

**ON THE REPRESENTATION OF SPATIAL AND TEMPORAL STRUCTURES:  
EFFECTS IN HUMAN VISUOSPATIAL WORKING MEMORY**

by

Tara Reneé Ghazi

A dissertation submitted to Johns Hopkins University in conformity with the  
requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

January 2022

© 2022 Tara Ghazi

All rights reserved

# Abstract

In a diverse range of environments, each replete with unique physical phenomena, humans are capable of acting and achieving with volition. To do so we capitalize upon structures that exist in the physical world, rapidly drawing associations and forming conceptual relationships between items and occurrences. In this dissertation work, I examine how structures in the domains of space and time impact the representations of information that we form and hold in working memory, in the service of goal-driven behavior. Three key findings arise from the studies I present herein.

First, representation of spatial structures in working memory is supported by oscillatory neural activity that *differs* between individuals based upon biological sex. The peak of posterior alpha frequency oscillatory activity is modulated in support of visuospatial representation maintenance more so in females than males. Among males but not females, successful representation of relative spatial structure is positively tied to an individual's peak frequency of alpha oscillatory activity.

Second, the interaction of spatial and temporal structures *across perceptual modalities* impacts representation in working memory. Shared temporal structure between a stream of visual targets and a stream of sounds promotes representation of the spatial structure of those sounds. This integration of perceptual information occurs whether helpful or harmful, differentially impacting performance.

Third, the representation of spatial information in working memory is impacted by a particular form of temporal structure — rhythm. The presence of rhythmic versus arrhythmic temporal structure within a visuospatial stream does not increase the precision of working memory representation, but rather increases the speed with which representations may be formed. Rhythmic structure spontaneously and consistently facilitates working memory performance. Arrhythmic structure may hinder temporal processing but can be behaviorally compensated for with the application of controlled attention to the temporal domain. A novel paradigm, designed and utilized to study effects of rhythmic temporal structure upon visuospatial working memory is described.

*Committee Members:*

Dr. Susan Courtney (Advisor and First Reader), Dr. Howard Egeth (Second Reader),

Dr. Christopher Honey, Dr. Richard Ivry, Dr. Jason Fischer, Dr. John Desmond and

Dr. Brenda Rapp (Chair)

# Acknowledgements

There has been no more significant relationship developed over the course of my Ph.D. than that with my Advisor, Dr. Susan Courtney. You granted me both a place, and the space, to grow into the scientist I wished to become. Thank you for these years of patient counsel, support, instruction, and guidance.

Upon my arrival in the Courtney Lab, I was met with kind support and friendship from my then lab-mates. Thank you to Dr. Thomas Hinault for generously sharing your time, advice, and expertise in EEG methodology. Thank you to Dr. Michelle DiBartolo and her wife Stephanie for the soft landing in Baltimore and for the reminders of what great need there is for rigorous scientific contributions that can be applied toward helping people and society. Joint meetings with the laboratories of Dr. Howard Egeth and Dr. Jason Fischer provided an open and welcoming environment in which to share, critique and learn from one another. Thank you to the members of these labs, to Dr. Ian Phillips, Dr. Ana Navarro, and in particular Dr. Sarah Cormeia for the ever positive and generous contribution of ideas. Thank you to Dr. Janice Chen and Dr. Christopher Honey for your down to earth advice at pivotal moments. I thank Dr. Egeth and Dr. Honey for expressing interest, asking questions, and providing helpful commentary about my research throughout my PhD.

For the years of training I received in the Cognitive Neuroscience of Aging, I am grateful to Dr. Marilyn Albert and Dr. George Rebok. Thank you to the past and present

members of Dr. Michaela Gallagher's laboratory for their hosting of the Aging Journal Club in which I have enjoyed great discussions and learned much.

My education preceding my arrival at Johns Hopkins University was shaped both by the mentorship of specific individuals as well as the guiding philosophies of the institutions I was able to engage with as a student. I am forever grateful to the Public College and University systems of my home State of California for empowering me to pursue and achieve what was once, for me, only a dream. The lesson that there is little I am entitled to but much I can claim has prompted me to extend and engage often. Thank you to my first academic mentor, Dr. Robin Roffey at Sacramento City College, for challenging me to grow as a student and for cultivating the then tiny and seemingly impossible hope I had of pursuing a PhD. At Berkeley, I benefitted immensely from the efforts of Dr. Ani Adhikari — a Professor devoted to educating. I am grateful to my Professors at Berkeley, Dr. Eve Sweetser and Dr. Terrence Deacon, who generously took me on as a mentee, offering support, guidance, and inclusion. I would not have come this far without them. Thank you to my research Supervisor, Dr. John Williams at the University of Cambridge, for the kind direction and practical advice.

To my family and close friends, thank you for your love and support through the challenges, and for sharing with me in celebration of the successes.

To the members of my committee, my thanks for your agreement to take part in this milestone. I thank you for your time, your reading, and your energies spent.

# Dedication

I dedicate this dissertation in part to my grandmothers Carmen and Margaret.

Their loving words, stories, and confidences have been sources of  
inspiration to me throughout this journey.

I also dedicate this dissertation to my husband, Peymon.

A visionary scientist, a loving father, the best partner one could hope for.

Your support and sacrifice as we embarked and daily partnership as we travel  
together on this *path less taken* make all the difference.

# Table of Contents

<b>Abstract.....</b>	<b>ii</b>
<b>Acknowledgements.....</b>	<b>iv</b>
<b>Dedication .....</b>	<b>vi</b>
<b>List of Tables.....</b>	<b>x</b>
<b>List of Figures.....</b>	<b>xi</b>
<b>Chapter 1: Introduction .....</b>	<b>1</b>
1.1 Background .....	2
1.2 Overview of Studies .....	4
1.2.1 Sex differences in the representation of structure.....	4
1.2.2 Bimodal structure impacts representation.....	5
1.2.3 Representation in the face of rhythm.....	7
1.3 References .....	8
<b>Chapter 2: Human Working Memory .....</b>	<b>12</b>
2.1 Introduction .....	12
2.2 Models of Working Memory.....	13
2.2.1 Cognitive Models .....	15
2.2.2 Neuroscientific Models .....	17
2.3 From Perception to Memory: Reducing Information Load.....	21
2.3.1 Biased Competition.....	23
2.3.2 Selective Attention.....	26
2.3.3 Strategies .....	30
2.4 The Storage and Processing of Representations .....	33
2.4.1 Interference and Oscillatory Activity .....	35
2.5 Open Questions .....	38
2.6 References .....	40
<b>Chapter 3: Modulation of Peak Alpha Frequency Oscillations during Working Memory is greater in Females than Males .....</b>	<b>49</b>
3.1 Introduction .....	49
3.1.1 Alpha Oscillations and Working Memory .....	50
3.1.2 Peak Alpha Frequency (PAF) .....	51
3.1.3 Influences of Sex on PAF .....	52

3.1.4 PAF and Cognitive Performance .....	53
3.1.5 Study Question.....	54
3.2 Methods.....	55
3.2.1 Participants .....	56
3.2.2 Working Memory Task.....	56
3.2.3 EEG Collection, and Pre-Processing .....	58
3.2.4 Peak Alpha Frequency.....	59
3.2.5 Statistical Analysis.....	62
3.3 Results.....	65
3.3.1 Task Performance .....	65
3.3.2 Peak Alpha Frequency within each Time Period.....	66
3.3.3 Modulation of Peak Alpha Frequency.....	67
3.3.4 Peak Alpha Frequency and Behavior .....	67
3.4 Discussion.....	69
3.5 Figures.....	75
3.6 References .....	82
<b>Chapter 4: Bimodal Spatial and Temporal Influences on Working Memory Performance.....</b>	<b>88</b>
4.1 Introduction .....	89
4.2 Methods.....	98
4.2.1 Participants .....	98
4.2.2 Stimuli and Task .....	99
4.2.3 Procedure.....	103
4.2.4 Design.....	103
4.3 Results.....	105
4.4 Discussion.....	109
4.5 Conclusion.....	114
4.6 Figures.....	115
4.7 References .....	119
<b>Chapter 5: Spatial Working Memory in the Presence or Absence of Visual Rhythm .....</b>	<b>124</b>
5.1 Introduction .....	124
5.2 Methods.....	130
5.2.1 Participants .....	130
5.2.2 Stimuli and Task .....	131

5.2.3 Design.....	135
5.2.4 Procedure.....	137
5.2.5 Analysis .....	137
5.3 Results.....	139
5.3.1 Experiment 1 .....	139
5.3.2 Experiment 2.....	140
5.3.3 Between-Experiment Outcomes.....	141
5.4 Discussion.....	142
5.5 Conclusion.....	147
5.6 Figures.....	149
5.7 References .....	154
<b>Chapter 6: Conclusions and Future Work .....</b>	<b>159</b>
6.1 Discussion.....	159
6.2 References .....	164

# List of Tables

<b>TABLE 1:</b> MEASURES OF CENTRAL TENDENCY IN TASK PERFORMANCE BY GROUP (CHAPTER 3).....	77
<b>TABLE 2:</b> MEASURES OF CENTRAL TENDENCY IN PEAK ALPHA FREQUENCY BY GROUP (CHAPTER 3). .....	80
<b>TABLE 3:</b> CORRELATIONS BETWEEN ACCURACY AND PEAK ALPHA FREQUENCY (CHAPTER 3) .....	81
<b>TABLE 4:</b> BEHAVIORAL MEASURES OF CENTRAL TENDENCY (CHAPTER 5).....	153

# List of Figures

<b>FIGURE 1:</b> WORKING MEMORY TASK AND TIMELINE, SELECTED TIME PERIODS AND CHANNELS FOR ANALYSIS OF EEG. (CHAPTER 3).....	75
<b>FIGURE 2:</b> ACCURACY DISTRIBUTIONS BY GROUP AND COMPARISON BY SHIFT FUNCTION. (CHAPTER 3) .....	78
<b>FIGURE 3:</b> RESPONSE TIME DISTRIBUTIONS BY GROUP AND COMPARISON BY SHIFT FUNCTION. (CHAPTER 3) .....	79
<b>FIGURE 4:</b> DISTRIBUTIONS OF PAF MODULATION BY SEX. COMPARISON BETWEEN MALE AND FEMALE DISTRIBUTIONS BY SHIFT FUNCTION. (CHAPTER 3) .....	80
<b>FIGURE 5:</b> OVERVIEW OF TASK CONDITIONS AND TRIAL TIMELINE. (CHAPTER 4) .....	115
<b>FIGURE 6:</b> ACCURACY (PROPORTION CORRECT) WITHIN BIMODAL TRIALS BY SAMPLE CONDITION. (CHAPTER 4) .....	116
<b>FIGURE 7:</b> PERFORMANCE ACROSS BIMODAL TRIALS BY SAMPLE CONDITION AND TEST TYPE. (CHAPTER 4) .....	117
<b>FIGURE 8:</b> COMPARISON OF PARTICIPANT SENSITIVITY, $d'$ , AND DECISION CRITERION, $b$ , ACROSS CONDITIONS. (CHAPTER 4) .....	118
<b>FIGURE 9:</b> TASK AND GENERAL TIMELINE OF A SINGLE TRIAL. (CHAPTER 5) .....	149
<b>FIGURE 10:</b> BEHAVIORAL PERFORMANCE BY TERMINAL TARGET SET GROUPING ACROSS RHYTHMIC AND ARRHYTHMIC CONDITIONS FOR EXPERIMENTS 1 AND 2. (CHAPTER 5).....	150
<b>FIGURE 11:</b> DECISION TIME BY RHYTHMIC AND ARRHYTHMIC CONDITIONS FOR EXPERIMENTS 1 AND 2. (CHAPTER 5) .....	152

## Chapter 1

# Introduction

Information is everywhere in our environment. Limitations in our capacity to attend and retain, constrain what information we can access. Conceptualizing and extracting structure from our environment enables us to reduce a sea of potential information into drops that we can make use of. Relationships and regularities are two types of informational structures we can utilize. In space, we can attend to relationships that are formed between objects, creating singular structure from disparate things. In time, we can attend to events that are discrete, coincidental, or sequential – structures that may shape the way we conceptually bind events together. When extracting information from our environment to accomplish an immediate goal, it may be the case that our propensity to attend to these types of structure can in some contexts facilitate our performance, yet in others lead us astray.

Through a series of studies in this dissertation work, I explore the question of how different structures in and of information impact working memory performance. I constrain the scope of this examination to a small subset of informational structures that

may arise – namely those arising from relationships in space, regularities in time, and the intersection of the two.

## 1.1 Background

The human brain is a powerful organ. It has evolved such that complex forms of cognition arise from the multitude of ongoing operations it carries out. The joint processes described holistically as “working memory,” are of particular importance as they enable us to achieve goal-driven behavior. Key among these processes are our abilities to control and constrain what we focus upon, and to manipulate, convert, and update what we are holding in mind. These are general descriptions of processes that we are capable of - which may be applied given a broad range of inputs. Of interest, though, is how the structure of those inputs may impact the way these processes are carried out. Information exists “out there” in the world. The activation of specific neural circuitries and networks, and the result of the computations they carry out, are what enable that information to be “represented” within the brain (Marr, 1982). It may be the case, that where we focus our attention, and what representations we form, are impacted by the structures of, and in, the information available to us.

In the physical world, two (among many) dimensions in which information may be structured are those of space, and time. Proximity and positioning in space, for example, each provide cues as to whether connections or relationships may be present between seemingly disparate things. Cues such as these may be processed in a seemingly automatic fashion, with or without conscious effort or awareness (Morton et al., 2017).

Similarly, in the ever-progressing stream of experience, proximity and position in time may serve as cues as well (Michalka et al., 2015; Rohenkohl et al., 2014).

A defining feature across various working definitions, models, and theories of working memory is its limitation in capacity (Courtney, 2004; Cowan et al., 2005; Kane et al., 2004; Oberauer, 2002). Evidence has pointed towards two primary factors for this limitation. The first arises from the focal nature of attention - a key mechanism of working memory (Cowan et al., 2005; Oberauer, 2019; Shipstead et al., 2012). The second from the neural mechanisms by which information is temporarily maintained in the brain (de Vries et al., 2020; Lisman & Idiart, 1995; Liu et al., 2020; Miller et al., 2018; Stokes, 2015). These subjects are reviewed in detail in Chapter 2.

*What* we “hold,” or represent, in working memory, is strongly impacted by what we attend to (Oberauer, 2002). Some information we choose to attend to – it may be aligned with our goals; other information we cannot help but attend to – it may claim our attention in spite our goals (Egeth & Yantis, 1997). *How* we represent information in working memory, may result from the use of automatic, conscious, and unconscious strategies (Courtney et al., 1998). Neural processes that support either filtering relevant from irrelevant information, or amplification of some representations and suppression of others, for example, are critical and occur automatically (Desimone & Duncan, 1995; Ikkai et al., 2014). Consolidating, chunking, converting, and associating separate pieces of information, are all descriptors of strategies we have conscious access to, that enable us to maximize what we store in a limited capacity system (Cowan, 2001; Gonthier, 2020; Miller, 1956).

It may be the case that different structures of and within information, impact both what we represent and how we do so. In the series of studies described herein, I present evidence in support of this indeed being the case.

## 1.2 Overview of Studies

### 1.2.1 Sex differences in the representation of structure

In the first study I present, “Modulation of Peak Alpha Frequency Oscillations during Working Memory is greater in Females than Males,” focus is placed upon the question *how?* Specifically, how does the automatic processing of different structures in spatial information differ across individuals?

Oscillatory neural activity within the *alpha* range of frequencies, has been well tied to the suppression of irrelevant information during working memory (Blacker et al., 2016; Ikkai et al., 2014). Further, the specific frequency of peak activity within the *alpha* range, has been directly tied to working memory performance (Klimesch et al., 1993, 2003). During the course of a lifespan from early development through aging, different trajectories in the maturation, stabilization, and decline of peak alpha frequency have been documented between females and males (Aurlen et al., 2004; Chiang et al., 2011). Sex-specific cyclical variation in peak alpha frequency among females, has also been well described (Bazanov et al., 2014; Wuttke et al., 1975).

These prior findings led to the hypothesis of this study: alpha frequency activity differentially supports working memory for different structures of information based

upon sex. Utilizing an in-laboratory dataset in which participants performed a series of visuospatial working memory tasks while undergoing electroencephalography (EEG) recording, I found support for this hypothesis. Females and males were equally adept in task performance, with similar levels of accuracy in representing different spatial information structures. How task performance was supported, however, differed. Among females, considerable modulation of peak alpha frequency in support of task performance was observed. Among males, peak alpha frequency was particularly tied to working memory performance for relational structures between spatial objects. These findings were published (Ghazi et al., 2021) and are presented in Chapter 3.

### 1.2.2 Bimodal structure impacts representation

In the second study I present, “Bimodal Spatial and Temporal Influences on Working Memory Performance,” I investigate the manner in which structures that are situated not just in space, but also in time, can impact working memory performance. In this study and the next, temporal structure is parameterized by way of manipulating information occurrences within dynamic streams that unfold over time. In the previous study, evidence was presented for sex differences in the neural implementation of representing structure. This study marks a shift in focus towards examination of spatial and temporal structures impact on behavior. Specifically, when presented with multiple sources and structures of information that seemingly interact, what can we learn of the representations gated into working memory by analyzing resulting behavior?

What we attend to has an impact on what representations we will hold in working memory (Oberauer, 2002; Shipstead et al., 2012). Even if engaged with and focused on

performing visual task, sounds can capture and draw our attention involuntarily (Hughes, 2014; Marsh et al., 2020). This draw is not always negative, however. The effects of processing seemingly uninformative sounds are sometimes beneficial – speeding up performance, for example (Van der Burg et al., 2008). This type of positive effect may arise from specialized cells and networks in the brain that rapidly process spatial and temporal conjunctions across multiple sensory modalities (Knudsen, 2011). The processing of conjunctions across modalities, can also facilitate formation of more holistic representations of information (Kondo et al., 2017).

In this study it was hypothesized that the intersection of structures in bimodally perceived domains of space and time have differential effects on working memory performance. When bimodal structures coincide and provide information in compliment to one another, representation in working memory will be facilitated. When structures coincide but are in conflict, representation will be negatively impacted. This hypothesis was tested by presenting bimodal stimulus streams that varied in temporal synchrony and spatial congruence to participants and comparing the resulting outcome on their behavioral performance. As predicted, both positive and negative effects in task performance were observed, with evidence pointing towards the integration of bimodal information having impacted representation. Interestingly, however, negative effects were of greater magnitude than positive effects, indicating distraction may have had a substantial role in this outcome. This study and its findings are presented in Chapter 4.

### 1.2.3 Representation in the face of rhythm

The final study presented herein is “Spatial Working Memory in the Presence or Absence of Visual Rhythm.”

It has been well demonstrated via studies of attention and perception, that individuals attend to the temporal informational structure that is naturally present in rhythmic stimuli (Breska & Deouell, 2014; Buzsáki et al., 2013). Temporal informational structures of this form contribute both to prediction and expectation. Irrespective of specific interval durations, each can facilitate perception and detection of stimuli that occur when predicted or expected (Breska & Deouell, 2014; Nobre & van Ede, 2018).

While behavioral effects driven by prediction and expectation arising from temporal structure have been thoroughly described in literature focusing solely on attention, far fewer studies have been reported that examine their impact on working memory. This is interesting given that it is not uncommon for working memory paradigms to utilize temporal presentation parameters that are, indeed, rhythmic (e.g. Berch et al., 1998; Corsi, 1972; Gmeindl et al., 2011). In this study, a novel visuospatial working memory paradigm was designed to test explicitly *if*, and if so, *how*, rhythmic temporal structure impacts working memory performance.

Two separate experiments were performed in which behavioral performance of a visuospatial working memory task was assessed for streams of spatial information presented with either rhythmic, or arrhythmic, temporal structure. I present evidence that rhythmic and arrhythmic temporal structures can impact the timing of

representation accessibility in visual working memory. The observed effects, however, and the degree to which they were reduced based upon attention being explicitly applied to the temporal domain, indicate that the absence of rhythm can be overcome by processes of control. This study and its findings are presented in Chapter 5.

## 1.3 References

- Aurlen, H., Gjerde, I. O., Aarseth, J. H., Eldøen, G., Karlsen, B., Skeidsvoll, H., & Gilhus, N. E. (2004). EEG background activity described by a large computerized database. *Clinical Neurophysiology*, 115(3), 665–673. <https://doi.org/10.1016/j.clinph.2003.10.019>
- Bazanov, O. M., Kondratenko, A. V., Kuzminova, O. I., Muravlyova, K. B., & Petrova, S. E. (2014). EEG alpha indices depending on the menstrual cycle phase and salivary progesterone level. *Human Physiology*, 40(2), 140–148. <https://doi.org/10.1134/S0362119714020030>
- Berch, D. B., Krikorian, R., & Huha, E. M. (1998). The Corsi block-tapping task: Methodological and theoretical considerations. *Brain and Cognition*, 38(3), 317–338. <https://doi.org/10.1006/brcg.1998.1039>
- Blackner, K. J., Ikkai, A., Lakshmanan, B. M., Ewen, J. B., & Courtney, S. M. (2016). The role of alpha oscillations in deriving and maintaining spatial relations in working memory. *Cognitive, Affective, & Behavioral Neuroscience*, 16(5), 888–901. <https://doi.org/10.3758/s13415-016-0439-y>
- Breska, A., & Deouell, L. Y. (2014). Automatic Bias of Temporal Expectations following Temporally Regular Input Independently of High-level Temporal Expectation. *Journal of Cognitive Neuroscience*, 26, 1555–1571. <https://doi.org/10.1162/jocn>
- Buzsáki, G., Logothetis, N., & Singer, W. (2013). Scaling brain size, keeping timing: Evolutionary preservation of brain rhythms. *Neuron*, 80(3), 751–764. <https://doi.org/10.1016/j.neuron.2013.10.002>
- Chiang, A. K. I., Rennie, C. J., Robinson, P. A., van Albada, S. J., & Kerr, C. C. (2011). Age trends and sex differences of alpha rhythms including split alpha peaks. *Clinical Neurophysiology*, 122(8), 1505–1517. <https://doi.org/10.1016/j.clinph.2011.01.040>
- Corsi, P. M. (1972). *Human memory and the medial temporal region of the brain*. McGill University, Montreal, QB.
- Courtney, S. M. (2004). Attention and cognitive control as emergent properties of information representation in working memory. *Cognitive, Affective and Behavioral Neuroscience*, 4(4), 501–516.
- Courtney, S. M., Petit, L., Haxby, J. V., & Ungerleider, L. G. (1998). The role of prefrontal cortex in

- working memory: Examining the contents of consciousness. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 353(1377), 1819–1828.  
<https://doi.org/10.1098/rstb.1998.0334>
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24(1), 87–114.  
<https://doi.org/10.1017/S0140525X01003922>
- Cowan, Nelson, Elliott, E. M., Saults, S. J., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, 51(1), 42–100.  
<https://doi.org/10.1016/j.cogpsych.2004.12.001>
- de Vries, I. E. J., Slagter, H. A., & Olivers, C. N. L. (2020). Oscillatory Control over Representational States in Working Memory. *Trends in Cognitive Sciences*, 24(2), 150–162.  
<https://doi.org/10.1016/j.tics.2019.11.006>
- Desimone, R., & Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. *Annual Review of Neuroscience*, 18, 193–222. <https://doi.org/10.1146/annurev-psych-122414-033400>
- Egeth, H. E., & Yantis, S. (1997). Visual attention: control, representation, and time course. *Annual Review of Psychology*, 48, 269–297.  
<https://doi.org/10.1146/annurev.psych.48.1.269>
- Ghazi, T. R., Blacker, K. J., Hinault, T., & Courtney, S. M. (2021). Modulation of Peak Alpha Frequency Oscillations During Working Memory Is Greater in Females Than Males. *Frontiers in Human Neuroscience*, 15(April), 1–14.  
<https://doi.org/10.3389/fnhum.2021.626406>
- Gmeindl, L., Walsh, M., & Courtney, S. M. (2011). Binding serial order to representations in working memory: a spatial/verbal dissociation. *Memory & Cognition*, 39(1), 37–46.  
<https://doi.org/10.3758/s13421-010-0012-9>
- Gonthier, C. (2020). Charting the Diversity of Strategic Processes in Visuospatial Short-Term Memory. *Perspectives on Psychological Science*.  
<https://doi.org/10.1177/1745691620950697>
- Hughes, R. W. (2014). Auditory distraction: A duplex-mechanism account. *PsyCh Journal*, 3(1), 30–41. <https://doi.org/10.1002/pchj.44>
- Ikkai, A., Blacker, K. J., Lakshmanan, B. M., Ewen, J. B., & Courtney, S. M. (2014). Maintenance of relational information in working memory leads to suppression of the sensory cortex. *Journal of Neurophysiology*, 112(8), 1903–1915. <https://doi.org/10.1152/jn.00134.2014>
- Kane, M. J., Tuholski, S. W., Hambrick, D. Z., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, 133(2), 189–217. <https://doi.org/10.1037/0096-3445.133.2.189>
- Klimesch, W., Sauseng, P., & Gerloff, C. (2003). Enhancing cognitive performance with repetitive transcranial magnetic stimulation at human individual alpha frequency. *European Journal of Neuroscience*, 17(5), 1129–1133. <https://doi.org/10.1046/j.1460-9568.2003.02517.x>

- Klimesch, W., Schimke, H., & Pfurtscheller, G. (1993). Alpha frequency, cognitive load and memory performance. *Brain Topography*, 5(3), 241–251. <https://doi.org/10.1007/BF01128991>
- Knudsen, E. I. (2011). Control from below: The role of a midbrain network in spatial attention. *European Journal of Neuroscience*, 33(11), 1961–1972. <https://doi.org/10.1111/j.1460-9568.2011.07696.x>
- Kondo, H. M., Van Loon, A. M., Kawahara, J. I., & Moore, B. C. J. (2017). Auditory and visual scene analysis: An overview. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1714). <https://doi.org/10.1098/rstb.2016.0099>
- Lisman, J. E., & Idiart, M. A. P. (1995). Storage of  $7 \pm 2$  Short-Term Memories in Oscillatory Subcycles. *Science*, 267(5203), 1512–1515.
- Liu, J., Zhang, H., Yu, T., Ni, D., Ren, L., Yang, Q., Lu, B., Wang, D., Heinen, R., Axmacher, N., & Xue, G. (2020). Stable maintenance of multiple representational formats in human visual short-term memory. *Proceedings of the National Academy of Sciences of the United States of America*, 117(51), 32329–32339. <https://doi.org/10.1073/pnas.2006752117>
- Marr, D. (1982). The Philosophy and the Approach. In *Vision : a computational investigation into the human representation and processing of visual information* (pp. 8–38). San Francisco: W.H. Freeman.
- Marsh, J. E., Campbell, T. A., Vachon, F., Taylor, P. J., & Hughes, R. W. (2020). How the deployment of visual attention modulates auditory distraction. *Attention, Perception, and Psychophysics*, 82(1), 350–362. <https://doi.org/10.3758/s13414-019-01800-w>
- Michalka, S. W., Kong, L., Rosen, M. L., Shinn-Cunningham, B. G., & Somers, D. C. (2015). Short-Term Memory for Space and Time Flexibly Recruit Complementary Sensory-Biased Frontal Lobe Attention Networks. *Neuron*, 87(4), 882–892. <https://doi.org/10.1016/j.neuron.2015.07.028>
- Miller, E. K., Lundqvist, M., & Bastos, A. M. (2018). Working Memory 2.0. *Neuron*, 463–475. <https://doi.org/10.1016/j.neuron.2018.09.023>
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *The Psychological Review*, 63(2), 81–97. <https://doi.org/10.1177/001088049003100202>
- Morton, N. W., Sherrill, K. R., & Preston, A. R. (2017). Memory integration constructs maps of space, time, and concepts. *Current Opinion in Behavioral Sciences*, 17, 161–168. <https://doi.org/10.1016/j.cobeha.2017.08.007>
- Nobre, A. C., & van Ede, F. (2018). Anticipated moments: Temporal structure in attention. *Nature Reviews Neuroscience*, 19(1), 34–48. <https://doi.org/10.1038/nrn.2017.141>
- Oberauer, K. (2002). Access to Information in Working Memory: Exploring the Focus of Attention. *Journal of Experimental Psychology: Learning Memory and Cognition*, 28(3), 411–421. <https://doi.org/10.1037/0278-7393.28.3.411>
- Oberauer, K. (2019). Working Memory and Attention – A Conceptual Analysis and Review. *Journal of Cognition*, 2(1), 1–23. <https://doi.org/10.5334/joc.58>

- Rohenkohl, G., Gould, I. C., Pessoa, J., & Nobre, A. C. (2014). Combining spatial and temporal expectations to improve visual perception. *Journal of Vision*, 14(4), 1–13.  
<https://doi.org/10.1167/14.4.8>
- Shipstead, Z., Redick, T. S., Hicks, K. L., & Engle, R. W. (2012). The scope and control of attention as separate aspects of working memory. *Memory*, 20(6), 608–628.  
<https://doi.org/10.1080/09658211.2012.691519>
- Stokes, M. G. (2015). “Activity-silent” working memory in prefrontal cortex: A dynamic coding framework. *Trends in Cognitive Sciences*, 19(7), 394–405.  
<https://doi.org/10.1016/j.tics.2015.05.004>
- Van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W., & Theeuwes, J. (2008). Pip and Pop: Nonspatial Auditory Signals Improve Spatial Visual Search. *Journal of Experimental Psychology: Human Perception and Performance*, 34(5), 1053–1065.  
<https://doi.org/10.1037/0096-1523.34.5.1053>
- Wuttke, W., Arnold, P., Becker, D., Creutzfeldt, O., Langenstein, S., & Tirsch, W. (1975). Circulating hormones, EEG, and performance in psychological tests of women with and without oral contraceptives. *Psychoneuroendocrinology*, 1(2), 141–152.  
[https://doi.org/10.1016/0306-4530\(75\)90006-2](https://doi.org/10.1016/0306-4530(75)90006-2)

## Chapter 2

# Human Working Memory

## 2.1 Introduction

Humans have remarkable abilities, empowered by the brain. Not only can we extract and distill information from without and within, we can convert and manipulate, analyze and examine, retain or discard. Conceiving of the brain as a highly complex and sophisticated organ that processes information, we can describe some of the computations that are carried out by the brain as processes that form “representations” of information (Marr, 1982). Once formed, representations can be stored, accessed, and further processed – abilities which give rise to complex thoughts and behaviors. While *potential* information in the environment takes many forms, as humans we have access to only a constrained range. Beyond limitations arising from what our sensory systems are capable of detecting, there are further bottlenecks of access through which some, but not all, information may pass. Some information may be prioritized, some discarded, some focused upon, some passed over entirely. The need or motivation to accomplish or perform can impact which information is prioritized for further processing and representation. In turn, the representations we ultimately form, store, and access, impact

our future thoughts and behavior. Because our final output of thought and behavior is contingent upon what subset of information passes through these “bottlenecks” to ultimately be represented within the brain, it is important to understand how and whether information itself impacts this process.

