Hz}$ ,  $f_{\text{distractor}} = 505 \text{ Hz}$ ,  $f_{\text{probe}} = 515 \text{ Hz}$ ), with an equal number of trials (30) in each condition.

Subjects were instructed to report whether the target and probe were of the same pitch or different pitches, and to report their decision by pressing ‘s’ (for “same”) or ‘d’ (for “different”) on the keyboard. There was a 1000 ms delay between subject response and the beginning of the next trial.

## Results and Discussion

Proportions of correct responses as a function of probe type and distractor type are reported in Table 3, as are proportions of correct responses in the away-shift and toward-shift

conditions. Shapiro-Wilk analyses found no violations of normality (all  $p$ -values  $> .160$ ).

**Table 3: Mean proportions of correct responses from Experiment 1**

Condition	Mean proportion of correct responses (SEM)
Same probe/same distractor	.67 (.04)
Same probe/different distractor	.64 (.03)
Different probe/same distractor	.57 (.04)
Different probe/different (away-shift) distractor	.67 (.04)
Different probe/different (toward-shift) distractor	.48 (.04)
Interference effect	.19

A 2 (probe type: same-frequency probe vs. different-frequency probe) X 2 (distractor frequency: same as target frequency vs. different from target frequency) repeated-measures ANOVA was performed on the proportions of correct responses. The main effect of probe type approached significance,  $F(1, 11) = 3.813$ ,  $MS_e = .023$ ,  $p = .077$ , partial  $\eta^2 = .257$ , with performance in the same-probe condition exceeding that in the different-probe condition. No main effect of distractor frequency was found,  $F(1, 11) = .169$ ,  $MS_e = .004$ ,  $p = .689$ , partial  $\eta^2 = .015$ . The interaction between probe type and distractor frequency was not significant,  $F(1, 11) = 1.074$ ,  $MS_e = .002$ ,  $p = .322$ , partial  $\eta^2 = .089$ .

A paired-samples  $t$ -test was used in a planned comparison between the different probe/different (away-shift) distractor and different probe/different (toward-shift) distractor conditions. Performance on the away-shift condition was significantly better than on the

toward-shift condition,  $t(11) = 5.786, p < .001$ . Additional paired-samples  $t$ -tests found better performance on the different probe/different (away-shift) distractor condition than on the different probe/same distractor condition,  $t(11) = 3.649, p = .004$ , and better performance on the different probe/same distractor condition than on the different probe/different (toward-shift) distractor condition,  $t(11) = 3.039, p = .011$ .

This experiment replicated the intrusion effect reported by Bancroft and Servos (2011), with better performance in the away-shift distractor condition than in the toward-shift distractor condition. This finding is consistent with the intrusion of distractors into STM: As the difference between the probe frequency and a combination of target and distractor frequencies is greater in the away-shift condition than in the toward-shift condition, performance is better in the away-shift condition. This effect is further evident in the significant differences in performance between the different (away-shift) distractor and the same distractor conditions, and between the same distractor and different (toward-shift) distractor conditions. Ironically, this intrusion into memory produces an improvement in performance in the away-shift condition. The beneficial effects of the distractor in the away-shift condition should not mask the fact that the distractor *is* intruding into and interfering with the contents of memory - that the interference has a beneficial effect merely reflects that the information intruding into memory differs from the probe to a greater degree than the original information.

The marginal effect of probe type (with performance better on same-probe than different-probe stimuli) is consistent with a perceptually challenging task. If subjects find it difficult to perceive the difference between two auditory stimuli differing by 5 Hz, they will make a “same” response - in a sense, the “same” response is the default response.

## Experiment 2

In Experiment 1, distractor stimuli could assume frequency values of 505, 510, 515, or 520 Hz. Notably, distractors of 510 and 515 Hz were more common than those of 505 or 520 Hz, as the former were found in both the same- and different/toward-shift distractor conditions, while the latter were found only in the different/away-shift condition. This may have led subjects to deliberately encode the distractor, as doing so would “refresh” memory when distractors had the same frequency as the target. In Experiment 2, we dropped the same-frequency distractor condition, allowing us to equalize the incidence rates of the various distractor frequencies.

### Method

**Subjects.** Twenty-seven undergraduate and graduate students from Wilfrid Laurier University participated for course credit or for a payment of \$8. Three subjects were excluded from analysis: One subject did not complete the experiment, and two subjects performed below chance, leaving a final sample of 24 subjects. All subjects signed a consent form stating that participants had to have normal or corrected-to-normal vision and hearing to participate. Sampling proceeded as in Experiment 1, and our effect of interest again was the comparison between the different-away and different-toward distractor conditions, in which a significant effect was found at  $n = 20$ , and data collection and confirmatory testing was continued until  $n = 24$ .

**Apparatus and materials.** Auditory stimuli were generated using Audacity and presented over Sony MDR-NC40 headphones (with the noise-cancelling function disabled), using SuperLab (San Pedro, California: Cedrus) running on a Windows 8 machine. Target and probe stimuli were pure tones, 1000 ms in duration, and all target and probe stimuli had a

frequency of either 510 or 516 Hz. Distractor stimuli were 250 ms in duration, and were 506, 512, 514, or 520 Hz. Different-frequency stimuli (both distractor and probe) differed from the target stimulus by 6 Hz. Distractors always differed from the probe frequency by +/- 4 Hz; unlike Experiment 1, there was no condition in which the distractor was the same frequency as the target.

**Procedure.** Subjects were presented with 240 trials. On each trial, subjects were presented with a 1000 ms target stimulus, followed by a 900 ms unfilled delay, a 250 ms distractor, a 350 ms delay, and a 1000 ms probe stimulus. The target and probe were either the same frequency ( $f_{\text{probe}} = f_{\text{target}}$ ) or different frequencies ( $f_{\text{probe}} = f_{\text{target}} \pm 6$  Hz). The distractor frequency differed from the target frequency by  $\pm 4$  Hz. Critically, in the different-probe condition, the distractor frequency could be shifted away from the target frequency and toward the probe frequency (for example,  $f_{\text{target}} = 510$  Hz,  $f_{\text{distractor}} = 514$  Hz,  $f_{\text{probe}} = 516$  Hz), or away from the target frequency and also away from the probe frequency (for example,  $f_{\text{target}} = 510$  Hz,  $f_{\text{distractor}} = 506$  Hz,  $f_{\text{probe}} = 516$  Hz). There were 120 trials in the same-frequency probe condition, and 60 trials in each of the different-frequency probe conditions.

Subjects were instructed to report whether the target and probe were of the same pitch or different pitches, and to report their decision by pressing 's' (for "same") or 'd' (for "different") on the keyboard. There was a 1000 ms delay between subject response and the beginning of the next trial.

## Results and Discussion

Proportions of correct responses as a function of probe type and distractor type are reported in Table 4, as are proportions of correct responses in the away-shift and toward-shift conditions.