To reason about links between input information and output human behavior, it is necessary to understand the cognitive processing that occurs between input and output. In other words, to reason about why a particular behavioral outcome resulted, it would be useful to consider the processes and processing leading up to it. Critical among the set of processes in the brain that enable access to and utilization of representations in the service of complex thought or behavior is *working memory*. In this review, then, while other cognitive processes are also considered, processes related to working memory are the subject of central focus. Of particular relevance to this dissertation are the cognitive control processes and properties of the stimulus input that facilitate the selection of the specific information allowed access to working memory, and its maintenance or updating.

## 2.2 Models of Working Memory

Working memory is a descriptive but singular term that encompasses a set of cognitive processes and mechanistic operations carried out within the brain (Miller et al., 2018). These processes include dynamic processing and conversion of information representations and the transient “short-term” storage of them (Courtney et al., 1997;

Cowan, 2008; Goldman-Rakic, 1995; Oberauer, 2002). At any given moment in the awake and conscious individual it might seem that a myriad of ongoing conversions from information as it exists in the environment, to information as represented by the brain, are taking place. In reality, only a small subset of the information available in our environment receives more than brief cursory neural processing (Kondo et al., 2017; Pressnitzer, 2011). Once represented in the brain, however, that information can then be both consciously and unconsciously made use of; it can be considered, manipulated, incorporated, or simply held to aid in the accomplishment of a goal.

Human memory has been the subject of scientific study for well over a century, but the term *working memory* wasn't coined and explicitly studied until about 60 years ago (Baddeley, 2010). Various models and conceptions of working memory began to be put forward in the 1960s and have continued to be reworked and adapted since. In 1968, Atkinson and Shiffrin first described an individual's "short-term memory store" as their *working memory* - a receiver of informational inputs from external sensory, and internal long-term memory sources. They proposed a framework within which targeted research questions could be built, and indeed were in the years that followed. At the core of their framework was the conception that all types of memory operated within two key dimensions – storage and control. While mechanisms of storage were deemed fixed and permanent, control processes were deemed dynamic and changing. Control processes were responsive to the goals of an individual, the instructions they followed, and the history of personal experience from which they might learn, adapt, and modify their behavior (Atkinson & Shiffrin, 1968). In this early conception, based upon a then limited

body of research, control and storage seemed woefully entangled. In the years and decades to follow, these two fundamental dimensions of memory would be teased apart and themselves subdivided into constituent mechanisms via cognitive and neuroscientific models.

### 2.2.1 Cognitive Models

In 1974, a *multicomponent model of working memory* was proposed (Baddeley & Hitch, 1974). Following Atkinson & Shiffrin, this model also gave a prominent role to attentional control processes in the short-term handling of information. Indeed, at visual center in the *multicomponent model* is a “central executive” that directs, allocates, and controls attentional processing. Unique in their model, at the time, was the designation of specific components for the handling of information dependent upon modality. For example, people often engage in an internal repetition or rehearsal of verbal information they are trying to hold in mind. The “Phonological loop” component was conceived of as being dedicated to the processing and temporary storage of auditory or verbalizable information; a system in which internal verbal rehearsal took place. A separate “Visuo-Spatial sketch pad” component was proposed as a system where processing and temporary storage of primarily visual information occurred. Whereas previously, working memory processes might have simply been conceived of as part of a unitary short-term memory system, this model accounted for findings in which the amounts of information and means of storing it differed based upon information modality (Cowan, 2008). It was apparent that all information was not being processed in the same way - working memory performance seemed to vary based upon information type.

Following the *multicomponent model*, an *Embedded Processes* model of working memory was described (Cowan, 1999). In this conception, processes of attention and long-term memory were assigned such critical roles in the working memory system, that storage mechanisms of working memory seemed to be an aspect of long-term memory rather than something separate from it. This view, or simply the language describing it, were modified to later clarify Cowan's conception of working memory as encompassing short-term memory storage coupled with the processing of those representations being temporarily stored (Cowan, 2008).

The *concentric model of working memory*, proposed not long after by Oberauer (Oberauer, 2002), built upon Cowan's conception, further centralizing and highlighting the role of attention. In the *concentric model*, working memory is described as a system of simultaneous storage and processing, whereby processing of only a small selection of representations at a given time can occur. This distinction was supported by empirical findings indicating that storing information in memory does not necessarily impact separate concurrent processing, but active processing of information already stored, does. Hence, passive storage versus active processing of that which is already stored, have distinct impacts. In this model a near infinite quantity of "inactive" representations can be housed in a long-term memory storage. A small subset of stored representations (approximately four) can be "activated" such that they are rendered accessible for processing. Of those accessible representations, only a single one of them can occupy the central "focus of attention" at a given time. The focus of attention enables intensive, conscious processing, to be carried out. With such strong emphasis placed on accessibility

for processing in this model, the role of “selection” of representations strongly impacts the outcome of working memory processing.

### 2.2.2 Neuroscientific Models

During the same span of time within which the above cognitive models were developed, neuroscientists were attempting to reveal the neural processes and pathways of working memory. In 1995, Patricia Goldman-Rakic reviewed findings then to date of the cellular ensembles and subregions in pre-frontal cortex (PFC) shown to have a critical role in working memory processes. It had been demonstrated that circuits involved in working memory were comprised of both excitatory and inhibitory neurons, and that dopamine might be an important neurotransmitter in initiating circuit level activity (Goldman-Rakic, 1995). At this time significant advances had been made in mapping both local circuits and long-range pathways between cortical and sub-cortical regions of the mammalian brain in non-human species (reviewed in Deacon, 1989). Insights gained from studies in humans and non-human primates in particular pointed towards the significance of PFC as a controller and coordinator in working memory processing.

Informed by these findings, and soon thereafter, a theoretical framework formalizing the critical role of PFC in cognitive control was presented (Miller & Cohen, 2001). That prefrontal cortex was anatomically situated to be extensively and vastly connected with other brain regions, coupled with behavioral and neuroimaging findings, pointed toward its critical role in cognitive control processes. As working memory specifically involves control over stored information, this framework has direct application. In the *Integrative Theory of Prefrontal Cortex Function* (Miller & Cohen, 2001),

a primary function of PFC in control processes related to working memory is to maintain goal representations. Maintenance of an abstract type of representation such as a goal was conceived to be carried out via combined mechanisms that allowed both flexibility in formation and updating of goals, as well as stable holding of them until a desired outcome was achieved. The numerous dopamine delivery pathways to PFC provided a mechanism by which circuits might be “opened” to instantiate a representation, and “closed” again to maintain it. This opening and closing was described as “gating” of information, and this terminology used explicitly in subsequent “Gating models” of working memory (Badre, 2012; O’Reilly & Frank, 2006).

Though some areas in PFC had been found to be generally active during control and maintenance processes (*e.g.* Cohen et al., 1997; Courtney et al., 1997), some regions within PFC were also found to be selective for specific types of representations. Dorsal and ventral pathways for processing of visual information based upon, respectively, whether that information was primarily spatial- or object-based had then been well described (Ingle et al., 1982; Mishkin & Ungerleider, 1982). The *what versus where* visual information pathways (Ungerleider et al., 1998) motivated studying whether parallel organization existed in PFC. Indeed similar divisions in representation processing within PFC were found, with dorsal regions of bilateral superior frontal sulcus exhibiting increased activation for spatial over identity information, and the more ventral middle and inferior frontal gyri exhibiting increased activation for identity information over spatial (Courtney, Petit, Maisog, et al., 1998). Comparative analysis of studies in humans with those in non-human primates further reinforced representation-specific

organization in primate PFC, with dorsal PFC processing of spatial information and more ventral PFC processing of object information (Courtney, Petit, Haxby, et al., 1998). Divisions of labor between PFC regions do not stop at *what* and *where*. Activity patterns across different neuronal populations within PFC pointed towards their specialization for specific types of processing, be it the features of items, their spatial positioning, or their broader semantic classification (reviewed in Goldman-Rakic, 1995). For a given type of informational input into working memory, an extensive amount of detail pertaining to that input can be extracted. In 2004, Courtney synthesized the then findings to date of the functional organization within and between PFC and other cortical regions based upon information type. Complex mappings and cross-regional activations for information processing were extensively discussed in this review (Courtney, 2004). Rising to the surface, was the importance of the holistic processing made possible via interactions *between* regions specialized by information-representation-type.

Efforts to further understand working memory processes through the lens of brain structure and function were fruitful. Building upon Miller and Cohen's synthesis of findings related to PFC function, and incorporating evidence about dopamine circuitry and connectivity between the PFC and basal ganglia, O'Reilly and Frank presented a detailed model of working memory function (O'Reilly & Frank, 2006). Whereas many of the models discussed above focused upon synthesizing behavioral evidence to tease apart mechanisms, the model they proposed was motivated by a computational approach and perspective. In their *prefrontal cortex and basal ganglia model of working memory* – the PBWM model – an “actor-critic architecture” is instantiated via ongoing interaction

between specific cortical and sub-cortical regions. This model is strongly grounded in the physiology of the brain yet motivated by what O'Reilly and Frank describe as a key problem a working memory system must solve: determining what information to maintain versus what to forget.

O'Reilly and Frank use the term, introduced earlier, “gating” as an analogue to mechanisms of representation transfer in and out of short-term storage within PFC. A situation that allows working memory representations to be *passed into* PFC is analogous to an open gate. Temporary storage of said representations, with fidelity, takes place when this gate is “closed.” Should the need to discard or overwrite stored representations arise, the gate is opened again. In formalizing the PBWM model, O'Reilly and Frank delve into how a neural system in particular might not only carry out gating, but also learn and adapt in the timing of when to *open* versus *shut* the gate. Note that the anthropomorphizing analogies and descriptions are used for conceptual understanding rather than implication of neural substrates as having some type of “agency” in these processes. In presenting reasoning for the PBWM model, it is argued that the same neural substrate cannot both carry out the act of storing representations as well as judge and teach itself if it did so to the desired effect. Hence, two separate but interacting systems are required. An “actor” – instantiated in PFC – allows information to be stored, discarded, or updated. A “critic” – instantiated in Basal Ganglia – learns from the outcome of the actor’s actions, and then trains the actor via dopamine signaling pathways accordingly (O'Reilly & Frank, 2006). What is not covered in these models, however, is the nature of

the information and the timing of its arrival at the gate, that might make it more or less likely to be stored or maintained.

## 2.3 From Perception to Memory: Reducing Information Load

Organisms with sensory capacities have means by which information in the environment can be transduced into forms that are meaningful, relevant, and useful. Within humans, perception is ongoing via the senses of vision, audition, and interoception among others. Working memory involves processing of information represented *in the brain*. Perceptual processes are an impactful precursor which contribute to how those information representations are formed. There is not, however, a simple and straightforward way to discretize the continuum between sensation, perception, and working memory representation into stages of processing. In part, this is because of how rapidly feed-forward signals are sent from the earliest regions of sensory processing to cortical regions of goal-oriented control, and because of the ongoing feed-back signaling from goal-oriented control regions which in turn bias early-stage sensory processing.

Simply introspecting on one's own lived experience makes it apparent that not all information "out there" and potentially available is actually being perceived. In 1964, Anne Treisman reviewed a then prominent model proposed by Broadbent which attempted to account for this (Treisman, 1964). In Broadbent's *Filter and Information Flow* model, sensory organs are able to transduce a vast amount of information, but that

information is immediately directed through a filtering process in which some information passes through to be further processed while other information is filtered out. Filtration was conceived of as a means by which powerful but limited neural resources would be dedicated toward the processing of a subset of information selected from the larger pool of what is fundamentally perceived.

Studies demonstrating, for example, that participants have a difficult time distinguishing separate auditory messages that originated from the same spatial location, or that people become habituated and less responsive to the same information presented repeatedly, provided support for the idea of a filtering system which selects based upon physical properties of incoming information. Treisman (Treisman, 1964) discussed evidence indicating that physical properties are actually not filtered out, but more conceptual properties may be. For instance, studies had shown that though focusing attention on one channel of auditory information (e.g., words presented to one ear), participants were able to report physical property features about information from a separate, unattended channel, but were *not* able to report or identify semantic content such as words (Treisman, 1964). If filtration means the complete removal of information so that it cannot be further processed, key inferences can be made from this kind of example. First, if filtration is indeed what was occurring, then it took place between, rather than before, the processing of physical features and semantic meaning. Second, the information being filtered out at this “between stage” resided in a channel not explicitly being attended to. This pointed towards attention being a process that might preserve information, keeping it from being filtered out.

### 2.3.1 Biased Competition

Filtration of information occurs not just early, but continually during processing and has a significant impact on working memory. Desimone and Duncan (1995) described filtration in the visual domain as the outcome of ongoing “biased competition” occurring throughout neural resources dedicated to early visual and later conceptual processing. The concept of “gating,” discussed earlier in the context of working memory processing within PFC, is also employed by Desimone and Duncan to describe how bias might be accomplished within the brain. They describe two forms of circuitry gating, wherein resources that process some subset of incoming information can be cut off, or the influence of a particular cell population performing some processing operation reduced (Desimone & Duncan, 1995). An alternate means of generating bias is through processing that might generate response gain, or increased neural response, within a particular population of neurons such that its signal dominates that of another population. Further stimulating neurons in a population that are already in a state of “preparatory activation” might lead to this, and has been demonstrated to occur with the direction of internal attention (Stokes & Duncan, 2014). Amplification or gain in itself, however, does not necessarily produce bias (Hillyard et al., 1998). Competition in both early and later stages of processing can be biased by attention. Attention is understood to play a critical role in resolving competition – the act of directing attention to one of two identical stimuli, for example, can break a competitive tie (Beck & Kastner, 2014; Desimone & Duncan, 1995).

Desimone and Duncan place significant focus on how the application of attention to a visual object representation can bias competition. Competitive processing, however,

is not restricted to vision, but is understood to occur throughout the brain (Miller & Cohen, 2001). Beginning with processes of perception, competition is necessitated by the availability of a finite number of neural resources with processing capabilities that are constrained by time. Access to increased neural resources, for example having a brain with greater versus lesser volume or containing more versus fewer neuronal cells, does not resolve this constraint (Deacon, 1990). At every stage of processing, biases tilt favor towards some information at the expense of others. Fecteau and Munoz (2006) frame this as a competition for *priority*. Dependent upon context, dependent upon stimuli, and dependent upon goals – priorities differ. Competition begins at perception – where cells in early sensory cortices may be tuned and therefore preferentially responsive to specific types of stimuli. Rather than resolving at this early stage of processing, however, competition impacts subsequent stages as well (Beck & Kastner, 2014). Because these processes can have a compounding influence with an ultimate impact on behavior, understanding what types of factors may contribute to positive or negative directions of bias in processing is important.

In studies of attention, biases have been found to arise from properties that are specific to the stimuli within an environment, or factors that are internal and specific to an individual. When stimulus-specific properties are responsible for generating attentional bias, these have been termed as “bottom-up” factors (Desimone & Duncan, 1995; Egeth & Yantis, 1997). In the visual domain an example of a temporal and spatial *bottom-up* factor is “newness” – or the appearance of something novel in a scene. For example, it has been shown that when performing a visual search task, participants are

faster to report targets that abruptly appear in previously unoccupied locations and slower to report targets that replaced pre-existing stimuli (Yantis & Jonides, 1984). This impact on behavioral performance – an individual responding faster – is understood to be the result of bottom-up factors influencing cognitive processes. The “salience” of a particular stimulus can also generate a bias in processing where the more salient an object, the more positive bias may be generated for its processing. Salience, in this context, is defined as “physical, bottom-up distinctiveness... that depends on the relationship of one object with respect to other objects in the scene” (Fecteau & Munoz, 2006).

If one is using the descriptor *bottom-up* it is not uncommon to apply the opposing term *top-down* to capture all other biasing factors that are internally generated and specific to the individual doing the processing. Specifically, factors under the “volitional control” of an individual are described as *top-down* (Theeuwes, 2018). Usage of these terms, *bottom-up* and *top-down*, however, has been critiqued for comprising an oversimplified and misleading “theoretical dichotomy” (Awh et al., 2012). This critique arises because these terms have been long and frequently used (e.g. Buschman & Miller, 2007; Gazzaley et al., 2005; Gazzaley & Nobre, 2012; Shinn-Cunningham et al., 2005), but may lack a satisfactory degree of precision in light of more recent evidence that effects related to the recent history and context of experience exert bias that aren’t quite “top-down,” and aren’t quite “bottom-up.” Further contributing to this worry is that some sources of bias may be attributed to *top-down* causes by default, simply because they don’t seem to fit the definition of *bottom-up* (Awh et al., 2012). Rather than using *top-*

*down* as an umbrella catch-all term without distinguishing between control processes and effects of history, then, herein I will simply distinguish between biasing factors that are *exogenous*, or generated from without, from those that are *endogenous*, or generated within the individual (Rohenkohl et al., 2011; and e.g. van Ede et al., 2017).

In studies of working memory, participants are typically given explicit instructions regarding what content they should attempt to remember. In Courtney *et al.* (1998), for example, participants viewed square images containing different faces, located in different positions on-screen over the course of many trials. Participants were instructed to remember the face they saw on some trials, while on other trials the face was not important - they were to instead remember the location of the face image (Courtney, Petit, Maisog, et al., 1998). For the two different types of trial described in this example, participants had different goals regarding the type of information to remember for performing the task at hand. Goals and motivations, even if prompted by external instruction, generate *endogenous* biasing in information processing (Awh et al., 2012; Theeuwes, 2014).

### 2.3.2 Selective Attention

Exogenous and endogenous factors exert continual impacts on processing, having a particular impact on which items are “selected” for attending to. Desimone and Duncan (1995) defined selectivity as “the ability to filter out unwanted information.” In the attention domain selection is understood as the prioritization of some information over others (Myers et al., 2017). Neither is more correct than the other. There are two sides in the competition for selection: some information is subjected to further processing; other

information less. The selection of some information representations over others is conceptualized as being directed by an internally generated “priority map” that incorporates both exogenous and endogenous factors (Bisley & Goldberg, 2010; Wolfe, 2021). The parietal lobe in particular is thought to be a region in which exogenous and endogenous representations converge and undergo signal gain or attenuation subject to the influence of a priority map. The outcome impacts direction of attention (Bisley & Goldberg, 2010), which the parietal lobe has a significant role in enacting (Yantis et al., 2002).

Selective attention has long been understood to have a critical role in working memory processes. The so-called “focus of attention” features prominently in both the *embedded processes model* (Nelson Cowan, 1999) and the *concentric model* (Oberauer, 2002) of working memory described earlier, for example. In Oberauer (2019) a “Taxonomy of Attention” is presented which distinguishes between attention that is *selective* and driven by goals or intentions, from attention that is the result of *capture* by the salient or unexpected. Selective attention in particular is understood to be limited in the “number of channels” of information that can be attended to in parallel (Oberauer, 2019). This limitation in attentional capacity is strongly linked to limitations in working memory storage capacity (Cowan et al., 2005), and reflects a now deeper understanding of the stages at which information may become filtered.

In Broadbent’s mid-20<sup>th</sup> century *Filter and Information Flow* model, discussed earlier, selective attention could be applied to only a single item of information at a time. An updated understanding, is that multiple low-level channels of processing can occur in

parallel, but more rich processing is indeed restricted to items that are selectively attended to (Lachter et al., 2004). In frameworks that refer to an early sensory or perceptual representation as a “trace,” the application of selective attention to a trace is thought to initiate the consolidation of that trace into a short-term working memory representation (Jolicoeur & Dell’Acqua, 1998; Ricker, 2015). These frameworks can lead to the conceptualization of information flowing from perception and into memory as a linear sequence, with information either being selected and consolidated, or lost. Processes of selection, however, may occur both before and after information is encoded into working memory, differentiated by behavioral goal (Panichello & Buschman, 2021). In the context of working memory, selective attention is understood to play a role not just in determining what information should be encoded in working memory, but also in prioritizing among or selecting between those representations that have already been stored (van Ede & Nobre, 2021).

As discussed earlier, selective processing can involve the amplification, or the suppression, of information. In the visual domain it has been well demonstrated that neural populations oscillating within a specific range of frequencies can serve to suppress the processing of certain stimuli. In studies of the human brain, the range of approximately 8-14 Hz, has long been identified as the *alpha* range of frequencies (Klimesch, 1996; Lopes da Silva et al., 1973; Nunez & Srinivasan, 2006). In the fully-developed and neurotypical human adult, the peak of *alpha* oscillatory activity typically averages at a frequency near 10 Hz (Hanslmayr et al., 2011; Nunez et al., 2001). When populations of neurons are oscillating in complementary phase, they are said to be

“synchronized,” and the amplitude of the oscillatory signal produced by the electric field they generate increases. When amplitude of this signal decreases as the result of a reduction in the number of populations firing in phase at a particular frequency, this is described as “desynchronization” (Pfurtscheller & Lopes da Silva, 1999). These processes are explained because the *synchronization* of alpha oscillatory activity – or alpha *rhythms* – measured from posterior brain regions has been implicated as a mechanism of suppression (Klimesch et al., 2007).

In studies employing electroencephalography (EEG) as a measurement device, patterns of change in alpha frequency oscillations, or rhythms, have been of particular interest for understanding exogenous and endogenous impacts on how information is selected. For example, posterior alpha rhythms exhibit a large decrease in amplitude when the eyes are opened after having been closed – a phenomenon coined the “Berger Effect” (Başar, 2012; Bazanova et al., 2017). When the eyes are closed, alpha synchronization is observed; when opened, desynchronization. Because this change is not dependent upon the input of stimuli to the retina, however, and because alpha desynchronization is also observed with increases in cognitive load, the phenomenon is argued to relate to endogenous processes of attention and attentional control (Bazanova & Vernon, 2014; Klimesch et al., 2007). Further, in studies of non-human primates, alpha oscillations in the frontal eye field (FEF) have been found to provide a temporal structure, or rhythm, which anchors the periodic direction and re-direction of endogenous attention (Gaillard et al., 2020).

### 2.3.3 Strategies

In 1956, Miller wrote an engaging and influential article describing an interesting finding arising from “experiments on the capacity of people to transmit information.” The finding, in brief, was that people seemed capable of retaining and recalling over a brief period of time approximately 7 discrete pieces of information. Interestingly, this quantity of approximately 7 could be comprised of basic and fundamental “bits” of information, or they could be comprised of what he described as “chunks.” Miller described the conversion of multiple bits of information into a single chunk as a process of *recoding* (Miller, 1956), but it also became commonplace to refer to this type of information-quantity-reduction strategy simply as “chunking.” Without prompting or instruction, people spontaneously *chunk* information. A common classroom example of chunking is the act of simplifying a 7-digit phone number for ease of remembering. Instead of remembering 7 discrete single-digit numbers, remember two *chunks*: one three-digit number and one four-digit number (396-4900 instead of 3-9-6-4-9-0-0).

While the total number of chunks an individual can hold in short-term memory is rather small, a single chunk can potentially contain an arbitrarily large quantity of information (Halford et al., 1998; Miller, 1956; Simon, 1974). Chunking operations in which individuals incorporate representations they hold in long-term memory such as semantic, conceptual or even episodic memories, can significantly improve accuracy or even expand the amount of content that can be held within a single chunk (Hinault et al., 2017b; Roediger, 1980). An interesting study is discussed in Simon (1974) to illustrate these benefits in which novice and expert chess players were given lengthy chess piece

placement sequences to remember. While the novices could recall only a handful of piece placements, the expert players were able to accurately reproduce very long sequences. Examination of the timing of reproduction for each group revealed that the experts were making longer pauses in their reproductions after larger sequences of piece placements than were novices. This was taken as evidence the experts had the ability to recode a longer sequence of piece placements into a single chunk than could the novices, thereby allowing them to retain a greater amount of information short-term (Simon, 1974).

Many years after the publication of Miller's "Magical Number Seven," Cowan (2001) reviewed findings gleaned from four decades of research that followed. The evidence was definitive that despite the possibility of storing much content into a single chunk, there remain significant limitations in the amount of information one can actively hold in mind at a given time. Twenty years later, today, further accumulating evidence (Cowan, 2010; Vogel et al., 2001) has contributed to widespread scientific acceptance of working memory as a "limited capacity" system. It is noted that scientific agreement on this point is not universal, however. Other researchers have framed inaccuracies in remembering as stemming from limitations in precision rather than capacity, for example (Bays et al., 2009). While for Miller, that number defining capacity limits was seven, others settled on a lesser amount (*e.g.* five in Simon, 1974), and for Cowan that limit was approximately four (Cowan, 2001). If more is better, strategies to maximize the amount of information able to be held in a limited capacity system, then, are advantageous.

It has been suggested that recoding is such a common and significant process in working memory that models which don't account for recoding processes are incomplete

(Bor & Owen, 2007). This follows from the early and influential *multicomponent model of working memory* (Baddeley & Hitch, 1974), described earlier, which explicitly included *components* within which cognitive strategies employed to aid in information retention and recall are carried out. Telling in this model, is that entirely separate components were modeled dependent upon both the modality and type of information being processed. Visual and spatial information processing and recoding was modeled as occurring in the *visuo-spatial sketchpad*. Auditory and verbal information processing and recoding was modeled as occurring in the *Phonological loop*. This is because there is systematic difference in how information is processed, dependent upon modality and information type.

Different recoding, or other strategic information conversion approaches, can facilitate working memory dependent upon what one is trying to remember. Rehearsal, for example, is a frequently employed strategy for remembering without conversion of a representation (Mizrak & Oberauer, 2021). For information that can be verbalized, such as words or numbers perceived via vision or audition, one can simply repeat them over and over silently “in mind” to aid with recall (e.g. Hinault et al., 2017a). For navigating via landmarks to reach a destination, one might mentally visualize each important landmark in sequence, “rehearsing” a path of travel. Rehearsal can also be combined with other strategies, as on its own it may not be as effective as others (Roediger, 1980).

Strategies may be more or less effective not just based upon information modality, but upon the manner in which information is presented. When viewing spatial information that is presented sequentially, for example, people have been found hold

their gaze in a central, idealized location, as opposed to shifting their focus around. This has been argued to facilitate spatial recoding (Patt et al., 2014). Across a range of age, and regardless of formalized math instruction, people extract what are described as “geometrical primitives,” or the constituent components of simple shapes, from spatial sequences (Amalric et al., 2017). People have also been demonstrated to use these geometric primitives to create structure from complex spatial configurations (Al Roumi et al., 2021). Without necessarily forming shapes, memory for sequential spatial information has been shown to be facilitated when people group items based upon proximity, clustering nearest neighbors (Gmeindl et al., 2011). Differences in the timing of sequential presentation of spatially related objects has also been shown to impact how people chunk or recode (Bor et al., 2003). Further, passive strategies may be more effective than chunking when information is presented sequentially at a fast rate (Cowan et al., 2005).

## 2.4 The Storage and Processing of Representations

In 1967, Murdock used the term ‘sensory stores’ to specify a critical but distinct aspect of short-term memory which evidence showed to differ based upon perceptual modality. For example, short-term memory recall for lists of verbal information was repeatedly found to be greater when the information was initially perceived through audition than through vision. Murdock proposed a model in which a *sensory store* stage of memory was where sensory processing occurred and differed in a modality-specific manner. In this model short-term memory was comprised of a *primary memory*, where

errors of displacement (confusing or swapping to-be remembered items) occurred, and a *secondary memory* where interference between items might occur (Murdock Jr, 1967). Murdock's focus on different memory stores and the errors these stores are prone to, is at its root getting at the important question of *which* representations, even briefly, are stored. Representations biased for processing, and selectively attended to, are more likely to be *available* to be stored than those that are discarded, suppressed, or filtered out.

Several years after Murdock's attribution of memory errors to the characteristics of *primary* or *secondary* memory stores, Craik and Lockhart (1972) argued that the *level of processing* a perceived item was subjected to, provided a better account. Lesser processing might account for both the rapid forgetting of much sensory input and the frequent errors arising in short-term memory; greater processing the deep and vast associations and semantic meanings people were able to rapidly form after perception of information. In this conception the brain did not have regions of storage per se but rather regions of processing (Craik, 2002). They proposed that processing could be distinguished by degree or "level." An item processed solely by attention, for example, was being subjected to only the same, single, level of processing. Additional distributed, or deeper, subjection to processing such as association or rehearsal was described as processing at multiple levels. This subjection of information to multiple levels of processing was conceived as leading to greater memorability (Craik & Lockhart, 1972).

Evidence from neuroimaging studies of recent decades has provided evidence of deep, distributed, and continual processing of representations occurring throughout the brain in support of working memory (Christophel et al., 2017). Sensory representations

initially formed in sensory cortices may continue to be “held” in those same sensory regions when a task demands memory of specific sensory features (Scimeca et al., 2018). Control over those sensory representations may then be executed via a network including PFC and parietal regions specialized for processing of specific representation types or classes (Blacker & Courtney, 2016; Courtney et al., 1996; Ester et al., 2015; Roth & Courtney, 2007; Walsh et al., 2011; Yee et al., 2010). If task demands prompt strategic recoding or conversion, the network involved in representation transformation may place greater demands upon subregions of lateral PFC (Bor et al., 2003; Bor & Owen, 2007). Working memory related activity throughout lateral PFC in particular has been found to occur in a “hierarchical gradient” based upon the level of abstraction relevant for task demands (Nee & D’Esposito, 2016). More anterior regions integrating highly abstracted representations; posterior areas more specific and action directed (Nee & D’Esposito, 2016). Representations can be stored “silently” for short periods of time, via what have been posited as weighting changes in synaptic strength (M. G. Stokes, 2015). However, representations stored in this manner must be “reactivated” for manipulation and conversion — hallmark features of working memory (Trübutschek et al., 2019).

#### 2.4.1 Interference and Oscillatory Activity

While control regions contribute to processing a wide range of representations with fidelity, errors in discriminability can arise when the same substrate is utilized for both perceptual and memory based processing of items that are separate but similar in sensory features (Teng & Kravitz, 2019). Errors between perception and memory can be further exacerbated when separate but highly similar items are being encoded in a narrow

window of time (Fischer & Whitney, 2014). These kinds of errors - swap errors between similar items, or proactive interference arising from close proximity in time, are described as forms of internal distraction in working memory (Lorenc et al., 2021)

More pervasive, however, may be the threat of interference arising from external sources in the environment. Because the number of items that can be simultaneously attended to is limited (Cowan et al., 2005), mechanisms to prevent distraction and interference in working memory processes are critical. Alpha frequency oscillations, described earlier in the context of attentional selection, are understood to be part of such a mechanism. The *inhibition-timing* hypothesis, for example, posits that endogenous, inhibitory control (including distractor suppression) is achieved via alpha oscillatory activity (Klimesch et al., 2007). A specific and repeated observation, for example, is an increase in posterior alpha oscillatory power during maintenance of representations in working memory, ipsilateral to the visual field those maintained stimuli were initially perceived in. Further, the degree of alpha power change that occurs in support of working memory maintenance has also been shown to differ based upon the type of representation being held (Blacker et al., 2016; Ikkai et al., 2014). This increase in alpha power – or *synchronization* as explained earlier – has been interpreted as suppression of sensory cortices to reduce the potential for processing of new, irrelevant stimuli (Klimesch, 2012).

Expanding upon this (and lending support for the claim that *gating* processes are a prevalent mechanism occurring throughout the brain) is the *gating by inhibition* hypothesis of Jensen and Mazaheri (2010). They discuss alpha oscillations for their role in

context with oscillations occurring in other frequency ranges. Here, alpha oscillations are not considered in isolation, but rather as part of a partnership wherein they modulate the periods within which fast-spiking ( $> 60\text{Hz}$ ) gamma oscillations can occur. During the trough phases of an alpha frequency oscillatory cycle, gamma oscillations can occur in bursts. As the alpha cycle phase reaches its peak, gamma oscillations are inhibited. Gamma and alpha oscillations are juxtaposed, therefore, such that increases in gamma activity reflect active information processing in local networks, while increases in alpha activity reflect pulsed, functional inhibitions of said processing (Jensen & Mazaheri, 2010).

This “cross-frequency coupling” between alpha and gamma oscillations is not specific to their particular pairing. Rather, their pairing is representative of a general mechanism of organizing communication across brain regions (Fries, 2005). Cross-frequency coupling between theta and gamma oscillations in networks tied to the hippocampus, for example, also facilitate working memory processing (Lisman & Jensen, 2013). Coupling that facilitates memory exists not just within regions, but across regions of the brain as well. Popov *et al.*, for example, demonstrated an inverse power relationship between coupled theta oscillations in PFC and alpha oscillations in posterior parietal cortex that supported working memory performance of a challenging task (Popov *et al.*, 2018). Hinault *et al.*, identified specific networks by which alpha and gamma frequency coupling across anatomically distant regions supports working memory control processes. They demonstrated that poorly tuned coupling resulting from white matter degradation in these networks, slows down processing and impacts behavioral performance (Hinault *et al.*, 2020).

## 2.5 Open Questions

In a recent review on the subject of declarative memory formation, de Sousa, Chowdhury, and Silva present evidence that memories are formed, organized and structured based upon temporal, spatial, and conceptual relationships (de Sousa et al., 2021). Another study focused on episodic memory argues that we form a type of “knowledge structure” labeled as a “cognitive map” based upon experiences situated in time and space (Morton et al., 2017). Indeed, humans are organisms that move through and interact with space. We experience life as a continuum that temporally unfolds. And often, the meaning we extract from our environment is informed based upon relationships between the two. Interestingly, because our studies are often built around the isolation and manipulation of single variables within the laboratory setting, we have a great deal to learn about how these dimensions of time and space can impact working memory. The studies presented in the following chapters shed some light on this subject.

A number of researchers have focused their examination on differences that are specific to individuals, to understand how those differences may give rise to differences in behavior (e.g. Jarrold & Towse, 2006; Unsworth & Engle, 2007). Within paradigms that have been well studied at the group level, there remain interesting questions to pursue regarding how individuals may differ in either task performance or the underlying neural which supports their performance. In Chapter 3, I present a study in which a key difference between individuals — biological sex — is targeted as a factor of interest in spatial working memory.

Additional questions arise when we consider the impact of structures which intersect in space and time. Structure can exist not just within a single modality, but also across modalities. While many studies have focused upon examining working memory given a single input modality of information – such as audition or vision – there remain many open questions regarding possible interactions between the two. In Chapter 4, I pursue such a question, studying working memory performance in the presence of structures across space, time, and modality.

While some of the earliest studies of human capabilities utilized sequences of verbal information presented in continuous streams to test the limits and bounds of memory, relatively few paradigms exist for the study of nonverbal, visuospatial information presented as such. We do know that the temporal distance or proximity of information we are attending to can sometimes result in working memory errors (Fischer & Whitney, 2014; Kiyonaga et al., 2017; Makovski & Jiang, 2008; Ricker, 2015). However, these examples have typically demonstrated effects of interference relative to low-level features of objects such as object color or orientation. Further, depending upon the context, these same effects of interference could also be considered adaptive, even helpful (Kiyonaga et al., 2017). Unknown, is whether temporal factors may impact working memory performance beyond low-level features, perhaps interfering or aiding with recognition of entire objects themselves. In Chapter 5, I present a novel visuospatial paradigm designed specifically for this study. This paradigm is utilized to answer questions about working memory for visuospatial information presented with an understudied form of temporal structure — that of rhythm.

## 2.6 References

- Al Roumi, F., Marti, S., Wang, L., Amalric, M., & Dehaene, S. (2021). Mental compression of spatial sequences in human working memory using numerical and geometrical primitives. *Neuron*, 1–13. <https://doi.org/10.1101/2020.01.16.908665>
- Amalric, M., Wang, L., Pica, P., Figueira, S., Sigman, M., & Dehaene, S. (2017). The language of geometry: Fast comprehension of geometrical primitives and rules in human adults and preschoolers. In *PLoS Computational Biology* (Vol. 13, Issue 1). <https://doi.org/10.1371/journal.pcbi.1005273>
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In *Psychology of learning and motivation* (Vol. 2, pp. 89–195). Elsevier.
- Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences*, 16(8), 437–443. <https://doi.org/10.1016/j.tics.2012.06.010>
- Baddeley, A. (2010). Working memory. *Current Biology*, 20(4), 136–140. <https://doi.org/10.1016/j.cub.2009.12.014>
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47–89). Elsevier.
- Badre, D. (2012). Opening the gate to working memory. *Proceedings of the National Academy of Sciences of the United States of America*, 109(49), 19878–19879. <https://doi.org/10.1073/pnas.1216902109>
- Başar, E. (2012). A review of alpha activity in integrative brain function: Fundamental physiology, sensory coding, cognition and pathology. *International Journal of Psychophysiology*, 86(1), 1–24. <https://doi.org/10.1016/j.ijpsycho.2012.07.002>
- Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, 9(10), 1–11. <https://doi.org/10.1167/9.10.7>
- Bazanova, O. M., Nikolenko, E. D., & Barry, R. J. (2017). Reactivity of alpha rhythms to eyes opening (the Berger effect) during menstrual cycle phases. *International Journal of Psychophysiology*, 122, 56–64. <https://doi.org/10.1016/j.ijpsycho.2017.05.001>
- Bazanova, O. M., & Vernon, D. (2014). Interpreting EEG alpha activity. *Neuroscience and Biobehavioral Reviews*, 44(1), 94–110. <https://doi.org/10.1016/j.neubiorev.2013.05.007>
- Beck, D. M., & Kastner, S. (2014). Neural systems for spatial attention in the human brain. In A. C. Nobre & S. Kastner (Eds.), *The Oxford Handbook of Attention*. Oxford University Press.
- Bisley, J. W., & Goldberg, M. E. (2010). Attention, intention, and priority in the parietal lobe. *Annual Review of Neuroscience*, 33, 1–21. <https://doi.org/10.1146/annurev-neuro-060909-152823>
- Blacker, K. J., & Courtney, S. M. (2016). Distinct Neural Substrates for Maintaining Locations and Spatial Relations in Working Memory. *Frontiers in Human Neuroscience*, 10(November), 1–

17. <https://doi.org/10.3389/fnhum.2016.00594>
- Blacker, K. J., Ikkai, A., Lakshmanan, B. M., Ewen, J. B., & Courtney, S. M. (2016). The role of alpha oscillations in deriving and maintaining spatial relations in working memory. *Cognitive, Affective, & Behavioral Neuroscience*, 16(5), 888–901. <https://doi.org/10.3758/s13415-016-0439-y>
- Bor, D., Duncan, J., Wiseman, R. J., & Owen, A. M. (2003). Encoding strategies dissociate prefrontal activity from working memory demand. *Neuron*, 37(2), 361–367. [https://doi.org/10.1016/S0896-6273\(02\)01171-6](https://doi.org/10.1016/S0896-6273(02)01171-6)
- Bor, D., & Owen, A. M. (2007). A common prefrontal-parietal network for mnemonic and mathematical recoding strategies within working memory. *Cerebral Cortex*, 17(4), 778–786. <https://doi.org/10.1093/cercor/bhk035>
- Buschman, T. J., & Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science*, 315(5820), 1860–1864. <https://doi.org/10.1126/science.1138071>
- Christophel, T. B., Klink, P. C., Spitzer, B., Roelfsema, P. R., & Haynes, J. D. (2017). The Distributed Nature of Working Memory. *Trends in Cognitive Sciences*, 21(2), 111–124. <https://doi.org/10.1016/j.tics.2016.12.007>
- Cohen, J. D., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., Smith, E. E., & Perlstein, W. M. (1997). Temporal dynamics of brain activation during a working memory task. *Nature*, 386(April), 604–608.
- Courtney, S. M. (2004). Attention and cognitive control as emergent properties of information representation in working memory. *Cognitive, Affective and Behavioral Neuroscience*, 4(4), 501–516.
- Courtney, S. M., Petit, L., Haxby, J. V., & Ungerleider, L. G. (1998). The role of prefrontal cortex in working memory: Examining the contents of consciousness. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 353(1377), 1819–1828. <https://doi.org/10.1098/rstb.1998.0334>
- Courtney, S. M., Petit, L., Maisog, J. M., Ungerleider, L. G., & Haxby, J. V. (1998). An area specialized for spatial working memory in human frontal cortex. *Science*, 279(5355), 1347–1351. <https://doi.org/10.1126/science.279.5355.1347>
- Courtney, S. M., Ungerleider, L. G., Keil, K., & Haxby, J. V. (1996). Object and Spatial Visual Working Memory Activate Separate Neural Systems in Human Cortex. *Cerebral Cortex*, 6(1), 39–49. <https://doi.org/10.1093/cercor/6.1.39>
- Courtney, S. M., Ungerleider, L. G., Keil, K., & Haxby, J. V. (1997). Transient and sustained activity in a distributed neural system for human working memory. *Nature*, 386(6625), 608–611. <https://doi.org/10.1038/386608a0>
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24(1), 87–114. <https://doi.org/10.1017/S0140525X01003922>
- Cowan, Nelson. (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.  
<https://doi.org/10.1017/CBO9781139174909.006>
- Cowan, Nelson. (2008). Chapter 20 What are the differences between long-term, short-term, and working memory? In *Progress in Brain Research* (Vol. 169, Issue 07). Elsevier.  
[https://doi.org/10.1016/S0079-6123\(07\)00020-9](https://doi.org/10.1016/S0079-6123(07)00020-9)
- Cowan, Nelson. (2010). The Magical Mystery Four: How Is Working Memory Capacity Limited, and Why? *Current Directions in Psychological Science*, 19(1), 51–57.  
<https://doi.org/10.1177/0963721409359277>
- Cowan, Nelson, Elliott, E. M., Saults, S. J., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, 51(1), 42–100.  
<https://doi.org/10.1016/j.cogpsych.2004.12.001>
- Craik, F. I. M. (2002). Levels of processing: Past, present ... and future? *Memory*, 10(5–6), 305–318. <https://doi.org/10.1080/09658210244000135>
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11(6), 671–684.  
[https://doi.org/10.1016/S0022-5371\(72\)80001-X](https://doi.org/10.1016/S0022-5371(72)80001-X)
- de Sousa, A. F., Chowdhury, A., & Silva, A. J. (2021). Dimensions and mechanisms of memory organization. *Neuron*, 109, 1–14. <https://doi.org/10.1016/j.neuron.2021.06.014>
- Deacon, T. W. (1989). Holism and associationism in neuropsychology: An anatomical synthesis. In *Integrating Theory and Practice in Clinical Neuropsychology* (1st ed., pp. 1–47). Routledge. <https://doi.org/10.4324/9780429489464-1>
- Deacon, T. W. (1990). Fallacies of progression in theories of brain-size evolution. *International Journal of Primatology*, 11(3), 193–236. <https://doi.org/10.1007/BF02192869>
- Desimone, R., & Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. *Annual Review of Neuroscience*, 18, 193–222. <https://doi.org/10.1146/annurev-psych-122414-033400>
- Egeth, H. E., & Yantis, S. (1997). Visual attention: control, representation, and time course. *Annual Review of Psychology*, 48, 269–297.  
<https://doi.org/10.1146/annurev.psych.48.1.269>
- 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(4), 893–905. <https://doi.org/10.1016/j.neuron.2015.07.013>
- Fecteau, J. H., & Munoz, D. P. (2006). Salience, relevance, and firing: a priority map for target selection. *Trends in Cognitive Sciences*, 10(8), 382–390.  
<https://doi.org/10.1016/j.tics.2006.06.011>
- Fischer, J., & Whitney, D. (2014). Serial dependence in visual perception. *Nature Neuroscience*, 17(5), 738–743. <https://doi.org/10.1038/nn.3689>

- Fries, P. (2005). A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, 9(10), 474–480.  
<https://doi.org/10.1016/j.tics.2005.08.011>
- Gaillard, C., Ben Hadj Hassen, S., Di Bello, F., Bihan-Poudec, Y., VanRullen, R., & Ben Hamed, S. (2020). Prefrontal attentional saccades explore space rhythmically. *Nature Communications*, 11(1), 1–13. <https://doi.org/10.1038/s41467-020-14649-7>
- Gazzaley, A., Cooney, J. W., McEvoy, K., Knight, R. T., & D'Esposito, M. (2005). Top-down Enhancement and Suppression of the Magnitude and Speed of Neural Activity. In *Journal of Cognitive Neuroscience* (Vol. 17, Issue 3, pp. 507–517).  
<https://doi.org/10.1162/0898929053279522>
- Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: Bridging selective attention and working memory. *Trends in Cognitive Sciences*, 16(2), 129–135.  
<https://doi.org/10.1016/j.tics.2011.11.014>
- Gmeindl, L., Walsh, M., & Courtney, S. M. (2011). Binding serial order to representations in working memory: a spatial/verbal dissociation. *Memory & Cognition*, 39(1), 37–46.  
<https://doi.org/10.3758/s13421-010-0012-9>
- Goldman-Rakic, P. S. (1995). Cellular basis of working memory. *Neuron*, 14(3), 477–485.  
[https://doi.org/10.1016/0896-6273\(95\)90304-6](https://doi.org/10.1016/0896-6273(95)90304-6)
- Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences*, 21(06), 803–831.  
[http://journals.cambridge.org/abstract\\_S0140525X98001769](http://journals.cambridge.org/abstract_S0140525X98001769)
- Hanslmayr, S., Gross, J., Klimesch, W., & Shapiro, K. L. (2011). The role of alpha oscillations in temporal attention. *Brain Research Reviews*, 67(1–2), 331–343.  
<https://doi.org/10.1016/j.brainresrev.2011.04.002>
- Hillyard, S. A., Vogel, E. K., & Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: Electrophysiological and neuroimaging evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 353(1373), 1257–1270. <https://doi.org/10.1098/rstb.1998.0281>
- Hinault, T., Kraut, M., Bakker, A., Dagher, A., & Courtney, S. M. (2020). Disrupted Neural Synchrony Mediates the Relationship between White Matter Integrity and Cognitive Performance in Older Adults. *Cerebral Cortex*, 1–13.  
<https://doi.org/10.1093/cercor/bhaa141>
- Hinault, T., Lemaire, P., & Touron, D. (2017a). Aging effects in sequential modulations of poorer-strategy effects during execution of memory strategies. *Memory*, 25(2), 176–186.  
<https://doi.org/10.1080/09658211.2016.1146300>
- Hinault, T., Lemaire, P., & Touron, D. (2017b). Strategy combination during execution of memory strategies in young and older adults. *Memory*, 25(5), 619–625.  
<https://doi.org/10.1080/09658211.2016.1200626>
- Ikkai, A., Blacker, K. J., Lakshmanan, B. M., Ewen, J. B., & Courtney, S. M. (2014). Maintenance of relational information in working memory leads to suppression of the sensory cortex.

- Journal of Neurophysiology*, 112(8), 1903–1915. <https://doi.org/10.1152/jn.00134.2014>
- Ingle, D. J., Goodale, M. A., & Mansfield, R. J. W. (1982). *Analysis of visual behavior*. Mit Press Cambridge, MA.
- Jarrold, C., & Towse, J. N. (2006). Individual differences in working memory. *Neuroscience*, 139(1), 39–50. <https://doi.org/10.1016/j.neuroscience.2005.07.002>
- Jensen, O., & Mazaheri, A. (2010). Shaping Functional Architecture by Oscillatory Alpha Activity: Gating by Inhibition. *Frontiers in Human Neuroscience*, 4(November), 1–8. <https://doi.org/10.3389/fnhum.2010.00186>
- Jolicœur, P., & Dell'Acqua, R. (1998). The Demonstration of Short-Term Consolidation. *Cognitive Psychology*, 36, 138–202.
- Kiyonaga, A., Scimeca, J. M., Bliss, D. P., & Whitney, D. (2017). Serial Dependence across Perception, Attention, and Memory. *Trends in Cognitive Sciences*, 21(7), 493–497. <https://doi.org/10.1016/j.tics.2017.04.011>
- Klimesch, W. (1996). Memory processes, brain oscillations and EEG synchronization. *International Journal of Psychophysiology*, 24(1–2), 61–100. [https://doi.org/10.1016/S0167-8760\(96\)00057-8](https://doi.org/10.1016/S0167-8760(96)00057-8)
- Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*, 16(12), 606–617. <https://doi.org/10.1016/j.tics.2012.10.007>
- Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition-timing hypothesis. *Brain Research Reviews*, 53(1), 63–88. <https://doi.org/10.1016/j.brainresrev.2006.06.003>
- Kondo, H. M., Van Loon, A. M., Kawahara, J. I., & Moore, B. C. J. (2017). Auditory and visual scene analysis: An overview. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1714). <https://doi.org/10.1098/rstb.2016.0099>
- Lachter, J., Forster, K. I., & Ruthruff, E. (2004). Forty-five years after broadbent (1958): Still no identification without attention. *Psychological Review*, 111(4), 880–913. <https://doi.org/10.1037/0033-295X.111.4.880>
- Lisman, J. E., & Jensen, O. (2013). The Theta-Gamma Neural Code. *Neuron*, 77(6), 1002–1016. <https://doi.org/10.1016/j.neuron.2013.03.007>
- Lopes da Silva, F. H., van Lierop, T. H. M. T., Schrijer, C. F., & Storm van Leeuwen, W. (1973). Organization of thalamic and cortical alpha rhythms: Spectra and coherences. *Electroencephalography and Clinical Neurophysiology*, 35(6), 627–639. [https://doi.org/10.1016/0013-4694\(73\)90216-2](https://doi.org/10.1016/0013-4694(73)90216-2)
- Lorenc, E. S., Mallett, R., & Lewis-Peacock, J. A. (2021). Distraction in Visual Working Memory: Resistance is Not Futile. *Trends in Cognitive Sciences*, 25(3), 228–239. <https://doi.org/10.1016/j.tics.2020.12.004>
- Makovski, T., & Jiang, Y. V. (2008). Proactive interference from items previously stored in visual working memory. *Memory and Cognition*, 36(1), 43–52.

<https://doi.org/10.3758/MC.36.1.43>

- Marr, D. (1982). The Philosophy and the Approach. In *Vision : a computational investigation into the human representation and processing of visual information* (pp. 8–38). San Francisco: W.H. Freeman.
- Miller, E. K., & Cohen, J. D. (2001). An Integrative Theory of Prefrontal Cortex Function. *Annual Review of Neuroscience*, 24, 167–202.
- Miller, E. K., Lundqvist, M., & Bastos, A. M. (2018). Working Memory 2.0. *Neuron*, 463–475. <https://doi.org/10.1016/j.neuron.2018.09.023>
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *The Psychological Review*, 63(2), 81–97. <https://doi.org/10.1177/001088049003100202>
- Mishkin, M., & Ungerleider, L. G. (1982). Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. *Behavioural Brain Research*, 6(1), 57–77. [https://doi.org/10.1016/0166-4328\(82\)90081-X](https://doi.org/10.1016/0166-4328(82)90081-X)
- Mizrak, E., & Oberauer, K. (2021). What Is Time Good for in Working Memory? *Psychological Science*, 095679762199665. <https://doi.org/10.1177/0956797621996659>
- Morton, N. W., Sherrill, K. R., & Preston, A. R. (2017). Memory integration constructs maps of space, time, and concepts. *Current Opinion in Behavioral Sciences*, 17, 161–168. <https://doi.org/10.1016/j.cobeha.2017.08.007>
- Murdock Jr, B. B. (1967). Recent developments in short-term memory. *British Journal of Psychology*, 58(3–4), 421–433.
- Myers, N. E., Stokes, M. G., & Nobre, A. C. (2017). Prioritizing Information during Working Memory: Beyond Sustained Internal Attention. *Trends in Cognitive Sciences*, 21(6), 449–461. <https://doi.org/10.1016/j.tics.2017.03.010>
- Nee, D. E., & D’Esposito, M. (2016). The hierarchical organization of the lateral prefrontal cortex. *eLife*, 5(MARCH2016), 1–26. <https://doi.org/10.7554/eLife.12112>
- Nunez, P. L., & Srinivasan, R. (2006). *Electric fields of the brain: the neurophysics of EEG* (2nd ed.). Oxford University Press, USA.
- Nunez, P. L., Wingeier, B. M., & Silberstein, R. B. (2001). Spatial-temporal structures of human alpha rhythms: Theory, microcurrent sources, multiscale measurements, and global binding of local networks. *Human Brain Mapping*, 13(3), 125–164. <https://doi.org/10.1002/hbm.1030>
- O’Reilly, R. C., & Frank, M. J. (2006). Making working memory work: A computational model of learning in the prefrontal cortex and basal ganglia. In *Neural Computation* (Vol. 18, Issue 2). <https://doi.org/10.1162/089976606775093909>
- Oberauer, K. (2002). Access to Information in Working Memory: Exploring the Focus of Attention. *Journal of Experimental Psychology: Learning Memory and Cognition*, 28(3), 411–421. <https://doi.org/10.1037/0278-7393.28.3.411>
- Oberauer, K. (2019). Working Memory and Attention – A Conceptual Analysis and Review.

- Journal of Cognition*, 2(1), 1–23. <https://doi.org/10.5334/joc.58>
- Panichello, M. F., & Buschman, T. J. (2021). Shared mechanisms underlie the control of working memory and attention. *Nature*, 592(7855), 601–605. <https://doi.org/10.1038/s41586-021-03390-w>
- Patt, V. M., Thomas, M. L., Minassian, A., Geyer, M. A., Brown, G. G., & Perry, W. (2014). Disentangling working memory processes during spatial span assessment: A modeling analysis of preferred eye movement strategies. *Journal of Clinical and Experimental Neuropsychology*, 36(2), 186–204. <https://doi.org/10.1080/13803395.2013.877123>
- Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG / MEG synchronization and desynchronization : basic principles. *Clinical Neurophysiology*, 110, 1842–1857. [https://doi.org/10.1016/S1388-2457\(99\)00141-8](https://doi.org/10.1016/S1388-2457(99)00141-8)
- Popov, T., Popova, P., Harkotte, M., Awiszus, B., Rockstroh, B., & Miller, G. A. (2018). Cross-frequency interactions between frontal theta and posterior alpha control mechanisms foster working memory. *NeuroImage*, 181(July), 728–733. <https://doi.org/10.1016/j.neuroimage.2018.07.067>
- Pressnitzer, D. (2011). Auditory scene analysis: the sweet music of ambiguity. *Frontiers in Human Neuroscience*, 5(December), 1–11. <https://doi.org/10.3389/fnhum.2011.00158>
- Ricker, T. J. (2015). The role of short-term consolidation in memory persistence. *AIMS Neuroscience*, 2(4), 259–279. <https://doi.org/10.3934/Neuroscience.2015.4.259>
- Roediger, H. L. (1980). The effectiveness of four mnemonics in ordering recall. *Journal of Experimental Psychology: Human Learning and Memory*, 6(5), 558–567. <https://doi.org/10.1037/0278-7393.6.5.558>
- Rohenkohl, G., Coull, J. T., & Nobre, A. C. (2011). Behavioural dissociation between exogenous and endogenous temporal orienting of attention. *PLoS ONE*, 6(1), 1–5. <https://doi.org/10.1371/journal.pone.0014620>
- Roth, J. K., & Courtney, S. M. (2007). Neural system for updating object working memory from different sources: Sensory stimuli or long-term memory. *NeuroImage*, 38(3), 617–630. <https://doi.org/10.1016/j.neuroimage.2007.06.037>
- Scimeca, J. M., Kiyonaga, A., & D’Esposito, M. (2018). Reaffirming the Sensory Recruitment Account of Working Memory. *Trends in Cognitive Sciences*, 22(3), 190–192. <https://doi.org/10.1016/j.tics.2017.12.007>
- Shinn-Cunningham, B. G., Ihlefeld, A., Satyavarta, S., & Larson, E. (2005). Bottom-up and Top-down Influences on Spatial Unmasking. *Acta Acustica United with Acustica*, 91, 967–979.
- Simon, H. A. (1974). How Big Is a Chunk? *Science*, 183(4124), 482–488.
- Stokes, M., & Duncan, J. (2014). Dynamic brain states for preparatory attention and working memory. In A. C. Nobre & S. Kastner (Eds.), *The Oxford Handbook of Attention* (pp. 152–180). Oxford University Press.
- Stokes, M. G. (2015). “Activity-silent” working memory in prefrontal cortex: A dynamic coding framework. *Trends in Cognitive Sciences*, 19(7), 394–405.