Shapiro-Wilk analyses found violations of normality in the same-probe condition ( $p = .053$ ), the different-probe/away-shift condition ( $p = .007$ ), and the different-probe/toward-shift condition ( $p = .043$ ). A planned paired-sample  $t$ -test found a significant difference in performance between the different-probe/away-shift and different-probe/toward-shift conditions,  $t(23) = 2.864$ ,  $p = .009$ . As the data was non-normal, we also performed the Wilcoxon signed-rank test, which also found a significant difference between the different-probe/away-shift (median = .83) and different-probe/toward-shift conditions (median = .71),  $z = 2.716$ ,  $p = .007$ ,  $W = 44$ , consistent with the results of the  $t$ -test.

**Table 4: Mean proportions of correct responses from Experiment 2**

Condition	Mean proportion of correct responses (SEM)
Same probe/Different distractor	.79 (.02)
Different probe/Different (away-shift) distractor	.76 (.04)
Different probe/Different (toward-shift) distractor	.66 (.04)
Interference effect	.09

As in Experiment 1, an effect of probe type was present, with better performance in the same-probe condition (median = .81) than the different-probe condition (median = .76),  $t(23) = 2.193$ ,  $p = .039$ . The results of the Wilcoxon signed-rank test were consistent with those of the  $t$ -test,  $z = 1.963$ ,  $p = .050$ ,  $W = 73.5$ .

The results of Experiment 2 are consistent with those of Experiment 1, with better performance in the different-probe/away-shift condition than the different-probe/toward-shift condition. The magnitude of the interference effect was smaller in Experiment 2 (.09) than in

Experiment 1 (.19), possibly due in part to the smaller distractor frequency shift in Experiment 2 relative to the probe frequency shift in the different-probe condition (4 Hz relative to 6 Hz) than in Experiment 1 (5 Hz relative to 5 Hz). The marginally-significant effect of probe type found in Experiment 1 was significant in Experiment 2. In general, the results of Experiment 2 replicated those of Experiment 1, and are consistent with the intrusion of distractors into memory.

### **General Discussion**

The nature of interference in WM/STM is an open and active research topic. A substantial body of research supports overwriting as a mechanism of interference (e.g., Lange & Oberauer, 2005; Oberauer, 2009; Oberauer & Lange, 2008; in contrast, though, see Jünger, Kliegl, & Oberauer, 2014). In the present study, we have extended the evidence for intrusion as a mechanism of interference in short-term memory by replicating the intrusion effect previously found in the vibrotactile modality (Bancroft & Servos, 2011; Bancroft, Servos, & Hockley, 2011; Bancroft, Hockley, & Servos, 2012b, 2013).

Previous studies have demonstrated that distractors can negatively affect same-different auditory frequency discrimination (Deutsch, 1970, 1972, 1973; Ries & DiGiovanni, 2009). Nevertheless, as discussed above, an intrusion-based interference account of Deutsch's and Ries and DiGiovanni's results is complicated for two methodological reasons: First, the inclusion of multiple distractors during the delay period, only one of which was systematically manipulated relative to the target and probe; second, the fact that the distractors were equal in duration to the target and probe. The present study avoided both of these limitations by inserting a single, relatively short distractor between the target and probe. Consistent with Deutsch's and Ries and DiGiovanni's results, we found better discrimination on trials with away-shift distractors



compared to trials with same-frequency and toward-shift distractors, and better discrimination for trials with same-frequency distractors compared to trials with toward-shift distractors. To the best of our knowledge, the present study is the first to demonstrate that single brief auditory distractors can intrude into the memory of a substantially longer target.

Scalar STM tasks share common cognitive and neural properties, independent of sensory modality, and have proven to be useful for testing theories of interference, the integration of low-level neural data with higher-level imaging and behavioural research, and the development of computational models. In the present study, we have demonstrated that interference effects identified in vibrotactile scalar STM can also be found in auditory scalar STM. Further, we have demonstrated that interference in scalar STM may not appear in the same fashion as overwriting in WM/STM for more complex stimuli, due to the simple nature of scalar stimuli.

Why? We propose that the answer has to do with the fact that scalar stimuli may take on a wide range of possible values, whereas complex stimuli may be more restricted - for example, auditory pure tones can take any chosen frequency, while a letter stimulus only has 26 possible values (in English). Consider, for instance, the consonant-vowel-consonant trigrams employed by Lange and Oberauer (2005) in their serial-recall study. Lange and Oberauer's distractors were trigrams sharing features (i.e., single letters) with one of the trigrams in the study list. These stimuli have multiple features. Scalar stimuli, in contrast, have only one relevant dimension, and this dimension may be able to take on a wide range of values. Consequently, integration of the target and distractor (perhaps as weighted averaging - see Bancroft et al., 2012b, 2013) may improve performance, as in the away-shift condition in the present study. This is not the case with stimuli such as letters. (We cannot “average” two letters.) As such, it may be appropriate

to consider the present results as a special case of similarity-based interference, where stimuli have a single task-relevant feature (stimulus frequency, duration, amplitude, etc.), but there may be multiple feature units (frontal neurons), each of which can represent the feature.

In one possibility, a target stimulus is held in memory, and a distractor “competes” for the feature units. Some units will encode the distractor, while others will encode the target. In this case, the effects of a distractor on task performance do not come from degradation of the target stimulus, as the target feature can be represented in a single unit (i.e., neuron). Rather, in this case, subjects incorporate the distractor into the decision-making process, possibly due to an inability to select only a single representation in memory for comparison to the probe. In this case, we have intrusion without loss of the original stimulus representation. Note that this differs from interference due to lack of distinctiveness (e.g., Oberauer, Farrell, et al., 2012), in that interference effects can be produced by distractors that are substantially different from information held in memory. This would seem to make it unlikely that subjects are confusing memory items and distractors.

Another possibility is that competition between target and distractor results in feature units (neurons) encoding some combination of the target and distractor. This is consistent with feature overwriting accounts. The greater effect of dissimilar distractors is due to the limited number of features used to represent simple stimuli in memory. Consider: If there is a relatively large set of available features, then distractors similar to target items will tend to produce greater interference, as they are competing with the target for feature units. In contrast, if we are using very simple stimuli, where there is only one salient feature and items differ in the “value” of that feature, then we would expect less interference for similar distractors (as they are “overwriting”

the feature units with a value similar to the original) than for dissimilar distractors (which would “overwrite” the feature units with a different value). The existence of multiple redundant feature units (i.e., frontal neurons) does not change this prediction. In this case, we have intrusion with loss of the original stimulus representation. In the strongest case, the distractor may completely displace the target from memory in some or all trials.

The results of the present study do not allow us to discriminate between these possibilities. What is suggested by the present study, however, is that intrusion into memory may be able to produce interference without requiring overwriting. As in previous studies (e.g., Oberauer, Farrell, et al., 2012), the interference effects found in the present study cannot be explained by resource-sharing or time-sharing accounts of interference. Presumably, distractors of any frequency require similar amounts of time and resources to process. However, the effects of those distractors differ based on their similarity to target and probe.