<https://doi.org/10.1016/j.tics.2015.05.004>

- Teng, C., & Kravitz, D. J. (2019). Visual working memory directly alters perception. *Nature Human Behaviour*, 3(8), 827–836. <https://doi.org/10.1038/s41562-019-0640-4>
- Theeuwes, J. (2014). Spatial orienting and attentional capture. In A. C. Nobre & S. Kastner (Eds.), *The Oxford Handbook of Attention* (pp. 231–252). Oxford University Press Oxford.
- Theeuwes, J. (2018). Visual selection: Usually fast and automatic; Seldom slow and volitional. *Journal of Cognition*, 1(1), 1–15. <https://doi.org/10.5334/joc.13>
- Treisman, A. M. (1964). Selective attention in man. *British Medical Bulletin*, 20(1), 12–16. <https://doi.org/10.1093/oxfordjournals.bmb.a070274>
- Trübutschek, D., Marti, S., Ueberschär, H., & Dehaene, S. (2019). Probing the limits of activity-silent non-conscious working memory. *Proceedings of the National Academy of Sciences of the United States of America*, 116(28), 14358–14367. <https://doi.org/10.1073/pnas.1820730116>
- Ungerleider, L. G., Courtney, S. M., & Haxby, J. V. (1998). A neural system for human visual working memory. *Proceedings of the National Academy of Sciences of the United States of America*, 95(3), 883–890. <https://doi.org/10.1073/pnas.95.3.883>
- Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, 114(1), 104–132. <https://doi.org/10.1037/0033-295X.114.1.104>
- van Ede, F., Niklaus, M., & Nobre, A. C. (2017). Temporal expectations guide dynamic prioritization in visual working memory through attenuated  $\alpha$  oscillations. *Journal of Neuroscience*, 37(2), 437–445. <https://doi.org/10.1523/JNEUROSCI.2272-16.2016>
- van Ede, F., & Nobre, A. C. (2021). Toward a neurobiology of internal selective attention. *Trends in Neurosciences*, xx(xx), 1–3. <https://doi.org/10.1016/j.tins.2021.04.010>
- Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of Features, Conjunctions, and Objects in Visual Working Memory. *Journal of Experimental Psychology: Human Perception and Performance*, 27(1), 92–114. <https://doi.org/10.1037/0096-1523.27.1.92>
- Walsh, M., Montojo, C. A., Sheu, Y.-S., Marchette, S. A., Harrison, D. M., Newsome, S. D., Zhou, F., Shelton, A. L., & Courtney, S. M. (2011). Object working memory performance depends on microstructure of the frontal-occipital fasciculus. *Brain Connectivity*, 1(4), 317–329. <https://doi.org/10.1089/brain.2011.0037>
- Wolfe, J. M. (2021). Guided Search 6.0 : An updated model of visual search. *Psychonomic Bulletin and Review*.
- Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 10(5), 601–621. <https://doi.org/10.1037/0096-1523.10.5.601>
- Yantis, S., Schwarzbach, J., Serences, J. T., Carlson, R. L., Steinmetz, M. A., Pekar, J. J., & Courtney, S. M. (2002). Transient neural activity in human parietal cortex during spatial

attention shifts. *Nature Neuroscience*, 5(10), 995–1002. <https://doi.org/10.1038/nn921>

Yee, L. T. S., Roe, K., & Courtney, S. M. (2010). Selective involvement of superior frontal cortex during working memory for shapes. *Journal of Neurophysiology*, 103(1), 557–563. <https://doi.org/10.1152/jn.91299.2008>

## Chapter 3

# Modulation of Peak Alpha Frequency Oscillations during Working Memory is greater in Females than Males

## 3.1 Introduction

Oscillatory activity within the alpha range of frequencies, canonically defined as 8-14Hz, has been studied for its role in facilitating working memory (Klimesch, 1999; Roux & Uhlhaas, 2014). One critical characteristic of alpha activity – its peak frequency – has been tied to differences in working memory performance and is well known to differ across individuals based upon genetics, age, and the type of cognitive activity in which an individual is engaged (Knyazeva et al., 2018; Mierau et al., 2017; Posthuma et al., 2001). Despite these known differences, individuals are often treated as members of a single homogenous study group; their separate data points represented solely by group-level statistics. Here, we show that by accounting for an individual's binary sex (female or male) differences in task-related modulation of peak alpha frequency, and its relationship to working memory performance become apparent.

### 3.1.1 Alpha Oscillations and Working Memory

Oscillatory activity is thought to be one of the mechanisms by which functional organization and structured communication in the brain are achieved (Basar & Düzgün, 2016; Buzsáki et al., 2013; Voytek & Knight, 2015). In working memory, the power of alpha oscillations has been related to performance (Klimesch, 1999), working memory maintenance (Herrmann et al., 2004), and working memory capacity (Tuladhar et al., 2007). Working memory by definition reflects rapid and short-lived processes that are subject to interference (Courtney et al., 2007). An ascribed functional role of alpha oscillations arising during working memory maintenance, is to facilitate preservation of maintained information. This facilitation is achieved through minimization of potential interference by dampening incoming external sensory stimuli when previously encoded information is being maintained (de Vries et al., 2020; Ikkai et al., 2014; Jensen & Mazaheri, 2010). The role of these oscillations may not simply be reactive, but proactive, as their phase and power have been found to shift in advance of anticipated distractors during a working memory task (Bonnefond & Jensen, 2012). Increases in posterior alpha power in particular, are hypothesized to reflect neural processes that protect information being held in working memory from interference by current sensory inputs (Klimesch 2012; Ikkai et al. 2014; Roux and Uhlhaas 2014). Causal manipulations using transcranial alternating current stimulation and transcranial magnetic stimulation at alpha frequencies supports this hypothesis (Borghini et al., 2018; Riddle et al., 2020).

### 3.1.2 Peak Alpha Frequency

The alpha oscillatory activity of a neurotypical young to middle-age adult, measured when they are awake in a resting state, will on average exhibit a peak of power near the frequency of 10 Hz (Clark et al., 2004; Haegens et al., 2014; Posthuma et al., 2001). This apex of the power frequency distribution, known as the Peak Alpha Frequency (PAF), is known to be somewhat variable across individuals, however (Doppelmayr et al., 1998). Developmental studies have found that PAF increases throughout childhood and adolescence and only stabilizes at a frequency average near 10 Hz in late adolescence to early adulthood (Chiang et al., 2011; Smith, 1941). With aging, PAF will typically drift lower once again, with the average PAF of those near 70 years old being closer to 8 Hz (Dustman et al., 1999; Scally et al., 2018). The typical developmental trajectory of increasing working memory performance and capacity, followed by subsequent advanced-age declines (e.g. Hamilton, Coates, and Heffernan 2003; Hinault et al. 2020; Reuter-Lorenz and Park 2010; Park et al. 2002), occur in seeming parallel to these described increases and decreases in PAF.

PAF can also shift on a timescale of moments, rather than that of a lifespan. During engagement in a cognitive task, PAF may shift to higher or lower frequencies, and do so differently across different brain regions. For example, Haegens et al. (2014) found that average posterior PAF increased when participants were performing an n-back working memory task compared to when they were at rest. They also found PAF increased with greater task difficulty. Sauseng et al. (2005) contrasted occipital and prefrontal PAF during sequential split-halves of a working memory task maintenance period. During working

memory maintenance, PAF measured over posterior regions was consistently higher than PAF measured pre-frontally. When the task demanded information also be manipulated, however, for the first half of the analyzed time period prefrontal and posterior PAF were aligned. Their findings support different functional roles of alpha oscillations dependent upon the region from which they arise and the networks with which they are associated. Faster posterior alpha oscillations seemingly support working memory maintenance, while prefrontal supports working memory control.

### 3.1.3 Influences of Sex on PAF

Sex-linked differences in PAF have been reported starting in childhood and continuing on into advanced age. A developmental study of children, for example, found that males reached a higher and more stable average PAF than females by the age of 11 (Matthis et al., 1980). A more recent study found this difference continued up until the age of 16 (Chiang et al., 2011). A large study that utilized a clinical database with EEG recordings from patients ranging in age from infancy to the late 80s, found that from 20 years of age and continuing up to 85, females had higher alpha frequencies than males (Aurlen et al., 2004).

Studies from the 1970s onwards have demonstrated predictable sex-linked variance in PAF. Wuttke et al. (1975), for example, measured resting state alpha frequencies in two groups of women: one naturally cycling, and one taking hormone-based oral contraceptives. In the naturally cycling group, alpha frequencies were found to increase and decrease with the cyclical fluctuation of sex-hormones that occur across

menstrual phases. In the group taking oral contraceptives, which prevent natural sex-hormone level fluctuations, alpha frequencies remained relatively constant over time (Wuttke et al., 1975). Becker et al. (1982) also examined the connection between sex-hormones and resting-state alpha changes across the menstrual cycle in females, but also included cognitive task performance as a variable of interest. Their psychological battery included tests of short-term and working memory. Replicating the findings of Wuttke et al. (1975), they found that mean alpha frequency increased and decreased cyclically in conjunction with fluctuating hormone levels, yet they found no difference in cognitive performance tied to these fluctuations (Becker et al., 1982).

#### 3.1.4 PAF and Cognitive Performance

The lack of relationship between sex-linked PAF fluctuations and performance described above might be surprising considering the positive associations between PAF and behavior previously reported. Higher resting state anterior PAF has been linked to higher Reverse Digit Span performance, for example, independent of age in a large gender-balanced study (Clark et al., 2004). In another study comprised of 12 males and 4 females, the researchers employed a split-half analysis of working memory performers and found that the average PAF of the good performer group was 1.25Hz greater than that of poor performer group (Klimesch et al., 1993). Additional evidence for positive correlations between PAF and speed of information processing, and PAF and memory performance have also been reported (Bazanov & Vernon, 2014; Klimesch et al., 1990). Moreover, shifting PAF higher or lower using sensory entrainment has been shown to causally influence cognitive performance (Ronconi et al., 2018).

In other cases, however, the relationship between PAF and performance is less clear. Angelakis et al. (2004), for example, reported no PAF-performance relationships in a small gender-balanced experiment. In a follow-up experiment where PAF and task performance were measured on two separate days in 13 females and 6 males, they found correlations of posterior PAF with performance on one day, but not the other (Angelakis et al., 2004). A possible contributor to the discrepant findings between all of these studies considered may be the neglect of a key factor – biological sex. Indeed, there is a gap in the literature regarding potential sex differences in PAF, particularly potential differences while under cognitive load.

### 3.1.5 Study Question

We hypothesized that during working memory, females and males may differentially modulate oscillatory neural activity to support task performance. As discussed above, females of reproductive age exhibit predictable sex-specific variation in PAF, yet this variation does not necessarily impact their task performance. In some studies that included both male and female participants, relationships between PAF and performance have been found. It is possible, therefore, that cognitive and neural processes related to alpha oscillatory activity underlying working memory, are utilized differently in females versus males, particularly as they contribute to task performance. These differences may be tied to sex differences in PAF modulation, or differences in the relationship between individual variability in PAF and working memory performance.

## 3.2 Methods

To test our hypothesis, we re-analyzed data from a study that employed a task in which both female and male participants maintained different types of spatial information in working memory – either precise locations of individual stimuli or spatial locations of the stimuli relative to each other– while undergoing EEG (Blacker et al., 2017). Brain activity during working memory for these different types of information has been found to differ, both neuroanatomically as measured with fMRI (Blacker & Courtney, 2016) and electrophysiologically as measured with EEG (Blacker et al., 2016). We hypothesized, therefore, that there might be sex differences in the neural activity underlying working memory for both precise and relative locations, only precise, only relative, or neither.

Analysis of neural activity recorded by EEG during key working memory periods – particularly maintenance and baseline– enabled us to compare PAF across both time periods and task conditions. It also enabled examination of a possible differential relationship between PAF and behavioral performance for females versus males. A recent meta-analysis reported a small female advantage in visuospatial working memory (Grissom & Reyes, 2019) whereas previous meta-analyses have found a small male advantage in certain types of spatial tasks (Hyde, 2014). The paradigms employed in those studies, however, differ from the paradigm used here, and behavioral similarities between females and males are the more typical finding (Hyde, 2005). Hence, we predicted there would be no difference between females and males in behavioral

performance. We did, however, predict differences between these groups in PAF, given the prior evidence outlined above. Further, we predicted that within-group differences in PAF or task-related modulation of PAF might be tied to behavioral performance, but that this activity-behavior relationship might be different in females versus males.

### 3.2.1 Participants

Study participants were young adults recruited from both the college student population and the local community. Participants ranged in age from 18 to 31 years. Binary grouping of participant sex as female or male was based upon self-report.

Any participants with below chance behavioral performance in the primary task, or incomplete EEG data, were excluded. After these criteria were applied, data from a total of 110 participants (33 males and 77 females) remained and were analyzed. Groups were not statistically different in age ( $p > 0.3$ ; mean (SD): males 21.8(3.4), females 21.1(3.2) years).

### 3.2.2 Working Memory Task

A visuospatial working memory task (**Figure 1**) was performed while continuous EEG was recorded. Conditions (Task-types) were distinguished based upon the number of sample stimuli presented and what aspects of those stimuli were relevant. Instructions differed for each of the four Task-types but are described here first based upon their similarities to aid comprehension. In Precise location trials the specific locations of the sample stimuli were relevant, whereas in Relative trials, the locations of the sample stimuli relative to each other were relevant. For both Task-types, either two or three

colored circle stimuli were presented for 500ms during the sample period. Instructions for each of the four Task-types are as follows: In 2-Sample Precise trials, participants were instructed to imagine a single line connecting the two sample stimuli, and then, after a 2000ms delay, determine if the test stimulus, a single black circle, was located upon that imagined line. In 3-Sample Precise trials, participants were to determine if the test stimulus occupied the same location as any one of the sample stimuli. In 2-Sample Relative trials, participants were to maintain the vertical relationship between the sample stimuli and determine if the test stimuli, also two, colored circles, held the same relationship. In 3-Sample Relative trials, the vertical spatial relationship between any two of the three sample stimuli might be relevant at test. All stimuli were displayed in the same single quadrant of the screen during a trial. Participants were instructed to maintain fixation on a central cross for the duration of each trial.

Participants were able to indicate a match or non-match response by keypress as soon as the test stimuli were displayed, and for up to an additional 1000ms after. Feedback for each trial was presented via a 100ms color change in the on-screen fixation cross and was immediately followed by the start of the next trial.

Prior to the EEG recording session, participants completed 88 practice trials, 22 in each condition. During EEG recording, participants completed 8 blocks of 64 trials for a total of 512 trials. Task-type trial order was pseudorandomized within each block. For further, detailed task methods, see *Spatial Locations and Relations Task* (Blacker et al., 2017; Hinault et al., 2019).

### 3.2.3 EEG Collection, and Pre-Processing

EEG Recording was performed at a sampling rate of 512Hz with a 138Hz cutoff filter. A 47-electrode equidistant montage was utilized and re-referenced to the average of all channels offline (for further details of EEG acquisition, see Blacker et al. 2017). The Fieldtrip software package (Oostenveld et al., 2011) was utilized for raw EEG data cleaning, processing, and analysis. Per the Fieldtrip pipeline, raw data were segmented, cleaned, and sorted by trial condition prior to spectral decomposition for analysis of induced power in the time-frequency domain. Detailed steps of this procedure follow.

Prior to segmentation, the continuous raw EEG data was high-pass filtered at 0.5Hz, de-meanned, and low-pass filtered at 30Hz. Then, data from each of the 47 channels were divided into temporally overlapping 7.5 second segments that included the 5.5 second trial period and the 2 seconds preceding trial onset. This is described as overlapping due to the final 2 seconds of one segment being the same as the starting 2 seconds of the subsequent segment. This standard approach of buffering via extended segmentation (Cohen, 2014) ensured preservation of signal during the Baseline time period (ref **Figure 1**). Independent Components Analysis with 40 primary components (Oostenveld et al., 2011) was utilized to identify, and target for removal, artifacts resulting from blinks, eye movements or excessive facial muscle activity (e.g. jaw clenching). Following this algorithm-based cleaning, visual inspection of the EOG channels placed above the left and right eyes was performed to identify ocular artifacts. Trials containing voltage shifts greater than  $18.75\mu\text{V}$  in the EOG channels were removed from the data.

Finally, any remaining trials with large (e.g., an order of magnitude greater) voltage spikes not filtered out by independent component analysis were removed.

Clean, segmented data were sorted based upon trial condition. Only correct trials were selected for analysis. In preparation for subsequent frequency-based analyses, spectral decomposition was performed. Cleaned, sorted, correct trial data underwent spectral analysis at a frequency resolution of 0.5Hz from 2.0Hz to 20.0Hz via complex wavelet convolution with a 2.0 second sliding window, zero-padded with a Hann/hanning multi-taper time frequency transformation based upon multiplication in the frequency domain. The Hann taper is well-suited for broad-band low frequency estimation and does not introduce edge artifacts (Cohen, 2014). This approach allowed for measurement of power at precise and specific frequency increments, with sufficient frequency resolution for direct comparative analyses between the defined time periods of interest. Spectral analysis resulted in a 3-dimensional data matrix of power organized by time, frequency, and channel. This matrix was utilized for the analyses described below.

### 3.2.4 Peak Alpha Frequency

The *Center of Gravity* method was used for calculating Peak Alpha Frequency (PAF), (Goljahani et al., 2012; Klimesch, 1999). In this method upper and lower frequency bounds and a time period are used as constraints within which to establish the specific frequency at which the power of alpha oscillations is maximal. This method was selected for its potential to increase the signal-to-noise ratio when calculating PAF during a cognitive task (Klimesch, 1999; Klimesch et al., 1993). The canonical alpha range of 8-14Hz

was used to define the frequency boundaries within which PAF was calculated and three time periods of interest were selected.

As discussed in the introduction, alpha power arising from posterior brain regions is understood to facilitate working memory maintenance. Posterior electrode channels are a common target for measurement of PAF (e.g., Haegens et al. 2014). Additionally, this study incorporated a visual working memory paradigm in which there have been prior findings related to alpha oscillatory activity over posterior regions (Blacker et al., 2016). For the purposes of this study, therefore, of the 47 channels recorded, 14 posterior channels were selected *a priori* for PAF calculation and used in entirety for calculation of PAF during each time period of interest. Channels and time periods are shown in **Figure 1**. *A priori* channel selection allowed for direct comparison of participants based upon female or male grouping (see recommendations in Rippon et al. 2014) whereas a data driven approach might suffer confounds resulting from unforeseen sex-linked group differences in electrode distribution or gross brain anatomy.

Channels and time periods for analysis were selected with the aim of minimizing inclusion of (if not excluding) signal not primarily related to alpha and working memory processes of particular interest in this study. Mu band activity, for example, overlaps with the alpha frequency range but originates in relation to motor response and is typically measured over motor and pre-motor regions (e.g. Hadley 1941). Mastoid channels and the central topmost channel, which due to their position may be detecting activity related to motor response rather than working memory maintenance, were thus not included in the selection of posterior channels. Additionally, the time periods of interest (ref. **Figure**

1) did not include the Test or Response periods during which stimuli differed across conditions and participants were preparing for, or making, a button press response.

Three separate measures of PAF within the range of 8.0Hz to 14.0Hz at a frequency resolution of 0.5Hz for each participant were made: during Baseline, Maintenance, and across the Pre-Test Task period (which included the Baseline and Maintenance periods as shown in **Figure 1**). Two discrete time periods of interest during the task, labeled Baseline and Maintenance were targeted for comparative PAF analyses. Baseline was defined as the last 500ms of fixation, prior to the presentation of the cue and sample stimuli. The length of this time window allowed for measurement of approximately four oscillatory cycles at the lowest frequency of interest here (8Hz) – a number of cycles within the typically recommended range (Cohen, 2014). This Baseline period covers a pre-stimulus range of time similar to the baseline periods of other studies (e.g. Hinault et al. 2019; Blacker et al. 2016). Maintenance was defined as the last 1500ms of the delay, prior to the presentation of test stimuli. As the name of the time period suggests, this period allows for analysis of PAF as it relates to the maintaining of information in working memory. The naming and definitions of these time periods of interest is in line with previous working memory research that utilizes EEG (e.g., Hinault et al. 2019; Bae and Luck 2018; Blacker et al. 2016; Manza, Hau, and Leung 2014; and see Roux and Uhlhaas 2014).

As stated earlier, spectral analysis results in a 3-dimensional matrix. Within that matrix, power is organized by the channel measured, the frequency at which oscillation is occurring, and the time of measurement. Following the *Center of Gravity* method

(Klimesch, 1999), to attenuate noise from any single one of the 14 posterior channels selected, a weighted averaging of power across channels was performed prior to PAF calculation. To compute weights for the weighted averaging process, power across the 14 channels of interest was first summed at a single time point. The proportion of a single channel's contribution to that sum was deemed its weight at that time point. At each time point, and for each 0.5Hz frequency increment within the range of 8Hz to 14Hz, the vector of power for each channel and the vector of weights for each channel were multiplied together. The average of this product was then taken. These steps were performed for all frequencies and time periods of interest. The frequency at which the largest weighted average of power existed within the time period of interest was denoted the Peak Alpha Frequency.

### 3.2.5 Statistical Analysis

Dependent variables relevant to our predictions included Accuracy (proportion of correct responses), Response Time, PAF, and PAF Modulation. Both parametric and non-parametric analytical approaches were employed with these measures defined as outcome variables. To test our predictions regarding task performance, PAF, and PAF Modulation, group level distributions were statistically compared using a robust and non-parametric approach: the Kolmogorov-Smirnov test (Rousselet et al., 2017). Modulation of PAF between the Baseline and Maintenance time periods was calculated as [Maintenance PAF – Baseline PAF] in units of Hz.

In addition to null hypothesis testing for a difference in distributions, shift functions (Doksum, 1974; Doksum & Sievers, 1976) were computed. This method quantifies how two distributions differ by comparing corresponding quantiles of each distribution and delivering a function showing how one of the distributions would need to change, or “shift,” to match the other. All shifts between quantiles are statistically tested as a group with correction for multiple comparisons, providing a robust measure of where and how distributions differ (Doksum & Sievers, 1976; Rousselet et al., 2017). All shift functions calculated herein computed differences between the Male and Female distributions as a function of the Female distributions. Further description and illustration of this method are provided in detail elsewhere (Rousselet et al., 2017).

Parametric group level comparisons of Accuracy, Response Time and PAF Modulation by Task Type were performed with separate 2 (Sex) x 4 (Task-type) ANOVA tests. This 2x4 analysis was more appropriate than a 2x2x2 ANOVA design because task instructions differed for each Task-type dependent upon load and thus load was not an independent factor. In addition to this statistical rationale, our previous findings pointed towards a qualitative difference in the neural effects of load for Precise versus Relative conditions in this paradigm (Blacker & Courtney, 2016). For completeness, however, we also analyzed the data using a 2x2x2 ANOVA. The results of the two analyses were essentially identical. In the results below, we present the 2x4 ANOVA results. In cases where violations of sphericity were present, correction was applied and reported. Tukey’s Honestly Significant Difference (HSD) method at an alpha level of 0.05 was used in post-

hoc analyses. For models of task performance and PAF modulation, main effects of Sex, Task-type, and their interaction were fit.

To test our prediction of differential relationships between behavior and PAF, and behavior and PAF Modulation, between females and males, hierarchical multiple regression and correlational analyses were performed. Accuracy (proportion correct) and Response Time were the behavioral measures of interest and the outcome variables for all equation-based analyses. Baseline, Maintenance, and Pre-Test Task period PAF—encompassing both the Baseline and Maintenance time periods—was utilized in modeling. Separate two-step hierarchical regressions were performed for each group due to the prediction of different relationships by Sex in which PAF impacts behavior. In the first step, Task Performance (Accuracy and Response Time were each modeled separately) was predicted by Task-type. In the second step, PAF (Pre-Test Task PAF, Baseline PAF, and Maintenance PAF were modeled separately) was added into the model as an interacting factor. An ANOVA test was then used to determine if the first and second models differed significantly. In post-hoc correlational analyses, PAF from each of the three time periods of interest was used. The False Discovery Rate (FDR) method was used for multiple comparisons corrections of p-values.

Statistics were calculated in R (R Core Team, 2018) using the *dplyr* (Wickham et al., 2018), the *car* (Fox et al., 2012), the *ez* (Lawrence & Lawrence, 2016), and the *lme4* (Bates et al., 2015) packages. Distribution analyses were performed using Robust Graphical Methods For Group Comparisons (the *rogme* package), per the methods

described in (Rousselet et al., 2017). Figures were produced in R with the *ggplot2* (Wickham, 2016) and the *rogme* packages.

### 3.3 Results

Behavioral results are presented first. They are followed by results of analyses of PAF calculated from correct trials of all conditions (Precise and Relative, 2- and 3-Sample stimuli), and then results of analyses with PAF calculated separately for each Task-type.

#### 3.3.1 Task Performance

Behavioral measures of central tendency are reported in **Table 1**. Two separate 2x4 (Sex x Task-type) ANOVAs were performed with Accuracy and Response Time as outcome variables. Main effects of Task-type were present for Accuracy:  $F(3,324) = 356.35$ ,  $p < 0.0001$ ,  $\eta^2 = 0.546$ , and Response Time:  $F(3,324) = 182.68$ ,  $p < 0.0001$ ,  $\eta^2 = 0.217$ . Accuracy was lower for the higher load (3-Sample) conditions. Response time was greater for the Relative task conditions as has been previously observed (Ikkai et al., 2014). No main effects of Sex, nor interaction between Condition and Sex were present. Post-hoc analysis of Accuracy indicated significantly different groupings between 3-Sample Precise and 3-Sample Relative Task-types, and each in comparison to the 2-Sample Task-types together. No differences were present by Sex. Post-hoc analysis of Response Time indicated significantly different groupings between 2-Sample Precise compared to 3-Sample Relative and 3-Sample Precise Task-types, with no differences by Sex.

Group distributions and shift functions of Accuracy and Response Time across all correct trials are shown in **Figure 2** and **Figure 3**, respectively. Two-sample, two-sided Kolmogorov-Smirnov tests indicated no difference between female and male Accuracy distributions:  $D=0.09$ ,  $p > 0.5$ ; or Response Time distributions:  $D=0.17$ ,  $p > 0.4$ .

### 3.3.2 Peak Alpha Frequency within each Time Period

PAF was extracted for each participant during the time periods of interest. Measures of central tendency for PAF by Task-type and Sex are presented in **Table 2**. Separate 2x4 (Sex x Task-type) ANOVA tests were performed with PAF from the different time periods as outcome variables. Baseline PAF test results indicated no main effect of Sex:  $F(1,108) = 0.99$ ,  $p > 0.3$ , Task-type:  $F(3,324)=0.34$ ,  $p > 0.7$ , nor an interaction between them:  $F(3,324) = 0.99$ ,  $p > 0.3$ . With PAF during Maintenance as the outcome variable, test results indicated no main effect of Sex:  $F(1,108) = 0.24$ ,  $p > 0.6$ , no significant main effect of Task-type:  $F(3,324) = 1.82$ ,  $p > 0.1$ , nor an interaction between them:  $F(3,324)=1.06$ ,  $p > 0.3$ .

Two-sample, two-sided Kolmogorov-Smirnov tests were performed to test for differences between female and male distributions both averaged across, and within specific task conditions, during the Baseline and Maintenance time periods. Test results indicated no significant differences between group distributions at Baseline nor during Maintenance,  $p > 0.1$  in all instances.

### 3.3.3 Modulation of Peak Alpha Frequency

PAF Modulation measures of central tendency by Task-type are shown in **Table 2**. As predicted, the results of a 2x4 ANOVA (Sex x Task-type) with Modulation as the outcome variable shows a main effect of Sex:  $F(1,108) = 5.31$ ,  $p < 0.025$ ,  $\eta^2 = 0.03$ . No main effect of Task-type:  $F(3,324) = 1.56$ ,  $p > 0.1$ , or interaction between Sex and Task-type:  $F(3,324) = 0.63$ ,  $p > 0.5$ , was present. Females exhibited greater Modulation of PAF than males, across Task-types.