While previous research has demonstrated that intrusion-based interference can affect scalar STM (e.g., Bancroft & Servos, 2011; Bancroft, Servos, & Hockley, 2011; Bancroft et al., 2012b, 2013), it is less clear what other forms of interference may also operate on scalar STM. This is further complicated by the lack of a clear analogue to the span measure in scalar STM; examinations of interference in WM/STM have often focused on span tasks (e.g., Burgess, Gray, Conway, & Braver, 2011; Oberauer, Farrell, et al., 2012; Oberauer, Lewandowsky, Farrell, Jarrold, & Greaves, 2012). In contrast, we do not yet even have a consistent estimate of scalar STM capacity. Li et al. (2013) estimated a capacity of more than one stimulus for auditory scalar STM, and Bancroft et al. (2012a) suggested a capacity of more than one stimulus for vibrotactile scalar STM, but later modeling by Bancroft et al. (2013) suggested their 2012 results

did not conclusively support a capacity of more than one stimulus in the vibrotactile domain.

Many other open questions about scalar STM still exist. For example, are scalar short-term tasks actually working memory tasks, in that they recruit domain-general attentional and/or executive processes (e.g., Engle, 2002; Engle, Tuholski, Laughlin, & Conway, 1999)? Although scalar STM tasks have occasionally been referred to as working memory tasks (e.g., “parametric working memory”; e.g., Romo et al., 1999), and several experimental (Bancroft, Servos, & Hockley, 2011; Linke et al., 2011) and computational (Bancroft et al., 2013) studies have suggested that inhibitory mechanisms act to protect the contents of scalar STM from interference, we lack unequivocal evidence that they recruit the attentional and/or executive processes that are characteristic of working memory tasks.

The myriad ways in which interference can appear in short-term memory tasks is, perhaps, surprising, given the relative simplicity of many such tasks. In the present study, we have described similarity-based interference in a simple auditory short-term memory task. While some previous work (e.g., Lange & Oberauer, 2005; Oberauer, Farrell, et al., 2012) has suggested that distractors that are similar to information in memory are more likely to cause interference than dissimilar distractors, the present study builds on previous work suggesting that the opposite pattern can also be found. The relatively direct relationship between memory representations and the activity of the neurons involved in memory maintenance offers an exciting opportunity to study interference in a far deeper manner than has previously been done.

### **Acknowledgments**

The present research was supported by NSERC grants to WEH, JAJ, and PS, NSERC scholarships to TDB and TME, and an Ontario Graduate Scholarship to TDB. The present

manuscript comprises a portion of TDB's doctoral research.

## Chapter 6: General Discussion

Humans appear to possess a short-term memory system capable of maintaining representations of scalar information. The present work outlines the evidence for a common scalar STM representational scheme, reconciles a critical issue in the scalar STM literature, and extends previous behavioural and computational results to new domains.

In Chapter 2, I review the literature on the location of storage of information in WM/STM. While there is substantial support for theories suggesting that representational cortex (e.g., sensory cortex) is a storage substrate for WM/STM, there is also evidence that prefrontal cortex is involved in storage. Intriguingly, the complexity of the information being held in memory appears to influence the location of storage, with simple information being represented in frontal cortex, and more complex information being represented in other areas, such as sensory cortex.

Harris, Miniussi, Harris, and Diamond (2002) reported TMS to contralateral (but not ipsilateral) primary somatosensory cortex impaired vibrotactile STM, suggesting that this is evidence for somatotopically-organized STM storage in primary somatosensory cortex. This contrasts with single-cell recording work that shows information maintenance in frontal, but not sensory cortex (e.g., Romo, Brody, Hernández, & Lemus, 1999). In Chapter 3, I use computational modeling methods to demonstrate that Harris et al.'s (2002) results do not necessarily imply STM storage in sensory cortex. Rather, their findings can be replicated by assuming frontal storage of information, and a TMS-induced increase in noisy neural activity in sensory cortex, which in turn outputs into the frontal memory store, corrupting the memory representation.

In Chapter 4, I show that Miller and Wang's (2006) model, which has been applied to vibrotactile scalar STM (in Chapter 3, as well as by Bancroft, Hockley, and Servos, 2013), can easily be extended to model scalar STM for stimulus duration. Notably, depending on parameter selection, a classic psychophysical effect (the subjective shortening effect) emerges from simulations. The argument for a common scalar STM representation is strengthened by the finding that a model designed for vibrotactile STM can also easily be applied to scalar STM for duration. This is particularly interesting, as previously modeled vibrotactile STM tasks (Ch. 3; Bancroft et al., 2013) involve the comparison of stimuli of fixed duration but potentially different frequencies. Scalar STM for duration, however, involves stimuli of potentially different durations. The model, however, represents both in a similar fashion. This is consistent with the notion of a common scalar STM representation, and also with previous experimental reports that stimulus frequency and stimulus duration may interact in memory (e.g., Khoshnoodi, Motiei-Langroudi, Omrani, Diamond, Abbassian, 2008).

In Chapter 5, I extend previous experimental research on interference in vibrotactile scalar STM to the auditory domain. Bancroft and Servos (2011, also Bancroft, Servos, & Hockley, 2011) found that irrelevant distractor stimuli presented during the delay period of a vibrotactile STM task appeared to intrude into memory, overwriting the contents of memory or otherwise being incorporated into the process by which the probe is compared against the contents of memory. In Chapter 5, I demonstrate that similar effects can be obtained when stimuli are auditory pure tones, rather than tactile vibrations, consistent with a common representational scheme in STM for auditory pure tones and tactile vibrations. Intriguingly, this suggests that present theory on overwriting and interference in WM/STM requires extension, and

that the complexity of information held in memory is a factor in the effects of interference on WM/STM. Under certain circumstances, the intrusion of irrelevant information can actually improve performance on a task.

These studies are not without their limitations. The most significant is that each individual study does not provide overwhelming evidence for a common scalar STM system. While each study is consistent with the existence of a common scalar STM system, arguing for a common scalar STM system based on these studies is best done by viewing these studies together, and in the context of the literature. It could be argued, for example, that the presence of intrusion-based interference in both auditory scalar STM and vibrotactile scalar STM does not indicate a common scalar STM system, but rather a more general property of STM that is not confined to scalar STM, and that the ability of a single model to fit a variety of results from the scalar STM domain is due to the flexibility of the model.

One could counter that it is not a weakness to interpret the results of a study in the context of the overall literature. One could also argue that the similarities between results found in auditory scalar STM and vibrotactile scalar STM are much greater than the simple pattern of results, but are also inherent in the task design and the similar neural representations reported in previous neuroscience research (e.g., Romo & Salinas, 2003; Spitzer et al., 2010, 2014; Spitzer & Blankenburg, 2011, 2012).

The computational model used in the present study (adapted from Miller and Wang, 2006) is relatively abstract and at a relatively coarse level of representation. While it retains fundamental operational characteristics of the neurons thought to be responsible for scalar STM maintenance (e.g., Romo & Salinas, 2003), and has been shown to be effective at modeling



behavioural data (aside from the studies in this dissertation, also see Bancroft et al., 2013), it is less effective at replicating the observed firing patterns of individual neurons (Jun et al., 2010; note that the Miller and Wang model is not the only model with this problem). As such, its biological accuracy is somewhat limited, and this makes it difficult to make inferences about the low-level operation of the scalar STM system from the modeling work in this dissertation.

A reasonable counter to these limitations might be as follows: Yes, the model has some well-defined limitations. However, it is able to adequately model behavioural data from a variety of tasks, to do so with fewer free parameters than datapoints, and that interesting emergent effects are produced by the model.

Perhaps the greatest question left unanswered is why scalar STM would have a separate neural representation from other forms of information. I do not claim to have an answer to this question, nor is the present body of work intended to produce an answer. The aim of the present dissertation is to provide further evidence for a scalar STM system and to resolve certain inconsistencies in the literature.

This manuscript supplements the literature on scalar STM, and suggests that scalar information held in STM is represented in a similar manner, independent of sensory modality. It also suggests that the neural substrate for scalar STM storage resides in the frontal lobe, in contrast with recent studies reporting storage in sensory cortex. This raises an important question.

### **Where is information in WM/STM represented in the brain?**

Models of WM/STM have proposed variously that maintenance takes place in prefrontal cortex, or in sensory/representational cortex. Scalar STM serves as an example of an STM task

where information is clearly represented in frontal cortex. However, there is strong evidence that other STM tasks recruit sensory cortex (e.g., Christophel, Hebart, & Haynes, 2012; Zhou & Fuster, 1996). Further, recent studies have identified multiple representations of information across various areas of cortex (e.g., Ester, Sprague, & Serences, 2015). Studies using some form of decoding method (whether based on fMRI, EEG, or single-cell recording) have found evidence for WM/STM maintenance in early visual cortex (Christophel et al., 2012; Serences, Ester, Vogel, & Awh, 2009), posterior parietal cortex (Christophel et al., 2012; Ester et al., 2015), the superior intraparietal sulcus (Bettencourt & Xu, 2016), the inferior prefrontal gyrus in humans (e.g., Spitzer, Wacker, & Blankenburg, 2010) or inferior convexity of the prefrontal cortex in monkeys (e.g., Romo et al., 1999), and the pre-supplementary motor area (Vergara, Rivera, Rossi-Pool, & Romo, 2016). Information may be encoded in multiple regions, such as visual cortex and parietal cortex (Christophel et al., 2012). In essence, any area that is capable of representing features of a to-be-remembered stimulus appears to be a candidate as a WM/STM storage system.

These findings are not simply based on elevated, sustained activity in a region, but rather on the presence of a signal interpretable as representing a stimulus. In fact, such a signal may exist in the absence of sustained activity increases in a system (Serences et al., 2009). Even more remarkably, it has recently been proposed that storage need not involve sustained changes in firing rate at all, but may occur in subthreshold changes in membrane potentials, or short-term synaptic plasticity on the order of seconds (Ester et al., 2015; Mendoza-Halliday, Torres, & Martinez-Trujillo, 2014; Stokes, 2015). In these cases, a neural signature for WM/STM maintenance could be difficult to detect, particularly using non-invasive methods.

We are left with a remarkably convoluted picture of the neural basis for working and short-term memory. It is clear that WM/STM need not rely on memory storage in the form of persistent activity in frontal cortex, as proposed in the “standard model” (Postle, 2006), however, it is also apparent that persistent activity in frontal cortex *does* sometimes represent information held in WM/STM. To further complicate matters, memory storage does not appear limited to a single cortical area, but may co-occur across multiple regions. In light of the apparent decoupling between elevated activity and information encoding, of what utility are the numerous studies of WM/STM that have not used decoding-based analysis methods? If we exclude that segment of the literature, we are left with the relatively small (although growing) subset of studies that have used decoding methods as the basis for constructing a model of WM/STM maintenance. This approach would ignore the majority of studies published on the neuroscience of WM/STM, but it is not presently clear how to interpret such studies, in light of recent developments.

Perhaps the first question that must be answered is this: What mechanism or mechanisms operate on a cellular (or perhaps molecular) level to support WM/STM? A variety of mechanisms have been implicated, including firing rate, synaptic modification, and subthreshold membrane potential changes (e.g., Ester et al., 2015). Fully understanding where information in WM/STM is encoded may require us to first improve our understanding of how information in WM/STM is encoded. Methodological shifts and innovations may be necessary to develop an accurate understanding of where, when, and how WM/STM storage takes place in the brain.

This potentially foreshadows a larger issue: Is the answer to this question consistent across the brain? Or will we have to understand the operations of each and every neural system

separately before we understand how they interact to produce the phenomenon of memory? If so, working and short-term memory may be one of the last neural functions to be fully understood.

## References

- Abrahamyan, A., Clifford, C. W. G., Arabzadeh, E., & Harris, J. A. (2011). Improving visual sensitivity with subthreshold transcranial magnetic stimulation. *Journal of Neuroscience*, *31*, 3290-3294.
- Acheson, D. J., Hamidi, M., Binder, J. R., & Postle, B. R. (2011). A common neural substrate for language production and verbal working memory. *Journal of Cognitive Neuroscience*, *23*, 1358-1367.
- Amassian, V. E., Cracco, R. Q., Maccabee, P. J., & Cracco, J. B. (2002). Visual system. In A. Pascual-Leone, N. J. Davey, J. Rothwell, E. M. Wasserman, & B. K. Puri (Eds.), *Handbook of transcranial magnetic stimulation* (pp. 323-334). Arnold: London.
- Anderson, D. E., Ester, E. F., Serences, J. T., & Awh, E. (2013). Attending multiple items decreases the selectivity of population responses in human primary visual cortex. *Journal of Neuroscience*, *33*, 9273-9282.
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation*, Vol. 2 (pp. 89-195). Oxford: Academic Press.
- Auksztulewicz, R., Spitzer, B., & Blankenburg, F. (2012). Recurrent neural processing and somatosensory awareness. *Journal of Neuroscience*, *32*, 799-805.
- Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*, 417-423.
- Baddeley, A. (2010). Working memory. *Current Biology*, *20*, R136-R140.

- Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63*, 1-29.
- Baddeley, A. D. (2001). Is working memory still working? *American Psychologist, 56*, 851-864.
- Baddeley, A., & Hitch, G. J. (2010). Working memory. *Scholarpedia, 5*, 3015.
- Bancroft, T., & Servos, P. (2011). Distractor frequency influences performance in vibrotactile working memory. *Experimental Brain Research, 208*, 529-532.
- Bancroft, T. D., Hockley, W. E., & Servos, P. (2011). Vibrotactile working memory as a model paradigm for psychology, neuroscience, and computational modeling. *Frontiers in Human Neuroscience, 5*, 162.
- Bancroft, T. D., Hockley, W. E., & Servos, P. (2012a). Can vibrotactile working memory store multiple items? *Neuroscience Letters, 514*, 31-34.
- Bancroft, T. D., Hockley, W. E., & Servos, P. (2012b). Diffusion modeling of interference in vibrotactile working memory. *NeuroReport, 23*, 255-258.
- Bancroft, T. D., Hockley, W. E., & Servos, P. (2013). Irrelevant sensory stimuli interfere with working memory storage: Evidence from a computational model of prefrontal neurons. *Cognitive, Affective, & Behavioral Neuroscience, 13*, 23-34.
- Bancroft, T. D., Hockley, W. E., & Servos, P. (2014). Does stimulus complexity determine whether working and short-term memory storage relies on prefrontal cortex or sensory cortex? *Attention, Perception, & Psychophysics, 76*, 1954-1961.
- Bancroft, T. D., Hogeveen, J., Servos, P., & Hockley, W. E. (2014). TMS-induced neural noise in sensory cortex interferes with short-term memory storage in prefrontal cortex.

*Frontiers in Computational Neuroscience*, 8, 23.

Bancroft, T. D., Servos, P., & Hockley, W. E. (2011). Mechanisms of interference in vibrotactile working memory. *PLoS ONE*, 6, e22518.

Barak, O., Tsodyks, M., & Romo, R. (2010). Neuronal population coding of parametric working memory. *Journal of Neuroscience*, 30, 9424-9430.

Barrouillet, P., Portrat, S., & Camos, V. (2011). On the law relating processing to storage in working memory. *Psychological Review*, 118, 175-192.

Bestmann, S., Baudewig, J., Siebner, H. R., Rothwell, J. C., & Frahm, J. (2005). BOLD MRI responses to repetitive TMS over human dorsal premotor cortex. *NeuroImage*, 28, 22-29.

Bettencourt, K. C., & Xu, Y. (2016). Decoding the content of visual short-term memory under distraction in occipital and parietal areas. *Nature Neuroscience*, 19, 150-157.

Bisley, J. W., Zaksas, D., Droll, J. A., & Pasternak, T. (2004). Activity of neurons in cortical area MT during a memory for motion task. *Journal of Neurophysiology*, 91, 286-300.

Bohning, D. E., Shastri, A., McConnell, K. A., Nahas, Z., Lorberbaum, J. P., Roberts, D. R., Teneback, C., Vincent, D. J., & George, M. S. (1999). A combined TMS/fMRI study of intensity-dependent TMS over motor cortex. *Biological Psychiatry*, 45, 385-394.

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, 1196-1207.

Burgess, G. C., Gray, J. R., Conway, A. R. A., & Braver, T. S. (2011). Neural mechanisms of interference control underlie the relationship between fluid intelligence and working

- memory span. *Journal of Experimental Psychology: General*, *140*, 674-692.
- Burton, H., & Sinclair, R. J. (2000). Attending to and remembering tactile stimuli: A review of brain imaging data and single-neuron responses. *Journal of Clinical Neurophysiology*, *17*, 575-591.
- Christophel, T. B., Hebart, M. N., & Haynes, J.-D. (2012). Decoding the contents of visual short-term memory from human visual and parietal cortex. *Journal of Neuroscience*, *32*, 12983-12989.
- Constantinidis, C., & Steinmetz, M. A. (1996). Neuronal activity in posterior parietal area 7a during the delay periods of a spatial memory task. *Journal of Neurophysiology*, *76*, 1352-1355.
- Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, *7*, 547-552.
- Conway, R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*, 769-786.
- Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62-101). Cambridge University Press.
- Dade, L. A., Zatorre, R. J., Evans, A. C., & Jones-Gottman, M. (2001). Working memory in another dimension: Functional imaging of human olfactory working memory.



- NeuroImage*, 14, 650-660.
- D'Esposito, M., Detre, J. A., Alsop, D. C., Shin, R. K., Atlas, S., & Grossman, M. (1995). The neural basis of the central executive system of working memory. *Nature*, 378, 279-281.
- Deutsch, D. (1970). Tones and numbers: Specificity of interference in immediate memory. *Science*, 168, 1604-1605.
- Deutsch, D. (1972). Effect of repetition of standard and of comparison tones on recognition memory for pitch. *Journal of Experimental Psychology*, 93, 156-162.
- Deutsch, D. (1973). Interference in memory between tones adjacent in the musical scale. *Journal of Experimental Psychology*, 100, 228-231.
- Duncan, J. (2001). An adaptive coding model of neural function in prefrontal cortex. *Nature Reviews Neuroscience*, 2, 820-829.
- Duncan, J., & Miller, E. K. (2002). Cognitive focus through adaptive neural coding in the primate prefrontal cortex. In *Principles of Frontal Lobe Function*, Stuss, D. T., & Knight, R. T. (eds.). Oxford: Oxford University Press.
- Durstewitz, D. (2004). Neural representation of interval time. *NeuroReport*, 15, 745-749.
- Emrich, S. M., Riggall, A. C., LaRocque, J. J., & Postle, B. R. (2013). Distributed patterns of activity in sensory cortex reflect the precision of multiple items maintained in visual short-term memory. *Journal of Neuroscience*, 33, 6516-6523.
- Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, 11, 19-23.
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory,

- short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*, 309-331.
- Ester, E. F., Sprague, T. C., & Serences, J. T. (2015). Parietal and frontal cortex encode stimulus-specific mnemonic representations during visual working memory. *Neuron*, *87*, 893-905.
- Ester, E. F., Anderson, D. E., Serences, J. T., & Awh, E. (2013). A neural measure of precision in visual working memory. *Journal of Cognitive Neuroscience*, *25*, 754-761.
- Freedman, D. J., Riesenhuber, M., Poggio, T. & Miller, E. K. (2001). Categorical representation of visual stimuli in the primate prefrontal cortex. *Science*, *291*, 312-316.
- Fuster, J. M. (2003). More than working memory rides on long-term memory. *Behavioral and Brain Sciences*, *26*, 737.
- Gerwig, M., Kastrup, O., Meyer, B.-U., & Niehaus, L. (2003). Evaluation of cortical excitability by motor and phosphene thresholds in transcranial magnetic stimulation. *Journal of the Neurological Sciences*, *215*, 75-78.
- Guilford, J. P., & Park, D. G. (1931). The effect of interpolated weights upon comparative judgments. *American Journal of Psychology*, *43*, 589-599.
- Haegens, S., Osipova, D., Oostenveld, R., & Jensen, O. (2010). Somatosensory working memory performance in humans depends on both engagement and disengagement of regions in a distributed network. *Human Brain Mapping*, *31*, 26-35.
- Haegens, S., Nácher, V., Luna, R., Romo, R., & Jensen, O. (2011).  $\alpha$ -Oscillations in the monkey sensorimotor network influence discrimination performance by rhythmical inhibition of neuronal spiking. *Proceedings of the National Academy of Sciences*, *108*, 19377-19382.