Considering that males Modulated PAF less than females, it would be expected that PAF at Baseline would be more highly correlated with PAF during Maintenance among males compared to females. This was the case. For males, Pearson's  $r = 0.82$ , for females:  $r = 0.58$ . The difference between male and female correlations is significant:  $z = 4.71$ ,  $p < 0.0001$ . Distributions of Modulation of PAF (the difference, calculated within each individual, between PAF during the Maintenance versus Baseline time periods) are shown in **Figure 4**. A two-sample, two-sided Kolmogorov-Smirnov test indicates a statistically significant difference between the female and male distributions of Modulation:  $D = 0.29$ ,  $p < 0.05$ .

### 3.3.4 Peak Alpha Frequency and Behavior

To test our prediction of a relationship between PAF and task performance, regression analyses using behavioral measures (Accuracy and Response Time) as outcome measures predicted by PAF during the time periods of interest were performed. Separate

correlational analyses were then performed to further examine the relationship between PAF and Accuracy for each Task-type.

Females exhibited greater Modulation of PAF than Males (above results) yet did not have an apparent behavioral advantage in performance across task conditions. It is possible that different relationships between PAF, Modulation of PAF, and task-performance exist for males versus females. Female and male groups were modeled separately, based upon the prediction of different PAF-behavior relationships between them.

In all second-step hierarchical models with Response Time as the outcome variable, the addition of PAF as a predictor did not improve model fit above and beyond Task-type alone. This was the case for both males and females, and PAF from each time periods of interest. For the female group, the addition of PAF to Task-type as a predictor of Accuracy did not improve model fit over Task-type as a sole predictor. This was the case for PAF during all three time periods of interest (Baseline, Maintenance, and the Pre-Test Task period). For the male group, however, model fit improved significantly when PAF was incorporated as a predictor of Accuracy. In males, Accuracy predicted by the interaction of Task-type and Baseline PAF was a better fit than Accuracy predicted by Task-type alone:  $F(2,124)=3.30$ ,  $p<0.05$ . The second-step model incorporating Baseline PAF was a significant predictor of Accuracy:  $F(7,124)=36.6$ ,  $p<0.0001$ ,  $R^2=0.66$ . The second-step model incorporating Maintenance PAF as an interacting predictor in males, exhibited a trending difference over Task Type as the sole predictor:  $F(2,124)=2.22$ ,  $p<0.08$ . This second-step model was also a significant predictor of Accuracy in males:

$F(7,124)=34.9$ ,  $p<0.0001$ ,  $R^2=0.64$ . As expected, then, given that the Pre-Test Task period incorporates both the Baseline and Maintenance period, the second-step model adding Pre-Test Task period PAF also proved a significantly better fit than the first-step model:  $F(2,124)=4.07$ ,  $p<0.01$ . This second-step model was a significant predictor of Accuracy in males:  $F(7,124)=37.8$ ,  $p<0.0001$ ,  $R^2=0.66$ .

Further examining this relationship between Accuracy and PAF in males, separate correlational analyses were performed. Correlations by group, separately for each task condition and corrected for multiple comparisons, between Accuracy and PAF during the three time periods of interest are shown in **Table 3**. In Precise Conditions, there was no significant correlation between Accuracy and PAF for either females or males. A significant positive correlation between PAF and Accuracy among males was present in the 3-Sample Relative Condition ( $r=0.49$ ,  $p<0.05$ ) and trending in the same direction in the 2-Sample Relative Condition ( $r=0.42$ ,  $p=0.11$ ) during the Pre-Test Task time period of interest. There was no correlation between these measures in females ( $r=-0.01$ ,  $p>0.05$ ). A significant difference was present between male and female group correlations in the 3-Sample Relative Pre-Test Task time period ( $z=2.51$ ,  $p=0.01$ ).

### 3.4 Discussion

The present study presents evidence for sex-differences in the modulation of PAF during visuospatial working memory. Across four different Task-types, females modulated PAF more than males. In males, PAF during task performance could be

described as more constant. In males, overall PAF across Baseline and working memory Maintenance periods was more highly correlated with task performance than it was in females.

While PAF has been described as a “stable neurophysiological trait” in adults (Grandy et al., 2013), this phrase is not an accurate descriptor as it applies to adult females. In females, resting-state PAF varies predictably in relation to the menstrual cycle, as discussed in the Introduction. Additionally, the difference between resting-state and task-related PAF has also been found to vary across individuals and within individual females across the menstrual cycle (Bazanov et al. 2014). In the present study, we find that females, as a group, also modulate PAF between time periods during performance of a task.

What the cognitive function of this modulation may be, however, is not clear. Modulation did not render a behavioral advantage for females. Between females and males, there was no statistical difference in task performance – either for Accuracy or Response Time – and group performance distributions were quite similar. A recent meta-review that focused on sex differences in visuo-spatial working memory found small effects of a female advantage for specific location memory and a male advantage for other visuo-spatial tasks (Voyer et al., 2017). The present task, however, which contrasts working memory for precise and relative spatial locations, did not find a female or male advantage for either task. Another recent review (Grissom & Reyes, 2019) examined the evidence for sex and gender differences in executive function in both animal models and human behavior. They highlight the importance of considering motivation, reward, and

strategy usage when evaluating whether small effects should reasonably be attributed to differences in sex or gender. In the present study, there are main effects of task condition in both accuracy and response time (**Table 1**), with lower accuracy and slower response times, generally, in 3-Sample compared to 2-Sample conditions. No speed-accuracy trade-off was present across conditions, nor were there different patterns of behavior by sex group that might indicate differences in motivation or response to reward.

In females, there was no significant relationship between either PAF Modulation and task performance, or overall PAF and task performance. For males, though, indication of a relationship between PAF and task performance was present. Specifically, in males, Accuracy within Relative, but not Precise, Task-types was positively correlated with PAF.

As reviewed in the Introduction, a study with predominantly male participants found that the better performing half of participants had higher PAFs than the worse performing half (Klimesch et al., 1993). In a study comprised solely of males, a positive correlation was found between PAF and an aspect of “non-verbal intellect” as assessed by a standardized test (Bazanov and Aftanas 2008). It would be prudent for future research, given the sex-specific findings here and the findings of prior studies, to account for sex as a factor when examining PAF and potential relationships with cognitive processes or task performance.

In the paradigm utilized here, Task-types varied—in working memory load demand and in the type of information required to be maintained. During both the Baseline and Maintenance time periods, only a fixation cross was visually present on

screen. The amount of information a Participant needed to hold in mind, however, increased considerably in the Maintenance period compared with Baseline. Across Task-types, rules needed to be maintained, but with each trial, new stimuli-specific information needed to be maintained. There was no effect of Task-type on either PAF during Maintenance, or Modulation of PAF. However, the significant positive correlation in males between PAF, particularly during Baseline and performance on the higher-load Relative Task-type raises interesting questions.

The Baseline time period preceded the presentation of stimuli pertaining to the current trial and followed the inter-trial interval of the preceding trial. It was not a period of true “rest,” but rather a brief pause in which the cognitive demands of one trial have passed and the demands of a new trial are pending. How alert and attentive an individual participant was during this time period prior to the more demanding high-load Relative trials may have had an impact on trial performance. Some studies dividing alpha activity into upper and lower bands have associated a reduction in the lower range of alpha power with attentional processes – specifically alertness and expectancy (Klimesch et al. 1998; and see reviews: Klimesch 1999, Bazanova and Vernon 2014). That PAF falls at a higher frequency for some individuals might be explained by reduced lower alpha power due to increased alertness and expectancy.

PAF provides only a single metric of the power distribution of alpha oscillatory activity and does not allow for conclusions to be drawn regarding the origin of variations in those distributions. We do not know, for example, whether PAF Modulation in females reflects a change in the rate of oscillations occurring within a single circuit or network, or

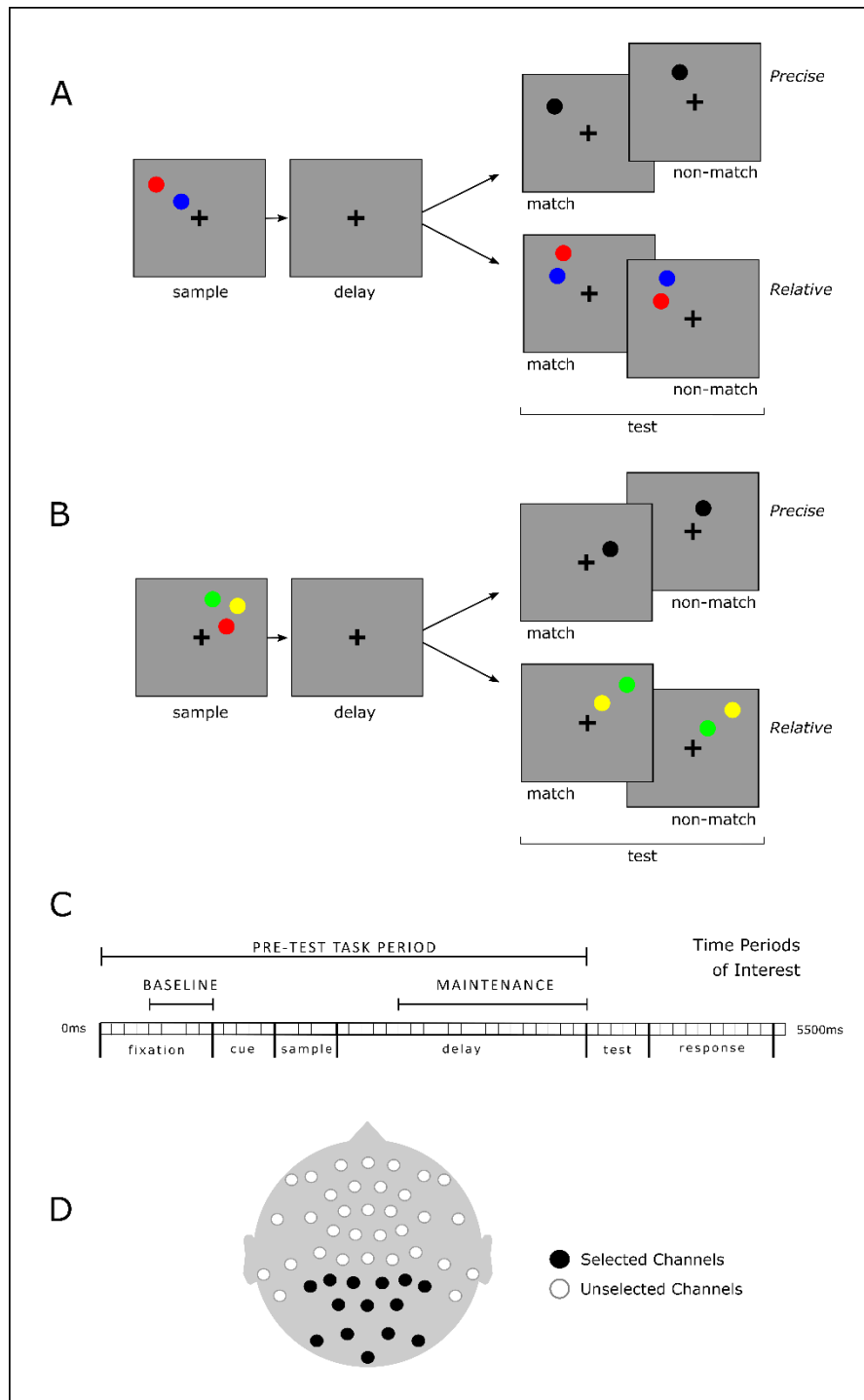
whether it reflects a decrease in the activity of one circuit or network and an increase in the activity of another. A recent study utilizing MEG, examined changes in the frequency of localized oscillatory activity as they contributed to working memory performance (Noguchi & Kakigi, 2020). They found better working memory performers exhibited more “flexible modulation” of localized frequency based on memory load demands. Of the 24 participants they studied, 18 were female. The results of the current study indicate the peak frequency of alpha oscillatory activity across time periods changes more in females than in males, but the relationship between peak frequency and performance was only observed in males. Perhaps females were able to more rapidly modulate attention (and thus PAF) across the different time periods of each trial, whereas males had a more constant attention state across trials. It is important for future research to consider sex in order to understand whether the alpha frequency distribution or its modulation is differentially related to cognitive task performance according to sex.

Throughout we have used the term “sex” as a potential factor of influence on neural activity. In instances where there is indication that sex-hormones are an associated, and perhaps even causal factor in differences between females and males, the term sex-difference may be appropriate. Gender as a term, however, encompasses many of the developmental events and life experiences that arise, in part, from the juxtaposition of biological sex and identity. Gender undoubtedly impacts experience, thinking, strategy, and behavior – and potentially the neural activity underlying them. In the present study, we did not find differences in task performance between females and males that might point towards differences in strategy, motivation, or other factors in

which gender has been shown to impact behavior. A limitation of the present study, however, is the smaller sample size of male participants compared to female. Echoing the recommendations espoused by experts in sex and gender research (e.g. Rippon et al. 2014; McCarthy and Arnold 2008; Rubin, Atwood, and Olson 2020; Jordan-Young and Rumiati 2012), future studies that attempt to disentangle effects of sex from those of gender would greatly increase our broader understanding of individual differences. Studies able to incorporate larger participant groups, balanced in number, would be ideal.

We hypothesized that females and males have different cognitive or neural processes related to oscillatory brain activity by which behavioral outcomes are accomplished. The results of our analyses of peak alpha frequency, measured while participants performed multiple variants of a spatial working memory task, provides support for this hypothesis. Researchers in cognitive neuroscience are more often coming to recognize the importance of studying oscillatory activity based upon frequency bands defined for each individual based on PAF. It is critically important that researchers also be aware of the implications of the potential for sex differences in task-related modulations within those frequency bands. Only by accounting for factors of sex, when appropriate, can we hope to move towards a deeper understanding of the brain basis of behavior.

### 3.5 Figures

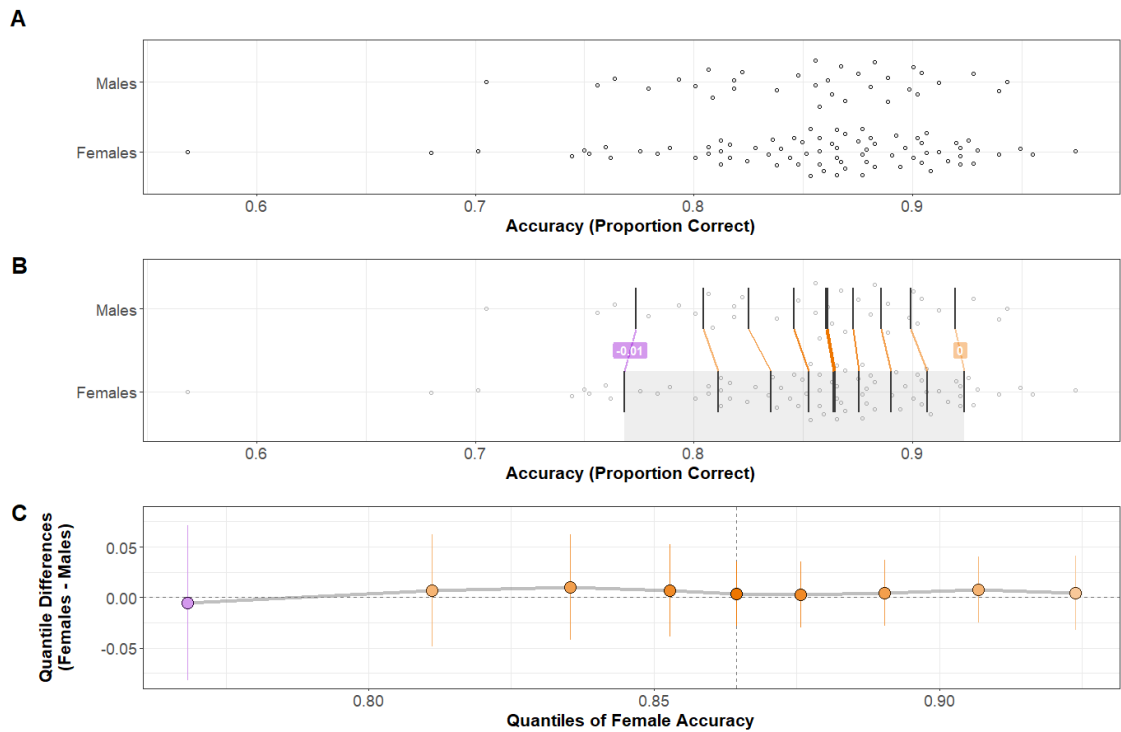


**Figure 1:** Working Memory Task and Timeline, selected time periods and channels for analysis of EEG. The four distinct Task-types included two or three sample stimuli within Precise and Relative conditions. All Task-types required a match/non-match decision based upon the specific instructions for the Task-type. Visualization is organized for clarity but note that sample load is not an independent factor from

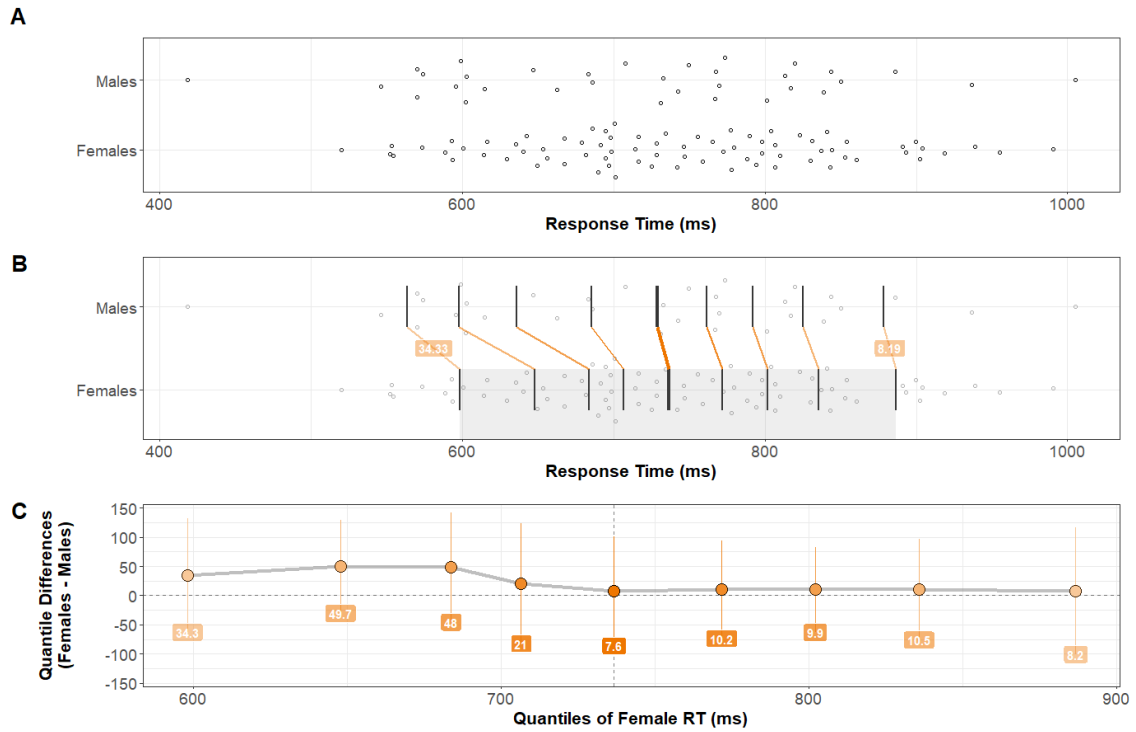
condition; instructions differed for each of the four Task-types. **A:** 2-Sample Task-type examples. In the Precise spatial Task-type, the test stimulus (black) matched if positioned upon an imaginary line segment (held in working memory) connecting the centers of the two sample stimuli. In the Relative spatial Task-type, the vertical spatial relationship of the colored test stimuli was relevant, but not their specific positions. In the example shown, the red sample stimulus is positioned higher than the blue sample stimulus. At test, a red stimulus above a blue stimulus would be a match, regardless of their particular vertical or horizontal positions. **B:** 3-Sample Task-type examples. With three sample stimuli, in the Precise Task-type, the test stimulus was a match if it occupied the exact spatial position of any one of the sample stimuli. In the Relative Task-type, test stimuli could be any two of the three sample stimuli colors, but were always in a different spatial location (though within the same visual quadrant); only the vertical spatial relationship between the corresponding color stimuli was relevant. **C:** Timeline for each Task-type and time periods of interest for EEG analysis. Baseline: 500ms prior to cue and visual sample stimuli presentation (the cue indicated whether a precise or relative judgment would be required). Maintenance: 1500ms during the working memory delay prior to presentation of test stimuli. **D:** EEG channel montage. All recorded channels are depicted. The 14 channels indicated in black were selected a priori for PAF analyses. All 14 selected channels were included in the analysis for each time period of interest.

**Table 1:** Measures of central tendency—Medians, Means, and Standard Deviations (SD)—in Task Performance by Group: females (F) and males (M), shown for each Task-type. Main effects of Task-type are present for Accuracy ( $p<0.05$ ) and Response Time ( $p<0.05$ ). No main effect of Sex or interaction between Task-type and Sex are present.

		Task-type											
		2-Sample Precise			2-Sample Relative			3-Sample Precise			3-Sample Relative		
Sex		Median	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD
Accuracy (proportion correct)	F	0.93	0.91	0.08	0.95	0.92	0.07	0.70	0.72	0.09	0.88	0.86	0.08
	M	0.95	0.92	0.06	0.96	0.93	0.06	0.69	0.70	0.08	0.88	0.86	0.09
Response Time (ms)	F	663.7	665.7	122.9	713.5	715.5	115.8	747.2	750.6	120.5	851.3	834.8	110.4
	M	668.6	641.6	132.6	673.7	681.4	141.7	738.1	722.2	135.6	835.7	830.7	132.1



**Figure 2:** Accuracy Distributions by group and comparison by Shift Function. **A:** Stripchart of Accuracy distributions by group. **B:** Stripchart shown in A with quantiles (vertical black bars) and differences between distribution quantiles characterized. When the difference between quantiles (Females–Males) is negative, the connecting line between corresponding quantiles is purple; when positive, orange. The heavier weight vertical black bar denotes the median (5<sup>th</sup> quantile) of each distribution. **C:** Shift function between Male and Female distributions. Range of the x-axis corresponds to the shaded region of the Female distribution in B. Y-axis shows the difference between group distributions by quantile: as in B, purple indicates a negative difference; orange indicates a positive difference. Points indicate how much the Male distribution would need to shift at a particular quantile to match the corresponding quantile in the Female distribution. Vertical lines at each point represent a bootstrapped 95% confidence interval about the difference. Note that difference points all near zero and confidence intervals at every quantile difference point cross over zero: there is no significant difference in Accuracy across Task-types between Females and Males. Two-sample, two-sided Kolmogorov-Smirnov test:  $D=0.09$ ,  $p > 0.5$ .



**Figure 3:** Response Time distributions by group and comparison by Shift Function. **A:** Stripchart of Response Time distributions (mean Response Time across Trial-types) for Males and Females. **B:** Stripchart shown in A with quantiles and differences between distribution quantiles characterized (see **Figure 2** caption for detailed explanation). **C:** Shift function between Male and Female distributions. Note that bootstrapped 95% confidence intervals at each quantile difference point cross zero. There is no significant difference in Response Time across Task-types between Females and Males. Two-sample, two-sided Kolmogorov-Smirnov test:  $D=0.17$ ,  $p > 0.4$ .



**Table 3:** Correlations between Accuracy and Peak Alpha Frequency (PAF) during the Pre-Test Task Period, Baseline and Maintenance time periods (as shown in **Figure 1**). Accuracy means by Trial-type are shown in **Table 1**. Pearson's *r* correlation statistics are shown along with *p*-values indicating significance of difference from zero. All *p* column values of “–” indicate a multiple-comparison-corrected  $p > 0.25$ . Values shown in parenthesis indicate the value of *p* without correction for multiple comparisons. Neither Males nor Females show a pattern of PAF correlating with Accuracy in Precise condition trials. For Males, there is indication of a positive correlation between PAF and Accuracy in Relative trials. A significant difference between Male and Female correlations—indicated in bold—is present for the 3-Sample Relative condition during the Pre-Test Task Period ( $z=2.56$ ,  $p=0.01$ ).

Task-type	Pre-Test Task Period PAF				Baseline PAF				Maintenance PAF			
	Females		Males		Females		Males		Females		Males	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
2-Sample Precise	0.05	–	0.02	–	0.08	–	-0.00	–	0.00	–	0.08	–
2-Sample Relative	0.13	–	0.40	0.14 (0.02)	0.09	–	0.37	0.14 (0.03)	0.13	–	0.30	0.23 (0.09)
3-Sample Precise	-0.09	–	0.08	–	0.10	–	0.05	–	-0.04	–	-0.00	–
3-Sample Relative	<b>-0.02</b>	–	<b>0.49</b>	0.03 (0.00)	0.03	–	0.44	0.04 (0.01)	-0.05	–	0.37	0.09 (0.04)

## 3.6 References

- Angelakis, E., Lubar, J. F., Stathopoulou, S., & Kounios, J. (2004). Peak alpha frequency: An electroencephalographic measure of cognitive preparedness. *Clinical Neurophysiology*, 115(4), 887–897. <https://doi.org/10.1016/j.clinph.2003.11.034>
- Aurlien, H., Gjerde, I. O., Aarseth, J. H., Eldøen, G., Karlsen, B., Skeidsvoll, H., & Gilhus, N. E. (2004). EEG background activity described by a large computerized database. *Clinical Neurophysiology*, 115(3), 665–673. <https://doi.org/10.1016/j.clinph.2003.10.019>
- Bae, G. Y., & Luck, S. J. (2018). Dissociable decoding of spatial attention and working memory from EEG oscillations and sustained potentials. *Journal of Neuroscience*, 38(2), 409–422. <https://doi.org/10.1523/JNEUROSCI.2860-17.2017>
- Basar, E., & Düzgün, A. (2016). The CLAIR model: Extension of Brodmann areas based on brain oscillations and connectivity. *International Journal of Psychophysiology*, 103, 185–198. <https://doi.org/10.1016/j.ijpsycho.2015.02.018>
- Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Grothendieck, G., Green, P., & Bolker, M. Ben. (2015). Package ‘lme4.’ *Convergence*, 12(1), 2.
- Bazanova, O. M., & Aftanas, L. I. (2008). Individual measures of electroencephalogram alpha activity and non-verbal creativity. *Neuroscience and Behavioral Physiology*, 38(3), 227–235. <https://doi.org/10.1007/s11055-008-0034-y>
- Bazanova, Olga M., Kondratenko, A. V., Kuzminova, O. I., Muravlyova, K. B., & Petrova, S. E. (2014). EEG alpha indices depending on the menstrual cycle phase and salivary progesterone level. *Human Physiology*, 40(2), 140–148. <https://doi.org/10.1134/S0362119714020030>
- Bazanova, Olga M., & Vernon, D. (2014). Interpreting EEG alpha activity. *Neuroscience and Biobehavioral Reviews*, 44(1), 94–110. <https://doi.org/10.1016/j.neubiorev.2013.05.007>
- Becker, D., Creutzfeldt, O. D., Schwibbe, M., & Wuttke, W. (1982). Changes in physiological, eeg and psychological parameters in women during the spontaneous menstrual cycle and following oral contraceptives. *Psychoneuroendocrinology*, 7(1), 75–90. [https://doi.org/10.1016/0306-4530\(82\)90057-9](https://doi.org/10.1016/0306-4530(82)90057-9)
- Blacker, K. J., & Courtney, S. M. (2016). Distinct Neural Substrates for Maintaining Locations and Spatial Relations in Working Memory. *Frontiers in Human Neuroscience*, 10(November), 1–17. <https://doi.org/10.3389/fnhum.2016.00594>
- Blacker, K. J., Ikkai, A., Lakshmanan, B. M., Ewen, J. B., & Courtney, S. M. (2016). The role of alpha oscillations in deriving and maintaining spatial relations in working memory. *Cognitive, Affective, & Behavioral Neuroscience*, 16(5), 888–901. <https://doi.org/10.3758/s13415-016-0439-y>