- Haegens, S., Luther, L., & Jensen, O. (2012). Somatosensory anticipatory alpha activity increases to suppress distracting input. *Journal of Cognitive Neuroscience*, *24*, 677-685.
- Harris, J. A., Harris, I. M., & Diamond, M. E. (2001). The topography of tactile working memory. *Journal of Neuroscience*, *21*, 8262-8269.
- Harris, J. A., Miniussi, C., Harris, I. M., & Diamond, M. E. (2002). Transient storage of a tactile memory trace in primary somatosensory cortex. *Journal of Neuroscience*, *22*, 8720-8725.
- Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, *458*, 632-635.
- Hayden, B. Y., & Gallant, J. L. (2013). Working memory and decision process in visual area V4. *Frontiers in Decision Neuroscience*, *7*, 18.
- Hernández, A., Zainos, A., & Romo, R. (2002). Temporal evolution of a decision-making process in medial premotor cortex. *Neuron*, *33*, 959-972.
- Jun, J. K., Miller, P., Hernández, A., Zainos, A., Lemus, L., Brody, C. D., & Romo, R. (2010). Heterogenous population coding of a short-term memory and decision task. *Journal of Neuroscience*, *30*, 916-929.
- Jünger, E., Kliegl, R., & Oberauer, K. (2014). No evidence for feature overwriting in visual working memory. *Memory*, *22*, 374-389.
- Khoshnoodi, M. A., Motiei-Langroudi, R., Omrani, M., Diamond, M. E., & Abbassian, A. H. (2008). Effect of tactile stimulus frequency on time perception: The role of working memory. *Experimental Brain Research*, *185*, 623-633.
- Klingberg, T. (1998). Concurrent performance of two working memory tasks: Potential

- mechanisms of interference. *Cerebral Cortex*, 8, 593-601.
- Konstantinou, N., & Lavie, N. (2013). Dissociable roles of different types of working memory load in visual detection. *Journal of Experimental Psychology: Human Perception and Performance*, 39, 919-924.
- Kumar, S., Joseph, S., Pearson, B., Teki, S., Fox, Z. V., Griffiths, T. D., & Husain, M. (2013). Resource allocation and prioritization in auditory working memory. *Cognitive Neuroscience*, 4, 12-20.
- Lange, E. B., & Oberauer, K. (2005). Overwriting of phonemic features in serial recall. *Memory*, 13, 333-339.
- Lee, S.-H., Kravitz, D. J., & Baker, C. I. (2013). Goal-dependent dissociation of visual and prefrontal cortices during working memory. *Nature Neuroscience*, 16, 997-999.
- Lemus, L., Hernández, A., & Romo, R. (2009a). Neural codes for perceptual discrimination of acoustic flutter in the primate auditory cortex. *Proceedings of the National Academy of Sciences*, 106, 9471-9476.
- Lemus, L., Hernández, A., & Romo, R. (2009b). Neural encoding of auditory discrimination in ventral premotor cortex. *Proceedings of the National Academy of Sciences*, 106, 14640-14645.
- Li, D., Cowan, N., & Saults, J. S. (2013). Estimating working memory capacity for lists of nonverbal sounds. *Attention, Perception, & Psychophysics*, 75, 145-160.
- Linden, D. E. J., Oosterhof, N. N., Klein, C., & Downing, P. E. (2012). Mapping brain activation and information during category-specific visual working memory. *Journal of Neurophysiology*, 107, 628-639.

- Linke, A. C., Vicente-Grabovetsky, A., & Cusack, R. (2011). Stimulus-specific suppression preserves information in auditory short-term memory. *Proceedings of the National Academy of Sciences, 108*, 12961-12966.
- Magnussen, S., Greenlee, M. W., Asplund, R., & Dyrnes, S. (1991). Stimulus-specific mechanisms of visual short-term memory. *Vision Research, 31*, 1213-1219.
- Mecklinger, A., Gruenewald, C., Besson, M., Magnie, M.-N., & Von Cramon, Y. (2002). Separable neuronal circuitries for manipulable and non-manipulable objects in working memory. *Cerebral Cortex, 12*, 1115-1123.
- Mendoza-Halliday, D., Torres, S., & Martinez-Trujillo, J. C. (2014). Sharp emergence of feature-selective sustained activity along the dorsal visual pathway. *Nature Neuroscience, 17*, 1255-1262.
- Menning, H., Roberts, L. E., & Pantev, C. (2000). Plastic changes in the auditory cortex induced by intensive frequency discrimination training. *NeuroReport, 11*, 817-822.
- Mercer, T., & McKeown, D. (2010a). Interference in short-term auditory memory. *Quarterly Journal of Experimental Psychology, 63*, 1256-1265.
- Mercer, T., & McKeown, D. (2010b). Updating and feature overwriting in short-term memory for timbre. *Attention, Perception, & Psychophysics, 72*, 2289-2303.
- Miller, P., & Wang, X.-J. (2006). Inhibitory control by an integral feedback signal in prefrontal cortex: A model of discrimination between sequential stimuli. *Proceedings of the National Academy of Sciences, 103*, 201-206.
- Miniussi, C., Harris, J. A., & Ruzzoli, M. (2013). Modelling non-invasive brain stimulation in cognitive neuroscience. *Neuroscience and Biobehavioral Reviews, 37*, 1702-1712.

- Moliadze, V., Zhao, Y., Eysel, U., & Funke, K. (2003). Effect of transcranial magnetic stimulation on single-unit activity in the cat primary visual cortex. *Journal of Physiology*, *553*, 665-679.
- Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition*, *18*, 251-269.
- Oberauer, K. (2009). Interference between storage and processing in working memory: Feature overwriting, not similarity-based competition. *Memory & Cognition*, *37*, 346-357.
- Oberauer, K., & Kliegl, R. (2006). A formal model of capacity limits in working memory. *Journal of Memory & Language*, *55*, 601-626.
- Oberauer, K., & Lange, E. B. (2008). Interference in verbal working memory: Distinguishing similarity-based confusion, feature overwriting, and feature migration. *Journal of Memory and Language*, *58*, 730-745.
- Oberauer, K., Farrell, S., Jarrold, C., Pasiiecznik, K., & Greaves, M. (2012). Interference between maintenance and processing in working memory: The effect of item-distractor similarity in complex span. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 665-685.
- Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (2012). Modeling working memory: An interference model of complex span. *Psychonomic Bulletin & Review*, *19*, 779-819.
- Ogden, R. S., & Jones, L. A. (2011). Modality effects in memory for basic stimulus attributes: A temporal and nontemporal comparison. *Quarterly Journal of Experimental Psychology*, *64*, 1354-1371.
- Oliveri, M., Caltagirone, C., Filippi, M. M., Traversa, R., Cicinelli, P., Pasqualetti, P., & Rossini,