- Blacker, K. J., Negoita, S., Ewen, J. B., & Courtney, S. M. (2017). N-back versus Complex Span Working Memory Training Predictors of Near Transfer. *Journal of Cognitive Enhancement*, 52(2010), 2012.
- Bonnefond, M., & Jensen, O. (2012). Alpha oscillations serve to protect working memory maintenance against anticipated distracters. *Current Biology*, 22(20), 1969–1974. <https://doi.org/10.1016/j.cub.2012.08.029>
- Borghini, G., Candini, M., Filannino, C., Hussain, M., Walsh, V., Romei, V., Zokaei, N., & Cappelletti, M. (2018). Alpha oscillations are causally linked to inhibitory abilities in ageing. *Journal of Neuroscience*, 38(18), 4418–4429. <https://doi.org/10.1523/JNEUROSCI.1285-17.2018>
- Buzsáki, G., Logothetis, N., & Singer, W. (2013). Scaling brain size, keeping timing: Evolutionary preservation of brain rhythms. *Neuron*, 80(3), 751–764. <https://doi.org/10.1016/j.neuron.2013.10.002>
- Chiang, A. K. I., Rennie, C. J., Robinson, P. A., van Albada, S. J., & Kerr, C. C. (2011). Age trends and sex differences of alpha rhythms including split alpha peaks. *Clinical Neurophysiology*, 122(8), 1505–1517. <https://doi.org/10.1016/j.clinph.2011.01.040>
- Clark, C. R., Veltmeyer, M. D., Hamilton, R. J., Simms, E., Paul, R., Hermens, D., & Gordon, E. (2004). Spontaneous alpha peak frequency predicts working memory performance across the age span. *International Journal of Psychophysiology*, 53(1), 1–9. <https://doi.org/10.1016/j.ijpsycho.2003.12.011>
- Cohen, X. M. (2014). Analyzing Neural Time Series Data. In *MIT Press* (Issue 1). <https://doi.org/10.1007/s13398-014-0173-7.2>
- Courtney, S. M., Roth, J. K., & Sala, J. B. (2007). A hierarchical biased-competition model of domain-dependent working memory. In N. Osaka, R. H. Logie, & M. D’Esposito (Eds.), *The cognitive neuroscience of working memory* (pp. 369–384). Oxford University Press, USA.
- de Vries, I. E. J., Slagter, H. A., & Olivers, C. N. L. (2020). Oscillatory Control over Representational States in Working Memory. *Trends in Cognitive Sciences*, 24(2), 150–162. <https://doi.org/10.1016/j.tics.2019.11.006>
- Doksum, K. (1974). Empirical Probability Plots and Statistical Inference for Nonlinear Models in the Two- Sample Case. *The Annals of Statistics*, 2(2), 267–277.
- Doksum, K. A., & Sievers, G. L. (1976). Plotting with confidence: Graphical comparisons of two populations. *Biometrika*, 63(3), 421–434. <https://doi.org/10.1093/biomet/63.3.421>
- Doppelmayr, M., Klimesch, W., Pachinger, T., & Ripper, B. (1998). Individual differences in brain dynamics: important implications for the calculation of event-related band power. *Biological Cybernetics*, 79(1), 49–57. <https://doi.org/10.1007/s004220050457>
- Dustman, R. E., Shearer, D. E., & Emmerson, R. Y. (1999). Life-span changes in EEG spectral amplitude, amplitude variability and mean frequency. *Clinical Neurophysiology*, 110(8), 1399–1409. [https://doi.org/10.1016/S1388-2457\(99\)00102-9](https://doi.org/10.1016/S1388-2457(99)00102-9)

- Fox, J., Weisberg, S., Adler, D., Bates, D., Baud-Bovy, G., Ellison, S., Firth, D., Friendly, M., Gorjanc, G., Graves, S., & others. (2012). Package 'car.' *Vienna: R Foundation for Statistical Computing*.
- Goljahani, A., D'Avanzo, C., Schiff, S., Amodio, P., Bisiacchi, P., & Sparacino, G. (2012). A novel method for the determination of the EEG individual alpha frequency. *NeuroImage*, 60(1), 774–786. <https://doi.org/10.1016/j.neuroimage.2011.12.001>
- Grandy, T. H., Werkle-Bergner, M., Chicherio, C., Schmiedek, F., Lövdén, M., & Lindenberger, U. (2013). Peak individual alpha frequency qualifies as a stable neurophysiological trait marker in healthy younger and older adults. *Psychophysiology*, 50(6), 570–582. <https://doi.org/10.1111/psyp.12043>
- Grissom, N. M., & Reyes, T. M. (2019). Let's call the whole thing off: evaluating gender and sex differences in executive function. *Neuropsychopharmacology*, 44(1), 86–96. <https://doi.org/10.1038/s41386-018-0179-5>
- Hadley, J. M. (1941). Some relationships between electrical signs of central and peripheral activity: II. During "mental work." *Journal of Experimental Psychology*, 28(1), 53–62. <https://doi.org/10.1037/h0062239>
- Haegens, S., Cousijn, H., Wallis, G., Harrison, P. J., & Nobre, A. C. (2014). Inter- and intra-individual variability in alpha peak frequency. *NeuroImage*, 92, 46–55. <https://doi.org/10.1016/j.neuroimage.2014.01.049>
- Hamilton, C. J., Coates, R. O., & Heffernan, T. (2003). What develops in visuo-spatial working memory development? *European Journal of Cognitive Psychology*, 15(1), 43–69. <https://doi.org/10.1080/09541440303597>
- Herrmann, C. S., Senkowski, D., & Röttger, S. (2004). Phase-locking and amplitude modulations of EEG alpha: Two measures reflect different cognitive processes in a working memory task. *Experimental Psychology*, 51(4), 311–318. <https://doi.org/10.1027/1618-3169.51.4.311>
- Hinault, T., Blacker, K. J., Gormley, M., Anderson, B. A., & Courtney, S. M. (2019). Value-driven attentional capture is modulated by the contents of working memory: An EEG study. *Cognitive, Affective and Behavioral Neuroscience*, 19(2), 253–267. <https://doi.org/10.3758/s13415-018-00663-2>
- Hinault, T., Kraut, M., Bakker, A., Dagher, A., & Courtney, S. M. (2020). Disrupted Neural Synchrony Mediates the Relationship between White Matter Integrity and Cognitive Performance in Older Adults. *Cerebral Cortex*, 1–13. <https://doi.org/10.1093/cercor/bhaa141>
- Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, 60(6), 581–592. <https://doi.org/10.1037/0003-066X.60.6.581>
- Hyde, J. S. (2014). Gender similarities and differences. *Annual Review of Psychology*, 65, 373–398. <https://doi.org/10.1146/annurev-psych-010213-115057>

- Ikkai, A., Blacker, K. J., Lakshmanan, B. M., Ewen, J. B., & Courtney, S. M. (2014). Maintenance of relational information in working memory leads to suppression of the sensory cortex. *Journal of Neurophysiology*, 112(8), 1903–1915. <https://doi.org/10.1152/jn.00134.2014>
- Jensen, O., & Mazaheri, A. (2010). Shaping Functional Architecture by Oscillatory Alpha Activity: Gating by Inhibition. *Frontiers in Human Neuroscience*, 4(November), 1–8. <https://doi.org/10.3389/fnhum.2010.00186>
- Jordan-Young, R., & Rumiati, R. I. (2012). Hardwired for sexism? Approaches to sex/gender in neuroscience. *Neuroethics*, 5(3), 305–315. <https://doi.org/10.1007/s12152-011-9134-4>
- Klimesch, W., Doppelmayr, M., Russegger, H., Pachinger, T., & Schwaiger, J. (1998). Induced alpha band power changes in the human EEG and attention. *Neuroscience Letters*, 244(2), 73–76. [https://doi.org/10.1016/S0304-3940\(98\)00122-0](https://doi.org/10.1016/S0304-3940(98)00122-0)
- Klimesch, Wolfgang. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews*, 29(2–3), 169–195. [https://doi.org/10.1016/S0165-0173\(98\)00056-3](https://doi.org/10.1016/S0165-0173(98)00056-3)
- Klimesch, Wolfgang. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*, 16(12), 606–617. <https://doi.org/10.1016/j.tics.2012.10.007>
- Klimesch, Wolfgang, Schimke, H., Ladurner, G., & Pfurtscheller, G. (1990). Alpha frequency and memory performance. *Journal of Psychophysiology*, 4, 381–390.
- Klimesch, Wolfgang, Schimke, H., & Pfurtscheller, G. (1993). Alpha frequency, cognitive load and memory performance. *Brain Topography*, 5(3), 241–251. <https://doi.org/10.1007/BF01128991>
- Knyazeva, M. G., Barzegaran, E., Vildavski, V. Y., & Demonet, J. F. (2018). Aging of human alpha rhythm. *Neurobiology of Aging*, 69, 261–273. <https://doi.org/10.1016/j.neurobiolaging.2018.05.018>
- Lawrence, M. A., & Lawrence, M. M. A. (2016). Package ‘ez.’ *R Package Version*, 4.
- Manza, P., Hau, C. L. V., & Leung, H.-C. (2014). Alpha Power Gates Relevant Information during Working Memory Updating. *Journal of Neuroscience*, 34(17), 5998–6002. <https://doi.org/10.1523/JNEUROSCI.4641-13.2014>
- Matthis, P., Scheffner, D., Benninger, C., Lipinski, C., & Stolz, L. (1980). Changes in the background activity of the electroencephalogram according to age. *Electroencephalography and Clinical Neurophysiology*, 49(5–6), 626–635.
- McCarthy, M. M., & Arnold, A. P. (2008). Sex differences in the brain: What’s Old and What’s New? In J. B. Becker, K. J. Berkley, N. Geary, E. Hampson, J. P. Herman, & E. Young (Eds.), *Sex differences in the brain: from genes to behavior* (pp. 15–33). Oxford university press.
- Mierau, A., Klimesch, W., & Lefebvre, J. (2017). State-dependent alpha peak frequency shifts: Experimental evidence, potential mechanisms and functional implications. *Neuroscience*, 360, 146–154. <https://doi.org/10.1016/j.neuroscience.2017.07.037>

- Noguchi, Y., & Kakigi, R. (2020). Temporal codes of visual working memory in the human cerebral cortex: Brain rhythms associated with high memory capacity. *NeuroImage*, 222(August), 117294. <https://doi.org/10.1016/j.neuroimage.2020.117294>
- Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, 2011. <https://doi.org/10.1155/2011/156869>
- Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging*, 17(2), 299–320. <https://doi.org/10.1037/0882-7974.17.2.299>
- Posthuma, D., Neale, M. C., Boomsma, D. I., & Geus, E. J. C. de G. (2001). Are Smarter Brains Running Faster? Heritability of Alpha Peak Frequency, IQ, and Their Interrelation. *Behavior Genetics*, 31(6), 567–579.
- R Core Team. (2018). *R: A Language and Environment for Statistical Computing*. <https://www.r-project.org/>
- Reuter-Lorenz, P. A., & Park, D. C. (2010). Human Neuroscience and the Aging Mind: at Old Problems A New Look. *Journals of Gerontology: Psychological Sciences*, 65B(4), 405–415. <https://doi.org/10.1093/geronb/gbq035>
- Riddle, J., Scimeca, J. M., Cellier, D., Dhanani, S., & D'Esposito, M. (2020). Causal Evidence for a Role of Theta and Alpha Oscillations in the Control of Working Memory. *Current Biology*, 30(9), 1748-1754.e4. <https://doi.org/10.1016/j.cub.2020.02.065>
- Rippon, G., Jordan-Young, R., Kaiser, A., & Fine, C. (2014). Recommendations for sex/gender neuroimaging research: Key principles and implications for research design, analysis, and interpretation. *Frontiers in Human Neuroscience*, 8(AUG), 1–13. <https://doi.org/10.3389/fnhum.2014.00650>
- Ronconi, L., Busch, N. A., & Melcher, D. (2018). Alpha-band sensory entrainment alters the duration of temporal windows in visual perception. *Scientific Reports*, 8(1), 1–10. <https://doi.org/10.1038/s41598-018-29671-5>
- Rousselet, G. A., Pernet, C. R., & Wilcox, R. R. (2017). Beyond differences in means: robust graphical methods to compare two groups in neuroscience. *European Journal of Neuroscience*, 46(2), 1738–1748. <https://doi.org/10.1111/ejn.13610>
- Roux, F., & Uhlhaas, P. J. (2014). Working memory and neural oscillations: Alpha-gamma versus theta-gamma codes for distinct WM information? *Trends in Cognitive Sciences*, 18(1), 16–25. <https://doi.org/10.1016/j.tics.2013.10.010>
- Rubin, J. D., Atwood, S., & Olson, K. R. (2020). Studying Gender Diversity. *Trends in Cognitive Sciences*, 24(3), 163–165. <https://doi.org/10.1016/j.tics.2019.12.011>
- Sauseng, P., Klimesch, W., Doppelmayr, M., Pecherstorfer, T., Freunberger, R., & Hanslmayr, S. (2005). EEG alpha synchronization and functional coupling during top-down processing in a

working memory task. *Human Brain Mapping*, 26(2), 148–155.  
<https://doi.org/10.1002/hbm.20150>

Scally, B., Burke, M. R., Bunce, D., & Delvenne, J. F. (2018). Resting-state EEG power and connectivity are associated with alpha peak frequency slowing in healthy aging. *Neurobiology of Aging*, 71, 149–155. <https://doi.org/10.1016/j.neurobiolaging.2018.07.004>

Smith, R. J. (1941). The frequency growth of the human alpha rhythms during normal infancy and childhood. *The Journal of Psychology*, 11(1), 177–198.

Tuladhar, A. M., Ter Huurne, N., Schoffelen, J. M., Maris, E., Oostenveld, R., & Jensen, O. (2007). Parieto-occipital sources account for the increase in alpha activity with working memory load. *Human Brain Mapping*, 28(8), 785–792. <https://doi.org/10.1002/hbm.20306>

Voyer, D., Voyer, S. D., & Saint-Aubin, J. (2017). Sex differences in visual-spatial working memory: A meta-analysis. *Psychonomic Bulletin and Review*, 24(2), 307–334.  
<https://doi.org/10.3758/s13423-016-1085-7>

Voytek, B., & Knight, T. (2015). Dynamic network communication as a unifying neural basis for cognition, development, aging, and disease. *Biological Psychiatry*, 77(12), 1089.  
<https://doi.org/10.1016/j.biopsych.2015.04.016>

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.  
<https://ggplot2.tidyverse.org>

Wickham, H., François, R., Henry, L., & Müller, K. (2018). *dplyr: A Grammar of Data Manipulation*. <https://cran.r-project.org/package=dplyr>

Wuttke, W., Arnold, P., Becker, D., Creutzfeldt, O., Langenstein, S., & Tirsch, W. (1975). Circulating hormones, EEG, and performance in psychological tests of women with and without oral contraceptives. *Psychoneuroendocrinology*, 1(2), 141–152. [https://doi.org/10.1016/0306-4530\(75\)90006-2](https://doi.org/10.1016/0306-4530(75)90006-2)

## Chapter 4

# Bimodal Spatial and Temporal Influences on Working Memory Performance

In the previous chapter I presented evidence for sex-based differences in the neural mechanisms underlying working memory for visual objects situated in space. In the present and following chapter, I move into an examination of differences in behavior that arise when visual working memory is employed for objects and locations not only structured in space, but also structured dynamically in time.

This chapter presents a study motivated by questions of how working memory for visuospatial information is impacted when *more* than visual information is present. In typical daily life not just visual, but also auditory, streams of information are present and available to be perceived. The task design employed to answer the questions posed herein is based upon the well-known *Corsi Blocks* paradigm - typically utilized to assess visuospatial working memory span (Berch et al., 1998; Corsi, 1972). In the Corsi Blocks paradigm, visuospatial targets are presented one at a time in a serial sequence such that

the spatial location of each target randomly varies while the temporal rate of presentation is constant. In the study herein, auditory stimulus streams (sequences) were presented in conjunction with visuospatial streams. In different task conditions, the presentation parameters of the auditory stimuli were manipulated such that their spatial and temporal features were either similar to, or different from, the spatial and temporal features of the visual stimuli.

The present study was motivated by the hypothesis that when streams of information in separate modalities interact, performance of a unimodal working memory task will be affected. Specifically, when *spatialized* visual and auditory stimulus streams temporally *coincide*, integration of information from both streams is thought to occur. When these streams do not temporally coincide, integration may be less, and cross-modal suppression might instead be facilitated. I hypothesized that dependent upon the congruence between bimodal streams, either positive or negative effects on visuospatial working memory performance will result.

## 4.1 Introduction

The environments we exist within offer up rich and ongoing sensory information, yet we typically pay attention to and make use of only a small subset of that information. Selective attention has a critical role in working memory processes, which we rely upon to achieve goal-driven tasks (van Ede & Nobre, 2021). What we attend to, therefore, can

have a large impact on the outcome of our performance. Where we allocate our attention, matters. In a busy city filled with talking people and rumbling vehicles, perhaps sounds coming from our left and right matters most in our determining whether to step out into the street. In a wooded forest with squirrels bounding over dry leaves and snakes creating a rustle as they carve out paths over the ground, if we are hoping to spot a rare bird perhaps only sounds from above will catch our interest. In a quiet library, where the turning of pages or clicking of fingers on a keyboard pose small violations to an otherwise silent atmosphere, if our goal is to remain focused on our studies, we may try to shift attention away from any sound in an effort to do so. During those moments in which our attention is focused and our working memory engaged, when both sights and sounds are available and potentially relevant, what information do we utilize? And to what effect on our performance? Herein we explore these questions as they relate to performance in a visuospatial working memory task when spatially parameterized stimuli in both visual and auditory modalities are present.

Sounds and sights may each provide us with potentially task-relevant information, though via different sensory modality streams. In cases where both modality streams seem to be reinforcing the same relevant information, it may be advantageous to integrate the two. For example, if there are cues available to us that different streams originate from the same singular source, by way of multimodal integration we may conceptually bind information from those streams together (Talsma et al., 2010).

Multisensory inputs that share temporal and spatial features, in particular, are likely to be integrated and processed as singular unified events (Welch, 1999).

This type of multimodal binding occurs automatically and without conscious effort — we do not have to think about it for it to happen. In part, this is because the brain has classes of neurons which are specialized for the rapid computation and processing of conjunctive events. Some multisensory neurons in superior colliculus, for example, are tuned to respond to temporal conjunctions between input from different modalities (Meredith et al., 1987). Cellular networks within the cerebellum, support discrimination between temporal occurrences (Ivry & Spencer, 2004). In primary sensory cortices, there exist unimodal and crossmodal neurons which become stimulated when neurons in other sensory regions are activated. These cells promote amplified responses to input in their native sensory modality when those inputs occur in temporal conjunction with inputs from other modalities (Allman & Meredith, 2007; Allman et al., 2008). At the population level, neuronal activity in primary auditory cortex can have a modulatory effect on activity in visual cortex (Mercier et al., 2013). These circuits of activity have been directly linked to facilitation of behavior (Stein et al., 1989).

Note that the terms “multimodal” and “bimodal” are used to indicate the presence of stimuli in different sensory modalities. At times, bimodal stimuli — such as a sight and a sound — may be attributed, correctly, to a singular source. At other times, however, bimodal stimuli may arise from difference sources. While the brain is well

adapted to rapidly bind discrete percepts together, to facilitate comprehension of them (Talsma et al., 2010), the stimuli themselves exist separately from any unifying representations we may form.

Past studies have found that when visual and auditory stimuli are available to us, the combination can benefit behavioral performance in visual tasks. A simple sound “pip” presented along with a busy visual scene, for example, can speed up a visual search for a target (Van der Burg et al., 2008). An oddball sound, embedded in a stream and presented in synchrony with a visual array can speed up visual detection (Vroomen & De Gelder, 2000). People are faster to detect location changes of bimodal visual and auditory stimuli, than of stimuli in either modality alone (Schröger & Widmann, 1998). At other times, however, the presence of multimodal streams can lead to errors. A single visual stimulus accompanied by a rapid succession of discrete sounds, for example, can generate the false perception of multiple visual stimuli despite the presentation of only one (Mishra et al., 2007; Shams et al., 2000). When participants must determine a direction of stimulus movement, if both visual and auditory stimuli are presented together, moving in opposite directions, the conflicting presentation has a negative impact on performance. Interestingly, however, it is only the case that performance is worse when reporting the direction of sound presented with conflicting visual information, but not vice-versa (Soto-Faraco et al., 2004). Errors in word perception can occur when the sound of words and the mouth movements of a speaker are in conflict (McGurk & Macdonald, 1976). Engaging

in a separate visual task while still viewing a speaker and listening to mismatches, though, eliminates these errors (Alsius et al., 2005).

The presence of sound may have different effects on performance in a visual task dependent upon either the demands of the task or the features of the sound. In a study by Marsh *et al.*, for example, participants were presented with a visual sequence of seven Navon letters to encode into verbal working memory while an irrelevant stream of sounds was simultaneously presented. When the focal demands of the task were increased – described in this study as high working memory load – performance suffered far less from the presence of an auditory stream than when working memory load was low. In a follow-up experiment, however, participants were exposed to a sound stream that continually changed while trying to encode a visually presented verbal sequence. The presence of this dynamically changing sound had a negative impact on performance regardless of working memory load (Marsh et al., 2020).

We have some degree of control over what information we prioritize and selectively attend to among multiple modality streams. Over time, repetitive and irrelevant sounds can be endogenously suppressed, for example, reducing the amount of attention directed to them (Todorovic et al., 2011). However, we are never able to flip a cognitive switch that can selectively cut off input from one modality entirely (Lorenc et al., 2021; Myers et al., 2017). Simple, irrelevant sounds that temporally co-occur with a visual target can alter subjective visual perception (Barry E. Stein et al., 1996). Covert

shifts of attention can be induced by auditory stimuli in modalities we are not actively or purposefully attending to (Driver & Spence, 1998). In these examples, unrelated information arising from separate modality streams can have crossmodal effects that arise from their juxtaposition. If one is performing a visuospatial task in which visual and auditory stimuli co-occur in synchrony, then, information from the auditory modality, despite being irrelevant, may indeed impact performance.

Seemingly automatic integration effects, however, can be modulated by attentional strategies. If bimodal streams co-occur but we deem them independent of one another based on conceptual or contextual knowledge, there are different ways in which we might process the information arising from each. For instance, if both streams might be independently relevant, we might shift attention between the two modality streams. Doing so, however, may lead to slowing of our responses, or to errors when the streams contain conflicting information (Lukas et al., 2010). Another approach we might take is to actively focus on one modality while passively experiencing or trying to suppress the other. With this approach, however, we may miss potentially helpful information from the unattended stream (Gherri & Eimer, 2011), or make errors as a result of conflicting information seeping through (Marsh et al., 2015). Errors of this latter form are just one example of an outcome resulting from *intersensory bias* – posited to be a means of reconciling discrepant perceptual information (Welch, 1999; Welch & Warren, 1980). In studies where there are discrepancies in locational information between visual and auditory streams, a bias toward the location derived from the visual content has often

been found (reviewed in Welch & Warren, 1980). Visual dominance has also been found in a number of studies in which stimuli are presented in singular instances, rather than in streams (Colavita, 1974; Diaconescu et al., 2013; Egeth & Sager, 1977; Lukas et al., 2010).

It may also be the case that different people process bimodal streams of information differently. In a study of visuospatial working memory by Gmeindl & Courtney (Gmeindl & Courtney, 2011), for example, participants completed a within-subjects task contrasting performance in the presence versus absence of sound. In the sound condition, while participants performed a visuospatial task in which a serial sequence of targets was presented, discrete noise bursts were presented in synchrony with the onset of each new visual target. Further, the spatial presentation of each noise burst was randomized such that the sound was either central, or completely or partially lateralized, in its presentation. Different results were found for different participants: in neurotypical individuals, the presence of the auditory stream indeed resulted in worse performance compared to silence – the spatially random sound was distracting. For individuals with a clinical multiple sclerosis diagnosis, however, performance was actually facilitated in the sound condition compared to when no sound was present (Gmeindl & Courtney, 2011).

If immersed in a primarily visual task, we might consider all sounds to be a distraction – as in the case of unexpected noises occurring while studying in a library. But yet in some instances, and for some individuals, even potentially conflicting sounds may

be helpful. If an event does indeed negatively impact or interfere with performance, objectively it can be classified as “distracting.” A sound causing distraction can typically be attributed to one of two things: either the sound pulled attentional resources away from where they were previously allocated, or it interfered with ongoing task-related processing of information (Hughes, 2014). Trying to remember an ongoing stream of visual information in the presence of an auditory stream that is continually changing, as in Gmeindl & Courtney (2011), is an example of the latter. As demonstrated in that study, however, distraction is not always clearly predictable.

Many studies that have informed our understanding of the behavioral impact of bimodal visual and auditory stimulus presentation have focused on effects of perception or attention in particular (Egeth & Sager, 1977; Van der Burg et al., 2008; Villena-González et al., 2016). In studies examining the impact of bimodal stimulus streams upon working memory, the effects upon memory for verbal information in particular have been well studied (Hughes et al., 2013; Hughes & Marsh, 2019; Marsh et al., 2015). The effects of bimodal presentation upon working memory for visuospatial information, however, have been far less examined. Within the described visuospatial working memory study that employed spatially random auditory distractors (Gmeindl & Courtney, 2011), distraction, rather than bimodal integration or interference in particular, was the subject of focus. The question, therefore, of how bimodal visual and auditory streams may affect visuospatial working memory performance remains an open one.

Summarizing what has been discussed above, temporal coincidence between bimodal stimuli in particular can encourage either multimodal binding or integration (Talsma et al., 2010). Further, if bimodal stimuli share properties such as spatial location, information from both modalities may be relevant and improve performance on a spatial task. On the other hand, distraction and impaired working memory performance can arise due to effects of incongruency between a target and an unrelated stimulus (Diaconescu et al., 2011, 2013; Lorenc et al., 2021). Studies that present visual and auditory stimulus streams concurrently have demonstrated that when task goals warrant attending primarily to the visual stream, an auditory stream that is changing unpredictably in space is likely to capture attention and impact performance (Gmeindl & Courtney, 2011; Marsh et al., 2020; Welch, 1999). But yet, many studies have found that for processing of spatial information in particular, vision is the dominant sense (Diaconescu et al., 2013; Egeth & Sager, 1977; Lukas et al., 2010). Further, processing of repetitive auditory stimuli with constant parameters (such as tone) and predictable timing, can be suppressed (Todorovic et al., 2011). Thus, it remains unclear under what circumstances irrelevant and potentially conflicting auditory stimuli will or will not interfere with working memory for spatial information in the visual modality.

We hypothesized that streams of sound providing spatial information will be integrated in working memory during performance of a visuospatial task, but only when those sounds temporally co-occur with visual targets. Further, we hypothesized that this synchrony-induced integration will have differential effects on performance dependent

upon spatial congruence. We predicted auditory stimuli that temporally co-occur with visual stimuli will facilitate performance when the spatial information provided by both are congruent, but hinder performance when the spatial information of each is incongruent, and therefore in conflict.

## 4.2 Methods

To test these hypotheses, we adapted the working memory paradigm utilized in Gmeindl & Courtney (2011). The task in this study was itself a modified version of the Corsi Blocks paradigm (Corsi, 1972), which has been employed for decades in the assessment of an individual's visuospatial working memory span (Berch et al., 1998; Monaco et al., 2013). Generally, in this paradigm, a sequence of squares ('blocks') are selected from an array and presented as memory targets, with one target square identified at a time. Participants must remember every target presented and reproduce the sequence following a brief delay.

This visuospatial working memory task is described in detail below. As in Gmeindl & Courtney (2011), spatial manipulations of sound stimuli were achieved by lateralizing their presentation using interaural intensity differences with headphones.

### 4.2.1 Participants

Young adult participants were recruited from the Johns Hopkins University student community for participation and were compensated via course credit.

Participants ranged in age between 18 and 28 years, and included 30 females (mean age = 19.8, SD = 1.20) and 30 males (mean age = 19.9, SD = 1.98), categorized based upon binary self-report of gender. Participants were screened via self-report for typical hearing in both ears, normal or corrected-to-normal vision, and lack of attention or memory related clinical diagnoses or traumatic brain injury.

#### 4.2.2 Stimuli and Task

##### *Auditory stimuli*

A single auditory stimulus was utilized for the task. This sound stimulus was a 300ms segment of pink noise, with a 10ms rise and fall at onset and offset, respectively. Pink noise, unlike white noise, has a decaying power frequency spectrum and therefore is akin to naturally occurring sounds, without being generally associated with any specific semantic meaning. As pink noise has been shown to elicit a strong exogenous response (Wetzel et al., 2016), it was deemed an adequately salient, yet semantically neutral, sound stimulus with which to test our hypotheses. Two versions of this stimulus were created, one with 100% left-, the other 100% right-channel lateralization. Sounds were created in Audacity® sound editing software (Audacity Team, 1999). Auditory stimuli were presented through Sennheiser HD 201 over-ear binaural headphones, at a volume well above threshold.

### *Visual Stimuli*

Visual stimuli consisted of 10 square “blocks” presented on a neutral gray colored background. Each block was dark gray in color and occupied approximately  $3.6^\circ$  in visual angle of both height and width. Viewing distance was fixed at approximately 59cm from the screen using a fixed-position chinrest. A configuration of five blocks was created wherein each block was separated by greater than  $2.0^\circ$  of visual angle in all directions. This five-block configuration was repeated on the left and right sides of the visual space, with a horizontal separation of approximately  $7.0^\circ$  of visual angle between them. On designation of a particular block as a *target*, the color of the block changed to cyan blue on the next screen refresh.