- P. M. (2000). Paired transcranial magnetic stimulation protocols reveal a pattern of inhibition of facilitation in the human parietal cortex. *Journal of Physiology*, *529*, 461-468.
- Pasley, B. N., Allen, E. A., & Freeman, R. D. (2009). State-dependent variability of neuronal responses to transcranial magnetic stimulation of the visual cortex. *Neuron*, *62*, 291-303.
- Pasternak, T., & Greenlee, M. W. (2005). Working memory in primate sensory systems. *Nature Reviews Neuroscience*, *6*, 97-107.
- Paus, T. (2005). Inferring causality in brain images: A perturbation approach. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*, 1109-1114.
- Paus, T., Jech, R., Thompson, C. J., Comeau, R., Peters, T., & Evans, A. C. (1997). Transcranial magnetic stimulation during positron emission tomography: A new method for studying connectivity of the human cerebral cortex. *Journal of Neuroscience*, *17*, 3178-3184.
- Pazzaglia, F. (1999). The role of distinct components of visuo-spatial working memory in the processing of texts. *Memory*, *7*, 19-41.
- Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience*, *139*, 23-38.
- Ptito, M., Fumal, A., Martens de Noordhout, A., Schonen, J., Gjedde, A., & Kupers, R. (2008). TMS of the occipital cortex induces tactile sensations in the fingers of blind Braille readers. *Experimental Brain Research*, *184*, 193-200.
- Raabe, M., Fischer, V., Bernhardt, D., & Greenlee, M. W. (2013). Neural correlates of spatial working memory load in a delayed match-to-sample saccade task. *NeuroImage*, *71*,

84-91.

Ragert, P., Becker, M., Tegenthoff, M., Pleger, B., & Dinse, H. R. (2004). Sustained increase of somatosensory cortex excitability by 5 Hz repetitive transcranial magnetic stimulation studied by paired median nerve stimulation in humans. *Neuroscience Letters*, *356*, 91-94.

Ragert, P., Franzkowiak, S., Schwenkreis, P., Tegenthoff, M., & Dinse, H. R. (2008).

Improvement of tactile perception and enhancement of cortical excitability through intermittent theta burst rTMS over human primary somatosensory cortex. *Experimental Brain Research*, *184*, 1-11.

Ray, P. G., Meador, K. J., Epstein, C. M., Loring, D. W., Day, L. J. (1998). Magnetic stimulation of visual cortex: Factors influencing the perception of phosphenes. *Journal of Clinical Neurophysiology*, *15*, 351-357.

Repovš, G., & Baddeley, A. (2006). The multi-component model of working memory:

Explorations in experimental cognitive psychology. *Neuroscience*, *139*, 5-21.

Ries, D. T., & DiGiovanni, J. J. (2009). Effects of recurrent tonal information on auditory working memory for pitch. *Hearing Research*, *255*, 14-21.

Riggall, A. C., & Postle, B. R. (2012). The relationship between working memory storage and elevated activity as measured with functional magnetic resonance imaging. *Journal of Neuroscience*, *32*, 12990-12998.

Rihs, T. A., Michel, C. M., & Thut, G. (2007). Mechanisms of selective inhibition in visual spatial attention are indexed by  $\alpha$ -band EEG synchronization. *European Journal of Neuroscience*, *25*, 603-610.

Rissman, J., Gazzaley, A., & D'Esposito, M. (2008). Dynamic adjustments in prefrontal,



- hippocampal, and inferior temporal interactions with increasing visual working memory load. *Cerebral Cortex*, *18*, 1618-1629.
- Robertson, E. M., Théoret, H., & Pascual-Leone, A. (2003). Studies in cognition: The problems solved and created by transcranial magnetic stimulation. *Journal of Cognitive Neuroscience*, *15*, 948-960.
- Romo, R., & Salinas, E. (2003). Flutter discrimination: Neural codes, perception, memory and decision making. *Nature Reviews Neuroscience*, *4*, 203-218.
- Romo, R., Brody, C. D., Hernández, A., & Lemus, L. (1999). Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature*, *399*, 470-473.
- Romo, R., Hernández, A., Zainos, A., & Salinas, E. (1998). Somatosensory discrimination based on cortical microstimulation. *Nature*, *392*, 387-390.
- Ruchkin, D. S., Grafman, J., Cameron, K., & Berndt, D. S. (2003). Working memory retention systems: A state of activated long-term memory. *Behavioral and Brain Sciences*, *26*, 709-777.
- Sakurai, Y., Takahashi, S., & Inoue, M. (2004). Stimulus duration in working memory is represented by neuronal activity in the monkey prefrontal cortex. *European Journal of Neuroscience*, *20*, 1069-1080.
- Santi, A., Hoover, C., & Simmons, S. (2011). Rats' memory for time and relational responding in the duration-comparison procedure. *Learning and Motivation*, *42*, 173-184.
- Sarma, A., Masse, N. Y., Wang, X.-J., & Freedman, D. J. (2016). Task-specific versus generalized mnemonic representations in parietal and prefrontal cortices. *Nature*

- Neuroscience*, 19, 143-149.
- Savolainen, P., Carlson, S., Boldt, R., Neuvonen, T., Hannula, H., Hiltunen, J., Salonen, O., Ma, Y.-Y., & Pertovaara, A. (2011). Facilitation of tactile working memory by top-down suppression from prefrontal to primary somatosensory cortex during sensory interference. *Behavioural Brain Research*, 219, 387-390.
- Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science*, 20, 207-214.
- Sereno, M. I., Pitzalis, S., & Martinez, A. (2001). Mapping of contralateral space in retinotopic coordinates by a parietal cortical area in humans. *Science*, 294, 1350-1354.
- 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, 21341-21346.
- Silvanto, J., & Muggleton, N. G. (2008). New light through old windows: Moving beyond the "virtual lesion" approach to transcranial magnetic stimulation. *NeuroImage*, 39, 549-552.
- Sinclair, R. J., & Burton, H. (1996). Discrimination of vibrotactile frequencies in a delayed pair comparison task. *Perception & Psychophysics*, 58, 680-692.
- Spetch, M. L., & Wilkie, D. M. (1983). Subjective shortening: A model of pigeons' memory for event duration. *Journal of Experimental Psychology: Animal Behavior Processes*, 9, 14-30.
- Spitzer, B., & Blankenburg, F. (2011). Stimulus-dependent EEG activity reflects internal updating of tactile working memory in humans. *Proceedings of the National Academy of*

- Sciences, 108*, 8444-8449.
- Spitzer, B., & Blankenburg, F. (2012). Supramodal parametric working memory processing in humans. *Journal of Neuroscience, 32*, 3287-3295.
- Spitzer, B., Fleck, S., & Blankenburg, F. (2014). Parametric alpha- and beta-band signatures of supramodal numerosity information in human working memory. *Journal of Neuroscience, 34*, 4293-4302.
- Spitzer, B., Wacker, E., & Blankenburg, F. (2010). Oscillatory correlates of vibrotactile frequency processing in human working memory. *Journal of Neuroscience, 30*, 4496-4502.
- Spitzer, B., Gloel, M., Schmidt, T. T., & Blankenburg, F. (2014). Working memory coding of analog stimulus properties in the human prefrontal cortex. *Cerebral Cortex, 24*, 2229-2236.
- Stewart, L. M., Walsh, V., & Rothwell, J. C. (2001). Motor and phosphene thresholds: A transcranial magnetic stimulation correlation study. *Neuropsychologia, 39*, 415-419.
- Stokes, M. G. (2015). 'Activity-silent' working memory in prefrontal cortex: A dynamic coding framework. *Trends in Cognitive Sciences, 19*, 394-405.
- Strafella, A. P., Vanderwerf, Y., & Sadikot, A. F. (2004). Transcranial magnetic stimulation of the human motor cortex influences the neuronal activity of subthalamic nucleus. *European Journal of Neuroscience, 20*, 2245-2249.
- Sugishita, M., & Takayama, Y. (1993). Paraesthesia elicited by repetitive magnetic stimulation of the postcentral gyrus. *NeuroReport, 4*, 569-570.
- Thorndike, E. L. (1910). The relation between memory for words and memory for numbers, and

- the relation between memory over short and memory over long intervals. *American Journal of Psychology*, *21*, 487-488.
- Tong, F., & Pratte, M. S. (2012). Decoding patterns of human brain activity. *Annual Review of Psychology*, *63*, 483-509.
- Vallar, G., & Baddeley, A. D. (1984). Fractionation of working memory: Neuropsychological evidence for a phonological short-term store. *Journal of Verbal Learning and Verbal Behavior*, *23*, 151-161.
- Van Rooyen, P., & Santi, A. (2009). Pigeons' memory for time: Assessment of the role of subjective shortening in the duration-comparison procedure. *Learning & Behavior*, *37*, 74-84.
- Vergara, J., Rivera, N., Rossi-Pool, R., & Romo, R. (2016). A neural parametric code for storing information of more than one sensory modality in working memory. *Neuron*, *89*, 1-9.
- Vigliocco, G., Warren, J., Siri, S., Archiuli, J., Scott, S., & Wise, R. (2006). The role of semantics and grammatical class in the neural representation of words. *Cerebral Cortex*, *16*, 1790-1796.
- Wang, X-J. (2001). Synaptic reverberation underlying mnemonic persistent activity. *Trends in Neurosciences*, *24*, 455-463.
- Wearden, J. H., & Ferrara, A. (1993). Subjective shortening in humans' memory for stimulus duration. *Quarterly Journal of Experimental Psychology B: Comparative and Physiological Psychology*, *46*, 163-186.
- Wearden, J. H., Parry, A., & Stamp, L. (2002). Is subjective shortening in human memory

- unique to time representations? *Quarterly Journal of Experimental Psychology B: Comparative and Physiological Psychology*, 55, 1-25.
- Wearden, J. H., Goodson, G., & Foran, K. (2007). Subjective shortening with filled and unfilled auditory and visual intervals in humans? *Quarterly Journal of Experimental Psychology*, 60, 1616-1628.
- Weinberger, N. M. (2004). Specific long-term memory traces in primary auditory cortex. *Nature Reviews Neuroscience*, 5, 279-290.
- Wilson, F. A., O Scalaidhe, S. P., & Goldman-Rakic, P. S. (1993). Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science*, 260, 1955-1958.
- Yu, E. C., Sprenger, A. M., Thomas, R. P., & Dougherty, M. R. (2014). When decision heuristics and science collide. *Psychonomic Bulletin & Review*, 21, 268-282.
- Zhou, Y. D., & Fuster, J. M. (1996). Mnemonic neuronal activity in somatosensory cortex. *Proceedings of the National Academy of Sciences*, 93, 10533-10537.

### **Appendix A: Copyright information for manuscripts**

Chapter 2: Copyright 2014, Psychonomic Society, Inc. “Author retains the right to use his/her article for his/her further scientific career by including the final published journal article in other publications such as dissertations and postdoctoral qualifications provided acknowledgement is given to the original source of publication.”

Chapter 3: Copyright 2014, Bancroft, Hogeveen, Hockley, and Servos. The manuscript is licensed under the Creative Commons Attribution License (CC BY).

Chapter 4: Copyright 2014, Canadian Psychological Association. "In consideration of the assignment of copyright, the CPA hereby grants to the Authors, and the employers for whom this work was performed (if applicable), royalty-free, non-exclusive, limited licenses to: 1. Reproduce, or have reproduced, the above manuscript for the author(s)' personal use or for the author(s)' intracompany use provided that (a) the source and CPA copyright notice are indicated and (b) the copies are not used in a way that implies CPA endorsement of a product or service."

Chapter 5: Copyright 2015, Psychonomic Society, Inc. “Author retains the right to use his/her article for his/her further scientific career by including the final published journal article in other publications such as dissertations and postdoctoral qualifications provided acknowledgement is given to the original source of publication.”

### Appendix B: Collaborator contributions to Manuscripts

**Bancroft, T. D., Hockley, W. E., & Servos, P. (2014). Does stimulus complexity determine whether working and short-term memory storage relies on prefrontal cortex or sensory cortex? *Attention, Perception, & Psychophysics*, 76, 1954-1961.**

Bancroft: Primary author on all drafts, wrote majority of paper. Devised original idea.

Hockley: Contributed intellectually and contributed to writing.

Servos: Contributed intellectually and contributed to writing.

**Bancroft, T. D., Hogeveen, J., Servos, P., & Hockley, W. E. (2014). TMS-induced neural noise in sensory cortex interferes with short-term memory storage in prefrontal cortex. *Frontiers in Computational Neuroscience*, 8, 23.**

Bancroft: Primary author on all drafts and wrote majority of paper. Primarily responsible for the design, coding, and execution of all simulations. Devised original idea.

Hogeveen: Contributed intellectually and to writing, including some material to manuscript.

Contributed to project planning, conceptualization, and interpretation of results.

Hockley: Contributed intellectually and contributed to writing.

Servos: Contributed intellectually and contributed to writing.

**Bancroft, T. D., Hockley, W. E., & Servos, P. (2014). A shared short-term memory system for stimulus frequency and duration. *Canadian Journal of Experimental Psychology*, 68, 236-241.**

Bancroft: Primary author on all drafts and wrote majority of paper. Primarily responsible for the design, coding, and execution of all simulations. Devised original idea.

Hockley: Contributed intellectually and contributed to writing.

Servos: Contributed intellectually and contributed to writing.

**Bancroft, T. D., Jones, J. A., Ensor, T. M., Hockley, W. E., & Servos, P. (in press).**

**Overwriting and intrusion in short-term memory. *Memory & Cognition*, 44, 435-443.**

Bancroft: Primary author on all drafts and wrote majority of paper. Primarily responsible for the design and execution of all experiments and analysis and interpretation of results. Devised original idea.

Ensor: Contributed intellectually and to experimental design. Performed confirmatory data analysis and contributed  $W$  values. Contributed to writing, including sections in the introduction and discussion.

Hockley: Contributed intellectually, contributed to experimental design, analysis, and interpretation of results, and contributed to writing.

Jones: Contributed intellectually, contributed to experimental design and interpretation of results, and contributed to writing.

Servos: Contributed intellectually, contributed to experimental design and interpretation of results, and contributed to writing.