### *Task*

In **Figure 5** the on-screen configuration of visual stimuli is shown (A). For each trial a random sequence of six visual targets was selected from among the ten blocks within the visual array to be the *Sample Sequence*. A memory set size of six targets was determined to result in a suitable range of performance based upon pilot testing in this population. Targets were identified by an abrupt color change from grey to cyan blue, lasting 500ms. The visual inter-stimulus interval was 500ms, following which the next target in the sequence was presented. Following presentation of the entire Sample Sequence there was a 2.5s delay before presentation of the *Test Sequence*. The Test Sequence was comprised of six visual targets that were either identical to, or different from, the Sample Sequence. On every trial, participants were to determine whether all

targets within the Sample Sequence and Test Sequence matched in both location and order of presentation. Participants were cued to the onset of both the Sample Sequence and Test Sequence by a centrally presented fixation cross that lasted for a duration of 1s. Response was prompted by the display of “MATCH?” on screen until a keypress was made. No time limit was placed on responding. After a response was made, a 1s inter-trial interval preceded the start of the next trial. No feedback was provided.

During presentation of the Sample Sequence, on some trials an auditory stimulus (described above) was presented at the same time as each of the visual targets – with the onset of each occurring synchronously. In other trials, the onset of the auditory stimulus just preceded, or shortly followed, the onset of the visual stimulus. These trials are described as temporally *asynchronous*. Asynchronous auditory stimulus onsets randomly occurred either 100ms, 200ms, or 300ms before, or after, onset of each visual target.

*Synchronous* and *asynchronous* trials were further defined by the spatial congruence between visual and auditory stimuli. In some trials, auditory stimuli were lateralized congruently with the spatial lateralization of the visual stimuli. These are described as *congruent* trials. In other trials, referred to as *incongruent* trials, the lateralization of each auditory stimulus was opposite of its corresponding visual stimulus in the sequence. These spatial relationships between auditory and visual stimuli are shown in section A of **Figure 5**.

During the Test Sequence, only visual stimuli were presented, with three possible *Test Types*. One third of the Test Sequences matched the Sample Sequence in both the identity of all 6 targets and their order of presentation. One third of the Test Sequences mismatched the Sample Sequence in identity. In this Test Type a single target block was randomly selected from among those outside of the Sample Sequence and displayed in the third or fourth (randomly determined on each trial) position of the Test Sequence. The final third of Test Sequences were a mismatch in order. For mismatched order trials, the third and fourth targets within the Test Sequence were swapped in their ordinal positions. No instructions were provided to participants regarding trial types or auditory stimuli.

The task included a total of 75 trials. Sixty of the trials were bimodal, with both auditory and visual stimuli presented during the Sample Sequence. Fifteen trials were unimodal; only visual stimuli were presented. The 60 bimodal trials were divided among four conditions. As previously described, there were four Sample Conditions: Synchronous-Congruent, Synchronous-Incongruent, Asynchronous-Congruent, and Asynchronous-Incongruent. Thus, there were 15 trials per Sample Condition, equally split among the three Test Types (5 Match, 5 Order-mismatch, and 5 Identity-mismatch trials).

The task was programmed in Python 3 (Python Software Foundation, 2008) utilizing the PsychoPy 3.2 (Peirce, 2007) library of code and software for delivery.

### 4.2.3 Procedure

After providing consent and receiving verbal and on-screen instructions and training, participants performed the task in a private booth seated in front of a computer screen with headphones on. For each participant, trials from each described Sample Condition were intermixed in a pseudorandomized order. Trials were separated into five experimental blocks with self-paced breaks permitted between blocks. On completion of the task, participants were given a demographics questionnaire and then guided by the experimenter through a brief exit interview. The duration of the entire study session was approximately one hour.

### 4.2.4 Design

The study utilized a within-subjects, repeated-measures design. Randomized presentation of trials over the duration of the experiment, rather than a blocked design, was utilized to minimize habituation to any specific temporal or spatial occurrence of sound. Accuracy (proportion of correct responses) was the primary dependent variable of interest.

In bimodal trials, auditory and visual stimuli were presented only during the Sample Sequence to allow for study and interpretation of effects as they relate specifically to processes of working memory encoding. The Test Sequence always included only visual stimuli. Unimodal trials, in which only visual stimuli were presented during the Sample

Sequence, provided a baseline with which to further compare and interpret results related to our hypotheses.

To test our hypotheses, parametric group-level comparisons were performed using omnibus ANOVA tests with Accuracy as the primary outcome variable. A 2x2 (Synchrony x Congruence) ANOVA test was performed to isolate the effects of different bimodal Sample Conditions. A 2x2x3 (Synchrony x Congruence x Test Type) ANOVA test was performed to determine whether additional effects of Test Type were present and interacting with Sample Conditions. In cases where violations of sphericity were present, corrected statistics are reported. A 1-way ANOVA comparing all five Sample Conditions (4 bimodal, 1 unimodal) was performed. ANOVA tests were followed, when relevant, with paired *t*-tests to examine underlying effects. Multiple comparisons correction of *t*-test statistics was performed via the False Discovery Rate method (Benjamini & Hochberg, 1995).

Usage of a Yes/No recognition memory response procedure allowed for analysis of condition-specific responses using basic signal detection methods. Match trials in which participants responded “Yes” were marked as *Hits*. Mismatch trials in which participants responded “Yes” were marked *False Alarms*. Proportions of Hits and False Alarms were utilized to calculate  $d'$  as a measure of an individual's sensitivity to detect a match;  $\beta$  their criterion in making a Yes/No decision (Green & Swets, 1974).

All statistical analysis was performed in R (R Core Team, 2018) with the *dplyr* (Wickham et al., 2021), *ez* (Lawrence, 2016), and *psycho* (Makowski, 2018) packages. Figures were produced in R with the *ggplot2* (Wickham, 2016), *hrbrthemes* (Rudis, 2018), and *cowplot* (Wilke et al., 2019) packages.

## 4.3 Results

Of the data collected, one participant was excluded from further analysis due to an attentional disorder and three due to failure to perform the task (below 60% overall Accuracy). Data analyzed was from 56 participants (27 female, 29 male).

Our working hypotheses was that incongruence of spatial information between auditory and visual stimuli has a differential effect on spatial working memory dependent on whether sounds are synchronous with visual stimuli. A 2x2 (Congruence x Synchrony) omnibus ANOVA with Accuracy as the outcome variable resulted in a trending main effect of Congruence,  $F(1,55) = 3.33$ ,  $p < 0.075$ ,  $\eta_G^2 < 0.01$ , and a significant interaction between Synchrony and Congruence,  $F(1,55) = 10.28$ ,  $p < 0.005$ ,  $\eta_G^2 = 0.03$ . This interaction can be seen in **Figure 6**. Follow-up paired *t*-tests were performed contrasting Accuracy between conditions. After correction for multiple comparisons, the difference in Accuracy between Synchronous-Congruent versus Synchronous-Incongruent conditions was highly significant,  $t = 3.7$ ,  $p < 0.002$ , 95% C.I. [0.027,0.092]. That is, performance when bimodal

stimuli synchronously co-occurred *and* exhibited spatial congruence was significantly greater than when stimuli synchronously co-occurred but were in spatial conflict. When sounds were presented *asynchronously* with the visual stimuli, the Accuracy difference between Congruent and Incongruent Conditions was not significant.

Comparison of Accuracy between Synchrony conditions when bimodal streams were spatially congruent was only a trending effect,  $t = 1.78$ ,  $p < 0.08$ , 95% C.I. [-0.004, 0.071]. For spatially congruent stimuli streams, temporal synchrony resulted in a marginal improvement of performance compared to temporal asynchrony. Alternatively, comparison of Accuracy between Synchrony conditions when bimodal streams were in spatial conflict remained significant after correction,  $t = -2.46$ ,  $p < 0.03$ , 95% C.I. [-0.082, -0.008]. Performance was better, despite the spatial conflict, when bimodal streams were temporally asynchronous compared to synchronous. Temporal synchrony coupled with spatial conflict resulted in the worst performance. These results are shown in **Figure 6**.

A coarse comparison of visuospatial working memory performance in the context of bimodal (auditory and visual) versus unimodal (visual only) stimuli was performed by t-test between Accuracy in the unimodal condition versus Accuracy averaged across all bimodal conditions. No overall difference in performance was present,  $t = 1.13$ ,  $p > 0.2$ . A detailed comparison was performed via one-way ANOVA contrasting Accuracy across all conditions. The overall effect of Condition on Accuracy was significant:  $F(4,220) = 3.74$ ,  $p < 0.006$ ,  $\eta_G^2 = 0.032$ . Among the five Conditions, most relevant to our hypothesis were the

direct contrasts of unimodal versus bimodal temporally synchronous conditions. Put otherwise — the comparison in performance of the visual task when no-sound versus synchronously occurring sound was present. These contrasts were performed via paired *t*-tests, corrected for multiple comparisons. Accuracy in the case of Synchronous-Incongruent bimodal stimuli was significantly worse than Accuracy for unimodal visual stimuli only,  $t = -2.73$ ,  $p < 0.02$ , 95% C.I. [-0.078, -0.012]. Accuracy for Synchronous-Congruent bimodal stimuli, while numerically greater, was not statistically different from Accuracy given unimodal stimuli,  $t = 1.062$ ,  $p > 0.2$ .

Throughout the course of the experiment, participants were presented with trials of three different Test Types. The sequence presented during the Test period either matched the Sample Sequence or mismatched it in *Identity* or *Order*. To determine how Test Type contributed to the above effects on performance, a 2x2x3 (Synchrony x Congruence x Test Type) ANOVA was performed. In addition to the same effects of the previously reported 2x2 ANOVA, a main effect of Test Type,  $F(2,110) = 38.06$ ,  $p < 0.001$ ,  $\eta_G^2 = 0.16$ , was present. A 3-way interaction did not exist. However, significant interactions between Congruence and Test Type,  $F(2, 110) = 4.04$ ,  $p < 0.03$ ,  $\eta_G^2 < 0.01$ , and between Synchrony and Test Type,  $F(2,110) = 4.19$ ,  $p < 0.03$ ,  $\eta_G^2 < 0.01$ , were present. These results are shown in **Figure 7**. Mean Accuracy is over 90% across conditions for *Identity*-mismatch Test Type trials. Mean Accuracy for both Match and *Order*-mismatch

Test Type trials ranged between 70-85%. Again, the contrasts most relevant to our hypothesis were between the different spatial Congruence Conditions when stimuli were Synchronous. While the pattern of performance in Synchronous conditions was the same for both Match and Order-mismatch Test Types, the difference between Congruent and Incongruent spatial Conditions given an *Order*-mismatch Test Type was borderline marginal,  $t = 1.67$ ,  $p = 0.1$ , 95% C.I. [-0.011, 0.118]. The difference between these conditions given a Match Test Type, however, was large,  $t = 3.84$ ,  $p < 0.001$ , 95% C.I. [0.050, 0.158].

Average sensitivity in detecting a Match, operationalized as  $d'$ , was computed and compared across Conditions. A significant difference in sensitivity was present based upon Condition:  $F(4,220) = 3.48$ ,  $p < 0.01$ ,  $\eta_G^2 = 0.03$ . Sensitivity was similar between Unimodal (visual stimuli only) and Bimodal Asynchronous Sample Condition trials. In Bimodal Synchronous trials, however, participants exhibited reduced sensitivity for spatially incongruent compared to congruent stimuli:  $t = -3.81$ ,  $p < 0.001$ , 95% C.I. [-0.607, -0.188].

The decision criterion,  $\beta$ , was found to significantly differ across Conditions,  $F(4,220) = 3.12$ ,  $p < 0.02$ ,  $\eta_G^2 = 0.03$ , but the effect is not driven by Conditions of Temporal Synchrony. Between the Synchronous conditions, there was little change in the average decision criterion,  $\beta$ , between spatial conditions, or compared to Unimodal visual-only stimuli:  $t < 1.0$ ,  $p > 0.3$  for each comparison. Participants were not increasingly biased to

respond *Yes* or *No* in the presence of synchronous sound. These results are shown in **Figure 8**.

## 4.4 Discussion

People utilize signals of timing and spatialization to selectively allocate attention and make inferences about the environment (Nobre & van Ede, 2018; Theeuwes, 2014). These signals are sources of information. Studies of perception and attention have found that distinct signals arising from different sensory modalities can have a strong effect on behavior, particularly when those signals occur at the same time. In the present study, we present evidence that these effects extend to visuospatial working memory.

A significant difference in performance resulted when bimodal streams had synchronous compared to asynchronous temporal structure. We consider two possible ways by which information representation in working memory may have been affected, that account for this difference. First, by way of multimodal binding. Second, by way of reinforcement or interference.

Within synchronous trials, because auditory and visual information temporally coincided it may be the case that the two modality streams were representationally bound together (Talsma et al., 2010). For each target presented, then, the representation of that target's location would be formed based upon both visual and auditory information. In the case of spatially congruent bimodal streams, only accurate target

location information would be available for representational binding, facilitating memory. In the case of spatially incongruent bimodal streams, only the visual stream would provide accurate location information while the auditory stream would be misleading. Binding auditory and visual streams in these instances would diminish the precision with which location was represented, hindering memory.

While this account does fit the resulting behavior observed, with Accuracy in task performance being greater for Synchronous-Congruent trials compared to Synchronous-Incongruent trials, there are reasons this might be questioned. While both visual and auditory streams shared similarity in that they exhibited spatial and temporal structure, for example, they were not unified in any other obvious means – such as semantic connection. Pink noise, rather than an instrumental tone or an animal call was specifically utilized in this study for its lack of semantic meaning. Further, auditory stimuli were always presented through headphones. Previous studies have demonstrated that binding between auditory and visual stimuli is less likely when the two do not emanate from “out there” in space (Soto-Faraco et al., 2004). Another account then, is that bimodal streams were represented separately.

As reviewed in Chapter 2, decades ago it was demonstrated that people were able to remember and report fundamental features of sounds presented to them even when explicitly focusing their attention elsewhere (Treisman, 1964). In the present study, the spatial location of auditory stimuli may have been a sufficiently low-level feature for

participants to perceive and encode, even if not explicitly directing attention to the auditory modality. When drawing upon memory during task performance, then, representations formed from visual stimuli may have been predominant, but auditory representations that had been separately formed and stored also had influence.

In considering the patterns of performance tied to different Test Types, of note is that the highest accuracy was achieved across conditions when *sample* and *test* target identities mismatched (see **Figure 7**). Rather than separately encoding *sample* and *test* sequences, and then internally comparing representations between the two, reliance upon recognition memory is a far less demanding approach. An entirely mismatching target, then, would be readily identified as unfamiliar. In Match and Order-mismatch Test Types, performance was far lower.

Previous studies have demonstrated that *visuospatial* working memory in particular is more prone to errors of order in representing sequences of information than are other forms of working memory (e.g. Gmeindl et al., 2011). While the direction of difference between Accuracy in Synchronous-Congruent and Synchronous-Incongruent trials holds when it comes to Order-mismatch Test trials, the effect is marginal. There is a stronger effect of temporal structure in the case of Congruent trials, however, which is interesting. Performance is by far worst for trials in which bimodal stimuli were spatially congruent, but temporally asynchronous. Although the bimodal stimuli occurring in closest temporal proximity within Asynchronous-congruent trials were indeed spatially

congruent, they may have been spatially incongruent relative to sequentially neighboring targets. Put otherwise, a single auditory stimulus in these trial types may have been spatially incongruent with an earlier, or later, visual stimulus. It is possible, then, that this effect may have arisen from proactive interference occurring across modalities. When auditory stimuli were asynchronous with visual stimuli, each auditory stimulus onset occurred randomly either before or after onset of each visual stimulus. In cases where the auditory stimulus was presented early, perceiving it may have generated interference in processes of consolidation related to the preceding visual target. In cases where the auditory stimulus was presented “late,” it may have interfered with subsequent visual target encoding. When spatial incongruence was present with earlier or later targets, errors in memory for order may have resulted.

The largest Congruence-based difference within Synchronous trials arose within Match Test Type trials (**Figure 7**). Indeed, in the case of Synchronous-Incongruent trials, participants exhibited the lowest sensitivity to detect a match (**Figure 8A**). As can be seen in **Figure 8B**, however, this was not due to a shift in decision Criterion to make a Yes/No response. Put otherwise, spatial conflict between synchronous bimodal stimuli did not lead participants to become more or less conservative in selecting a response than they were for Synchronous-Congruent or Asynchronous-Incongruent conditions. Only among Asynchronous-Congruent trials was there a comparatively marked shift in decision Criterion indicating a bias towards a Yes response.

As demonstrated by the behavioral differences across Sample Conditions and Test Types discussed, performance of a visuospatial working memory task was strongly influenced by the presence of information in the auditory modality. Synchronous inputs from different modalities may have generated some competition for attention. Even though visual dominance has been repeatedly demonstrated (e.g. Egeth & Sager, 1977), there could at times be an outcome of competitive selection that favors dedicating neural resources to processing information arising from the auditory modality (Desimone & Duncan, 1995; Knudsen, 2007). In the case of a dynamically changing auditory stream, change in the properties or features of specific sounds may lead to exogenous generation of bias and ultimately outcompete a visual stream for allocation of attention. On the other hand, endogenous signals, particularly those related to control processes, may win the resource “tug of war”, keeping attention oriented in the most goal-appropriate direction. The sound stimulus used in the present study did not change within a trial, but the timing and lateralization of its presentation did. The impact was not so great as to drastically harm performance – Accuracy across Conditions was greater than 70%, but it clearly had an effect.

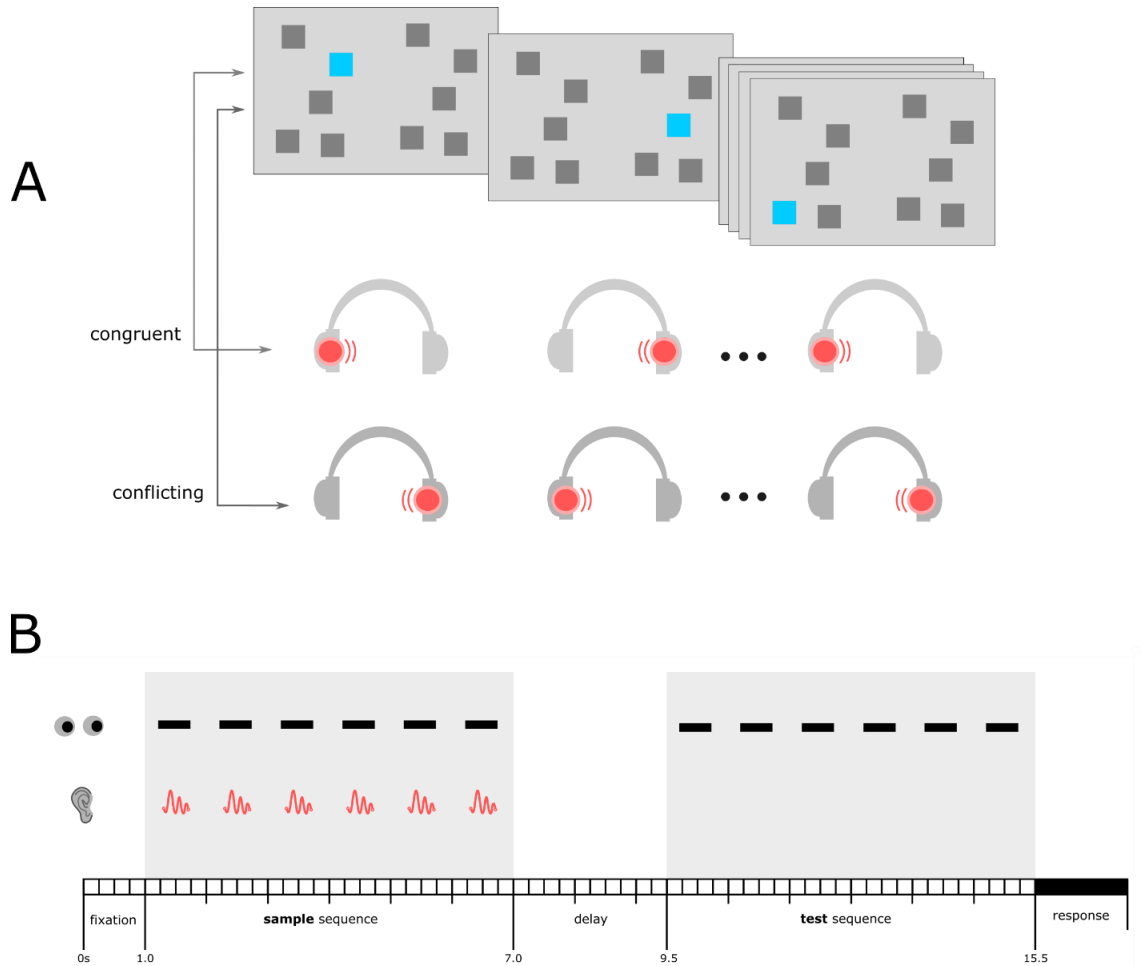
If the sound itself was consistently distracting, performance in Bimodal compared to Unimodal conditions should have suffered. This was not the case. Only when the sound was synchronous and spatially incongruent was there an effect of distraction – performance was worse in this case than in other Bimodal Conditions and worse than when no sound was present at all. Indeed  $d'$ , sensitivity, was also lowest in this

circumstance. This reinforces the similar finding when examining Accuracy by Test Type – given Synchronous and Incongruent sound, participants were least capable of distinguishing matches.

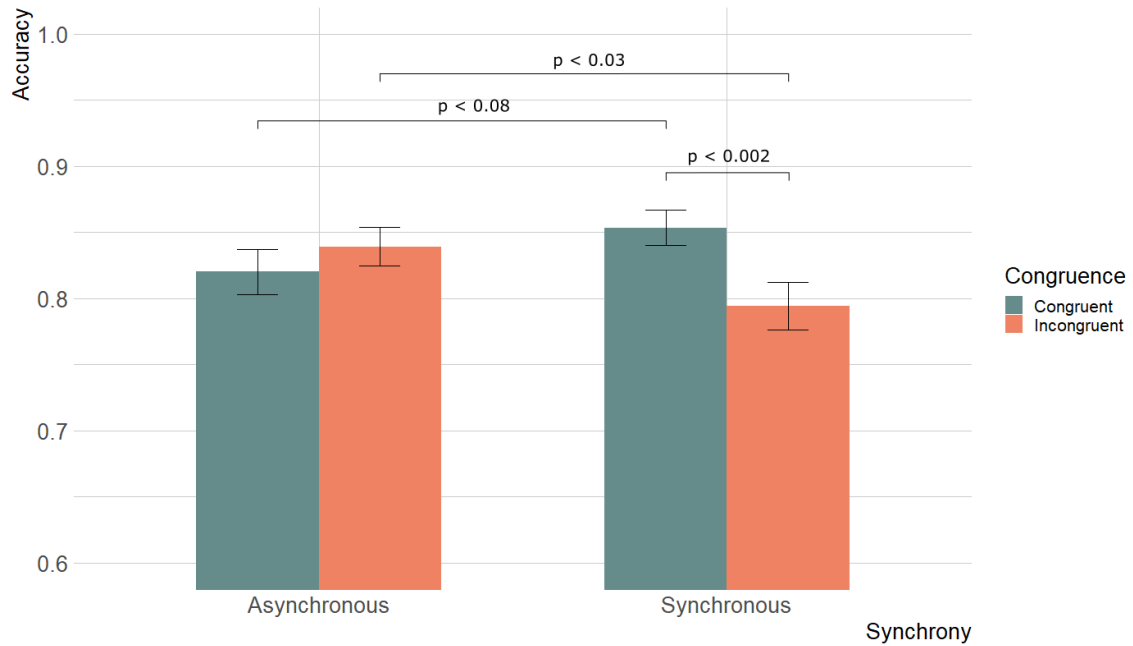
## 4.5 Conclusion

We hypothesized that the spatial information provided by bimodally presented visual and auditory stimuli would have the greatest impact on working memory performance when those stimuli occurred together in time. Indeed, this is exactly what we found. The effect, however, was not strictly beneficial. Instead, it differed based upon the congruence or incongruence of the spatial information provided by each sensory stream. While it may not be the case that information from each modality was integrated into a *single* representation, then stored in working memory, there is strong indication that representation arising from the perception of each modality did indeed impact performance. Temporal and spatial structure within, and across modalities, effect working memory representation. Future neuroimaging studies might shed light on the neural processing underlying this behavior.

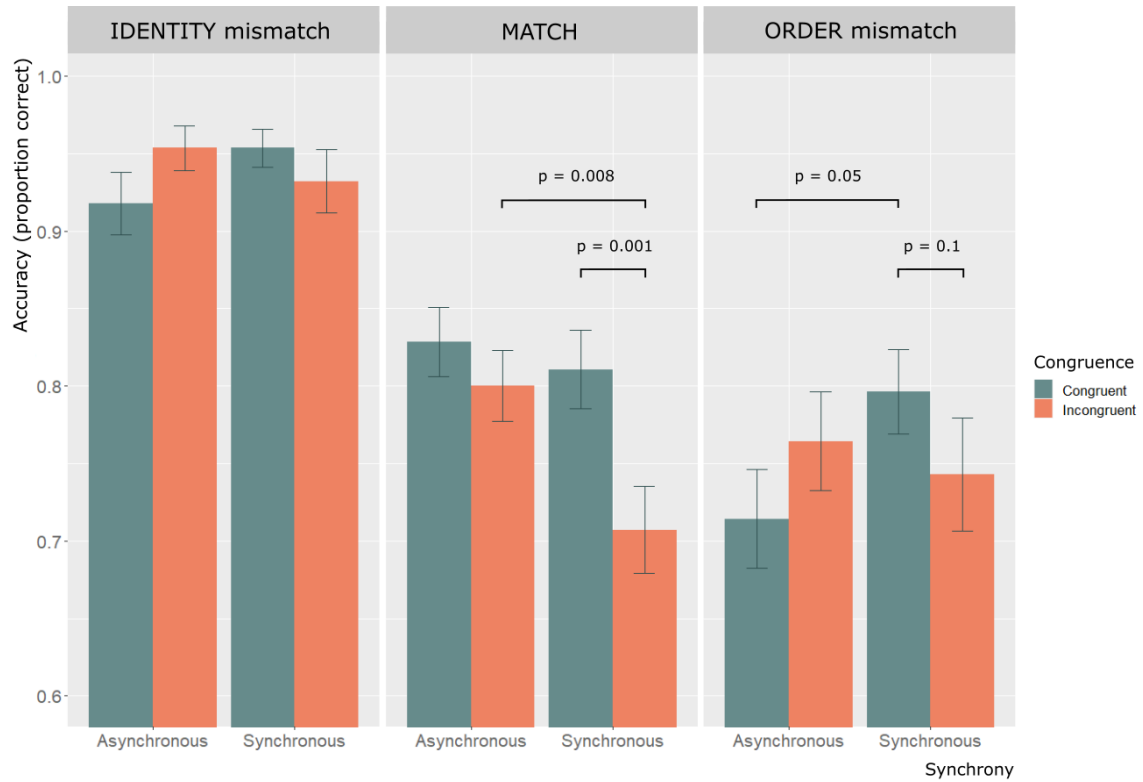
## 4.6 Figures



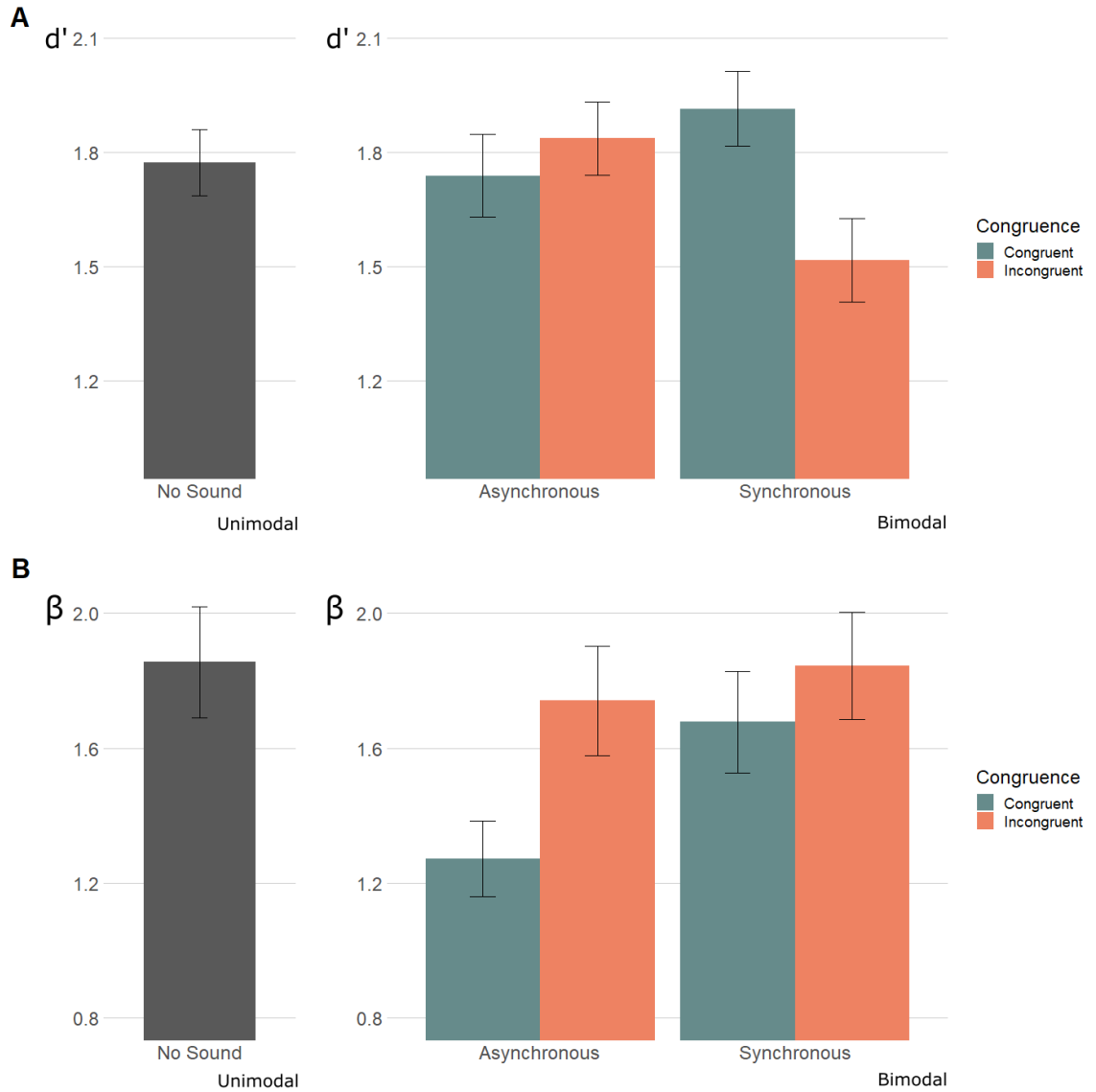
**Figure 5:** Overview of Task Conditions and Trial Timeline. **A:** Example of a *Sample Sequence* in which both visual and auditory stimulus streams are presented. Graphic depiction demonstrates the *spatial* relationship between visual targets (upper) and auditory (lower) stimuli in *Congruent* and *Incongruent* conditions. On every trial, six unique visual targets (cyan blue squares), selected at random, were presented. Ellipses and overlapping screens represent portions of the sequence presented during the experiment but not included in this visualization. **B:** Trial Timeline. Shaded rectangular regions indicate time periods of stimulus stream presentation. During presentation of the Sample Sequence, both visual and auditory stimulus streams were presented. During presentation of the Test Sequence, only visual stimuli were presented. In all task conditions, Participants were instructed to remember the Sample Sequence. After presentation of the Test Sequence, participants made a Yes or No response by keypress to indicate whether Sample and Test sequences matched. No instruction was given regarding the auditory stimuli.



**Figure 6:** Accuracy (proportion correct) within bimodal trials by Sample Condition. Error bars indicate standard errors of the mean. Performance across Conditions in which bimodal stimuli were spatially Congruent are shown in deep turquoise; spatially Incongruent are shown in salmon. A significant interaction is present between Synchrony and Congruence,  $F(1.55) = 10.28$ ,  $p < 0.005$ . Follow-up contrast statistics are shown.



**Figure 7:** Performance (Accuracy) across bimodal trials by Sample Condition and Test Type. Trials were evenly split between the three Test Types. While the general pattern of Accuracy within Synchronous-Congruent trials being higher than Accuracy within Synchronous-Incongruent trials holds across Test Types, the difference between these conditions was greatest among trials in which the Test and Sample sequences matched.



**Figure 8:** Comparison of participant sensitivity,  $d'$ , (**A**) and decision criterion,  $\beta$ , (**B**) across unimodal and bimodal Conditions. Main effects of Condition are present for both  $d'$  and  $\beta$  outcome variables. The largest difference in sensitivity is present between the Synchronous-Congruent and Synchronous-Incongruent Bimodal conditions, indicating greater and lesser sensitivity, respectively, to detect Matching Sample and Test sequences. This reinforces the results shown in **Figure 7**. As shown in (**B**), however, this was not driven by a change in decision criterion — these two conditions did not generate significantly more or less bias to make a Yes/No response. That  $\beta$  was lowest for Asynchronous-Congruent trials indicates that participants were more biased to respond Yes, further shedding light on the differences in Accuracy by Test Type (**Figure 7**) within this condition.

## 4.7 References

- Allman, B. L., & Meredith, M. A. (2007). Multisensory Processing in “Unimodal” Neurons: Cross-Modal Subthreshold Auditory Effects in Cat Extrastriate Visual Cortex. *Journal of Neurophysiology*, 98(1), 545–549. <https://doi.org/10.1152/jn.00173.2007>
- Allman, Brian L., Bittencourt-Navarrete, R. E., Keniston, L. P., Medina, A. E., Wang, M. Y., & Meredith, M. A. (2008). Do cross-modal projections always result in multisensory integration? *Cerebral Cortex*, 18(9), 2066–2076. <https://doi.org/10.1093/cercor/bhm230>
- Alsius, A., Navarra, J., Campbell, R., & Soto-Faraco, S. (2005). Audiovisual integration of speech falters under high attention demands. *Current Biology*, 15(9), 839–843. <https://doi.org/10.1016/j.cub.2005.03.046>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57, 289–300. <https://doi.org/10.2307/2346101>
- Berch, D. B., Krikorian, R., & Huha, E. M. (1998). The Corsi block-tapping task: Methodological and theoretical considerations. *Brain and Cognition*, 38(3), 317–338. <https://doi.org/10.1006/brcg.1998.1039>
- Colavita, F. B. (1974). Human sensory dominance. *Perception & Psychophysics*, 16(2), 409–412. <https://doi.org/10.3758/BF03203962>
- Corsi, P. M. (1972). *Human memory and the medial temporal region of the brain*. McGill University, Montreal, QB.
- Desimone, R., & Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. *Annual Review of Neuroscience*, 18, 193–222. <https://doi.org/10.1146/annurev-psych-122414-033400>
- Diaconescu, A. O., Alain, C., & McIntosh, A. R. (2011). The co-occurrence of multisensory facilitation and cross-modal conflict in the human brain. *Journal of Neurophysiology*, 106(6), 2896–2909. <https://doi.org/10.1152/jn.00303.2011>
- Diaconescu, A. O., Hasher, L., & McIntosh, A. R. (2013). Visual dominance and multisensory integration changes with age. *NeuroImage*, 65, 152–166. <https://doi.org/10.1016/j.neuroimage.2012.09.057>
- Driver, J., & Spence, C. (1998). Cross-modal links in spatial attention. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 353(1373), 1319–1331. <https://doi.org/10.1098/rstb.1998.0286>
- Egeth, H. E., & Sager, L. C. (1977). On the locus of visual dominance. *Perception & Psychophysics*, 22(1), 77–86. <https://doi.org/10.3758/BF03206083>
- Gherri, E., & Eimer, M. (2011). Active listening impairs visual perception and selectivity: An ERP study of auditory dual-task costs on visual attention. *Journal of Cognitive Neuroscience*,

- 23(4), 832–844. <https://doi.org/10.1162/jocn.2010.21468>
- Gmeindl, L., & Courtney, S. M. (2011). Deconstructing Spatial Working Memory and Attention Deficits in Multiple Sclerosis. *Neuropsychology*, 26(57). <https://doi.org/10.1037/a0026213>
- Green, D. M., & Swets, J. A. (1974). *Signal Detection Theory and Psychophysics*. R.E. Krieger Pub. Co.
- Hughes, R. W. (2014). Auditory distraction: A duplex-mechanism account. *PsyCh Journal*, 3(1), 30–41. <https://doi.org/10.1002/pchj.44>
- Hughes, R. W., Hurlstone, M. J., Marsh, J. E., Jones, D. M., & Vachon, F. (2013). Cognitive control of auditory distraction: Impact of task difficulty, foreknowledge, and working memory capacity supports duplex-mechanism account. *Journal of Experimental Psychology: Human Perception and Performance*, 39(2), 539–553. <https://doi.org/10.1037/a0029064>
- Hughes, R. W., & Marsh, J. E. (2019). When Is Forewarned Forearmed? Predicting Auditory Distraction in Short-Term Memory. *Journal of Experimental Psychology: Learning Memory and Cognition*, 46(3), 427–442. <https://doi.org/10.1037/xlm0000736>
- Ivry, R. B., & Spencer, R. M. C. (2004). The neural representation of time. *Current Opinion in Neurobiology*, 14(2), 225–232. <https://doi.org/10.1016/j.conb.2004.03.013>
- Knudsen, E. I. (2007). Fundamental components of attention. *Annual Review of Neuroscience*, 30, 57–78. <https://doi.org/10.1146/annurev.neuro.30.051606.094256>
- Lawrence, M. A. (2016). Package ‘ez.’ In *R package 3.1 version* (4.4-0). <https://github.com/mike-lawrence/ez>
- Lorenc, E. S., Mallett, R., & Lewis-Peacock, J. A. (2021). Distraction in Visual Working Memory: Resistance is Not Futile. *Trends in Cognitive Sciences*, 25(3), 228–239. <https://doi.org/10.1016/j.tics.2020.12.004>
- Lukas, S., Philipp, A. M., & Koch, I. (2010). Switching attention between modalities: Further evidence for visual dominance. *Psychological Research*, 74(3), 255–267. <https://doi.org/10.1007/s00426-009-0246-y>
- Makowski, D. (2018). The psycho package: an efficient and publishing-oriented workflow for psychological science. In *Open Source Software* (Vol. 3, Issue 22, p. 470).
- Marsh, J. E., Campbell, T. A., Vachon, F., Taylor, P. J., & Hughes, R. W. (2020). How the deployment of visual attention modulates auditory distraction. *Attention, Perception, and Psychophysics*, 82(1), 350–362. <https://doi.org/10.3758/s13414-019-01800-w>
- Marsh, J. E., Sörqvist, P., & Hughes, R. W. (2015). Dynamic cognitive control of irrelevant sound: Increased task engagement attenuates semantic auditory distraction. *Journal of Experimental Psychology: Human Perception and Performance*, 41(5), 1462–1474. <https://doi.org/10.1037/xhp0000060>
- McGurk, H., & Macdonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264(5588), 746–748. <https://www-nature->

com.libproxy1.usc.edu/articles/264746a0.pdf%0Ahttps://www.nature.com/articles/264746a0.pdf

- Mercier, M. R., Foxe, J. J., Fiebelkorn, I. C., Butler, J. S., Schwartz, T. H., & Molholm, S. (2013). Auditory-driven phase reset in visual cortex : Human electrocorticography reveals mechanisms of early multisensory integration. *NeuroImage*, 79, 19–29. <https://doi.org/10.1016/j.neuroimage.2013.04.060>
- Meredith, M. A., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 7(10), 3215–3229. <https://doi.org/10.1523/jneurosci.07-10-03215.1987>
- Mishra, J., Martinez, A., Sejnowski, T. J., & Hillyard, S. A. (2007). Early cross-modal interactions in auditory and visual cortex underlie a sound-induced visual illusion. *Journal of Neuroscience*, 27(15), 4120–4131. <https://doi.org/10.1523/JNEUROSCI.4912-06.2007>
- Monaco, M., Costa, A., Caltagirone, C., & Carlesimo, G. A. (2013). Forward and backward span for verbal and visuo-spatial data: Standardization and normative data from an Italian adult population. *Neurological Sciences*, 34(5), 749–754. <https://doi.org/10.1007/s10072-012-1130-x>
- Myers, N. E., Stokes, M. G., & Nobre, A. C. (2017). Prioritizing Information during Working Memory: Beyond Sustained Internal Attention. *Trends in Cognitive Sciences*, 21(6), 449–461. <https://doi.org/10.1016/j.tics.2017.03.010>
- Nobre, A. C., & van Ede, F. (2018). Anticipated moments: Temporal structure in attention. *Nature Reviews Neuroscience*, 19(1), 34–48. <https://doi.org/10.1038/nrn.2017.141>
- Peirce, J. W. (2007). PsychoPy-Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1–2), 8–13. <https://doi.org/10.1016/j.jneumeth.2006.11.017>
- R Core Team. (2018). *R: A Language and Environment for Statistical Computing*. <https://www.r-project.org/>
- Rudis, B. (2018). Hrbthemes: Additional themes, theme components and utilities for “ggplot2.” In *Hrbthemes Documentation*.
- Schröger, E., & Widmann, A. (1998). Speeded responses to audiovisual signal changes result from bimodal integration. *Psychophysiology*, 35(6), 755–759. <https://doi.org/10.1017/S0048577298980714>
- Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions. What you see is what you hear. *Nature*, 408(December), 788.
- Soto-Faraco, S., Kingstone, A., & Spence, C. (2004). Cross-Modal Dynamic Capture: Congruency Effects in the Perception of Motion Across Sensory Modalities. *Journal of Experimental Psychology: Human Perception and Performance*, 30(2), 330–345. <https://doi.org/10.1037/0096-1523.30.2.330>

- Stein, B. E., Meredith, M. A., Huneycutt, W. S., & McDade, L. (1989). Behavioral indices of multisensory integration: Orientation to visual cues is affected by auditory stimuli. *Journal of Cognitive Neuroscience*, 1(1), 12–24. <https://doi.org/10.1162/jocn.1989.1.1.12>
- Stein, Barry E., London, N., Wilkinson, L. K., & Price, D. D. (1996). Enhancement of perceived visual intensity by auditory stimuli: A psychophysical analysis. *Journal of Cognitive Neuroscience*, 8(6), 497–506. <https://doi.org/10.1162/jocn.1996.8.6.497>
- Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences*, 14(9), 400–410. <https://doi.org/10.1016/j.tics.2010.06.008>
- Theeuwes, J. (2014). Spatial orienting and attentional capture. In A. C. Nobre & S. Kastner (Eds.), *The Oxford Handbook of Attention* (pp. 231–252). Oxford University Press Oxford.
- Todorovic, A., van Ede, F., Maris, E., & de Lange, F. P. (2011). Prior expectation mediates neural adaptation to repeated sounds in the auditory cortex: An MEG study. *Journal of Neuroscience*, 31(25), 9118–9123. <https://doi.org/10.1523/JNEUROSCI.1425-11.2011>
- Treisman, A. M. (1964). Selective attention in man. *British Medical Bulletin*, 20(1), 12–16. <https://doi.org/10.1093/oxfordjournals.bmb.a070274>
- Van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W., & Theeuwes, J. (2008). Pip and Pop: Nonspatial Auditory Signals Improve Spatial Visual Search. *Journal of Experimental Psychology: Human Perception and Performance*, 34(5), 1053–1065. <https://doi.org/10.1037/0096-1523.34.5.1053>
- van Ede, F., & Nobre, A. C. (2021). Toward a neurobiology of internal selective attention. *Trends in Neurosciences*, xx(xx), 1–3. <https://doi.org/10.1016/j.tins.2021.04.010>
- Villena-González, M., López, V., & Rodríguez, E. (2016). Orienting attention to visual or verbal/auditory imagery differentially impairs the processing of visual stimuli. *NeuroImage*, 132, 71–78. <https://doi.org/10.1016/j.neuroimage.2016.02.013>
- Vroomen, J., & De Gelder, B. (2000). Sound enhances visual perception: Cross-modal effects of auditory organization on vision. *Journal of Experimental Psychology: Human Perception and Performance*, 26(5), 1583–1590. <https://doi.org/10.1037/0096-1523.26.5.1583>
- Welch, R. B. (1999). Meaning, attention, and the “unity assumption” in the intersensory bias of spatial and temporal perceptions. In G. Aschersleben, T. Bachmann, & J. Müsseler (Eds.), *Cognitive Contributions to the Perception of Spatial and Temporal Events*. Elsevier Science B.V.
- Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, 88(3), 638–667. <https://doi.org/10.1037/0033-2909.88.3.638>
- Wetzel, N., Buttelmann, D., Schieler, A., & Widmann, A. (2016). Infant and adult pupil dilation in response to unexpected sounds. *Developmental Psychobiology*, 58(3), 382–392. <https://doi.org/10.1002/dev.21377>

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.  
<https://ggplot2.tidyverse.org>

Wickham, H., François, R., Henry, L., & Müller, K. (2021). *dplyr: A Grammar of Data Manipulation* (1.0.7). <https://github.com/tidyverse/dplyr>

Wilke, C. O., Wickham, H., & Wilke, M. C. O. (2019). Package 'cowplot.' In *Streamlined Plot Theme and Plot Annotations for 'ggplot2'*.

## Chapter 5

# Spatial Working Memory in the Presence or Absence of Visual Rhythm

## 5.1 Introduction

Rhythmic temporal structures abound. They arise from within our environments, and within our bodies, creating ebbs and flows in the occurrence of events over time (Jones, 1976; Winfree, 1967; Zalta et al., 2020). Sequences of information can be described as streams of events with temporal structure. While rhythmically structured streams have been utilized in the study of a range of cognitive processes, whether and how rhythmic structure impacts *working memory* for streams of information remains an open question.

The simplest form of rhythm – an isochronous beat – can arise spontaneously, as with the pattern of our footsteps, or by design, as with the ticking of a clock. A repeated event that itself has a fixed duration, separated by repeated intervals of equal length, is what defines isochronous rhythmic structure (Nobre & van Ede, 2018). Perceptually, a series of discrete events may also be characterized as a rhythm if their sequence

generates an emergent recurring temporal pattern or structure (Nobre & van Ede, 2018; Parncutt, 1994; Ravignani et al., 2018). Beyond merely being defined as such, sequences with isochronous temporal structure are indeed perceived as being “rhythmic” (Breska & Deouell, 2017). On the other hand, repeated events whose instances are separated by intervals with lengths that vary, such that their sequence fails to form a perceptually discernible pattern, have temporal structure described as *arrhythmic*.

Within the brain, rhythms are ubiquitous. Rhythmic oscillatory activity facilitates communication within and between neural populations (Fries, 2015). Oscillations occurring at lower frequencies in particular organize the output of local computations across regions (Miller et al., 2012). Within the so-called “attention network” of the primate brain, rhythmic activity in the *theta-beta* range of frequencies (~4-20 Hz) generated from the thalamus and prefrontal and parietal cortices, rhythmically modulates the cycling between attentional shifting and sampling (Fiebelkorn & Kastner, 2019). This utilization of the term “rhythm” in reference to oscillations in the brain, is a characterization of continuous, periodic ebbs and flows. The similar structures over time between continuous events that change in their intensity or magnitude, and discrete events with clear onsets and offsets, can each be described as rhythmic.

Representation of a temporal structure such as rhythm, may require constituent representations of time, formed across brain regions. It has been argued that requisite representations of time are those of *interval* and *phase* (Gallistel, 2011). While an interval

is a representation of a temporal span between two events, phase accounts for where a given moment within that span resides (Gallistel, 2011; van Wassenhove, 2016). Representation of phase can be derived from two separate computations: “how long it will be, and how long it has been” (Gallistel, 2011). The latter requires a judgment regarding the amount of time lapsed from a starting point to the present — a representation of interval that has already occurred. The former, determining an amount of time from the present to a future point in time, requires a *prediction* about an interval to come. Representing temporal intervals enables a wide range of human functions from those reliant on perceptual judgments to the execution of control over movement (Buhusi & Meck, 2005; Ivry & Hazeltine, 1995). In the brain, both the basal ganglia and the cerebellum have been found necessary for interval representation (Buhusi & Meck, 2005; Ivry, 1996; Ivry & Keele, 1989). Performance of tasks requiring representation of a series of discrete instances which jointly form an “event structure,” is more variable in individuals with cerebellar lesions (Ivry & Spencer, 2004), indicating the cerebellum is important for processing stimuli streams with rhythmic or arrhythmic temporal structures. Recent evidence narrows this scope, however, pointing towards the cerebellum as critical for representation of single intervals; the striatum (within the basal ganglia) for representation of rhythms (Breska & Ivry, 2018).

Rhythm has been long studied in the auditory sensory domain. Perception and comprehension of rhythms that range from simple to complex, for example, have been long studied in the field of music cognition (Haegens & Zion Golumbic, 2018; Palmer &

Krumhansl, 1990; Parncutt, 1994). Evidence for exogenous neural entrainment by auditory rhythms in music has been found (Doelling et al., 2019; Doelling & Poeppel, 2015). It has also been found that individuals can exploit the regularity of temporal rhythm to help them identify targets within a stream (see Experiment 1 in Barnes & Johnston, 2010). Rhythmic perception is not only function within audition, however, but also vision.

Isochronous rhythmic structuring of visual stimuli enables inference and prediction about intervals yet to come (reviewed in Nobre & van Ede, 2018; Shalev et al., 2019). In Breska and Deouell (2014), for example, sequences of simple visual stimuli were presented in isochronous rhythm followed by a target that participants make a speeded response to. The interval between the final sequence stimulus and the target was manipulated such that for some trials its length could be validly predicted by the temporal structure generated by the sequence, while for other trials its length was not predictable. A large effect of rhythmic validity was found: participants were faster to respond to targets occurring after an interval that could be anticipated given the preceding sequence compared to when that interval was shorter or longer than expected. Further, rhythmicity had a greater effect on both behavior and evoked neural activity (measured by EEG) compared to when a non-temporal feature, in this case stimulus color, was used to guide behavior (Breska & Deouell, 2014). In another study with similar contrasts between temporal and color cues it was found that rhythm in particular induces both exogenous,

stimulus-driven, and endogenous, control-driven effects on attention, and that these separately impact behavior (Rohenkohl et al., 2011).

Valid predictions regarding the duration of a single interval can facilitate working memory performance. In a study by van Ede *et al.*, participants were presented with two lateralized oriented colored bars during an encoding period. The color of each bar was associated with either a short or long delay interval before being probed to reproduce the remembered stimulus orientation. Participants were more accurate in their reproduction when probed after an expected, versus unexpected, delay interval. Analysis of oscillatory activity (measured with EEG) during the task revealed significant modulation of posterior alpha power contralateral to the anticipated target dependent upon the expectation of being probed (van Ede et al., 2017). This study demonstrated that in working memory, expectation of a temporal interval can be utilized to prioritize the selection of some encoded items over others for reproduction at a specific time. As increases in alpha power contralateral to an attended target have been interpreted as a mechanism of suppressing irrelevant sensory inputs (Okazaki et al., 2014), their findings may point towards preparatory suppression to prevent distraction. Indeed, it has recently been put forward that distractors which occur at temporally predictable intervals can be more effectively suppressed than those that are unpredictable, supporting maintenance of items with fidelity in working memory (Gresch et al., 2021). Taken together, there is indication that a single, predictive temporal interval can facilitate working memory performance.

Rhythm, however, is a structure that emerges from *sequences* of intervals. In van Ede et al (2017), described previously, it was demonstrated that temporal expectation can facilitate working memory retroactively. This was accomplished by using expectation to guide internal attention and prioritizing processing of some encoded representations over others at a given time. In Rohenkohl et al., (2011), also described earlier, rhythm was shown to have independent, additive exogenous and endogenous effects on attention. Might rhythm — a predictable *series* of intervals — generate proactive benefits for working memory? If so, are effects exogenously driven, or do they vary based upon endogenous control?

It has been well established in the domains of attention and perception, that expectations generated by rhythmic structures can speed responses, enhance sensitivity to detect stimuli, and increase precision (reviewed in Haegens & Zion Golumbic, 2018; Nobre & van Ede, 2018). Because perception and attention are critical to working memory (Cowan et al., 2005; Oberauer, 2002, 2019; Ricker, 2015), and attention and working memory have been shown to fluctuate together (deBettencourt et al., 2019), benefits arising from rhythmic temporal structure may extend to working memory for streams of information. It may be the case that there is greater precision in encoding for stimuli presented with rhythmic compared to arrhythmic temporal structure, leading to greater accuracy in task performance. It may also be the case that rhythmically presented stimuli are encoded into working memory more rapidly than stimuli presented arrhythmically, leading to faster preparedness to respond.

## 5.2 Methods

To test the hypothesis that rhythmic temporal structure facilitates working memory for streams of information, we designed a study contrasting visuospatial working memory performance given stimulus streams with rhythmic or arrhythmic temporal structure. A novel task (described below and shown in **Figure 9**) was designed to test for effects of rhythmic temporal structure upon working memory processing for visuospatial information. To shed light on the specific processing that may be associated with rhythm, and to examine the extent of the exogenous impact rhythm might have upon working memory, we conducted two separate experiments. The first experiment contrasted task performance for the short-term memory of target locations presented either rhythmically or arrhythmically. In the second experiment, in addition to the requirement for memory of location, memory for target order in rhythmically or arrhythmically presented sequences was imposed.

### 5.2.1 Participants

Participants were recruited from the Johns Hopkins University student population and compensated via course credit. Musical experience of participants is noted for each experiment; it did not significantly differ between experiments. Gender labels are based upon participant self-report of birth sex and gender identity, in accordance with recent recommendations (Rubin et al., 2020).

Thirty-one individuals were recruited for Experiment 1. Data from two individuals were excluded from analysis due to atypical perception. Data from one individual was not available for analysis due to technical failure during the experiment. Analyzed data was collected from 18 cis-females and 10 cis-males, aged 19.7 (SD=1.1), and 19.8 (SD=1.2) years, respectively. Approximately 64% of participants reported having musical experience or training, with an average of 8.3 (SD=4.5) years of experience.

Twenty-nine individuals participated in Experiment 2 with two excluded due to technical failure during the experiment. Data available for analysis was collected from 17 cis-females and 10 cis-males aged 19.5 (SD=0.9), and 20.1 (SD=2.0) years, respectively. Approximately 70% of participants were musicians, with an average of 8.3 (SD=2.6) years of experience.

### 5.2.2 Stimuli and Task

#### *Overview*

The visuospatial working memory paradigm designed for this study, utilized in both Experiment 1 and 2, is depicted in **Figure 9**. The task required working memory for the spatial locations of targets configured in an arrangement that discouraged shape or eye movement recoding strategies. Targets were presented in a sequence, with either Rhythmic or Arrhythmic temporal structure determined by Condition.

The total number of targets presented was variable across trials. The task, no matter the number of targets displayed, was always to remember and reproduce the

locations of the *final three* targets presented – the *terminal target set*. A set size of three was held fixed as this falls within an expected working memory capacity range among neurotypical young adults (Cowan, 2010). The positioning of the target set at the end of the sequence provided time for temporal structure to be perceived. Half of the trials had an isochronous rhythmic structure and half had an arhythmic structure.

Experiments 1 and 2 differed in the specific task instruction given to participants. In Experiment 1, Participants were instructed that the final three target locations could be reproduced in *any* order. In Experiment 2, Participants were instructed to reproduce the final three target locations in the *same* order as presented. In all other aspects both Experiments were identical, as described below.

### *Visual Stimuli*

During the experiment, a geometric framework was presented visually on screen for the duration of each trial. Targets — colored shapes — appeared one at a time within the bounds of this framework, forming a sequence. Visual targets each occupied a distinct location within the boundaries of the framework. No framework edges or regions were shared between targets. As the sequence of targets was displayed, only a single target was presented at any given time.

The geometric framework was arranged circumferentially around a central fixation. The fixation cross occupied approximately 1° of visual angle in diameter, and the geometric framework's inner- and outermost extents fell between 2° and 10° of visual

angle in diameter. This framework was displayed for the entirety of each trial, excepting when feedback was given. Participants were instructed to hold their gaze at the central fixation cross throughout stimulus presentation. Viewing distance was fixed at approximately 59cm from the screen using a fixed-position chinrest.

### *Cue Sequence*

A 500ms fixation period occurred at the start of each trial, followed by a *cue sequence*. A visual cue: a colored circle stimulus occupying 1.2° of visual angle and presented centrally overlaying the fixation cross, and an auditory cue: a 320Hz tone with a 50ms rise and 100ms fall, were presented together, three times. Each cue presentation lasted 250ms, followed by an interval of a condition-specific length. The Auditory stimulus was created with Audacity® sound editing software (Audacity Team, 1999) and delivered through Sennheiser HD 201 over-ear binaural headphones, with presentation volume fixed at a level well above threshold. The combined visual + auditory cue was presented a total of three times, comprising the *cue sequence* within each trial. This sequence provided a cue regarding the impending start of the sample sequence, but also provided a valid indicator of its temporal structure. Both auditory and visual cues were presented to increase the salience of this structure.

### *Sample Sequence*

Immediately following the *cue sequence*, a visual *sample sequence* of targets was displayed without any accompanying sound. Within 86% of the trials, the total number of

targets displayed within the sample sequence was 6, 7, or 8, with equal probability. Within the remainder of trials, labeled as “catch trials,” a total of 4 targets were displayed. Each target in the sequence was presented onscreen for a duration of 250ms. The final three targets, the *terminal target set* of every sequence, were the targets to be maintained in working memory. The length of the sample sequence was pseudorandomized within experimental blocks. Participants were not provided with advance notice of the total length of any given sequence. Any target set being maintained in working memory would require updating until the end of the sample sequence was reached. Catch trials were included to promote attending to and encoding of targets throughout the duration of the sample sequence, and to discourage the adoption of a strategy to ignore early targets. Catch trials, however, were excluded from analysis as their short overall duration may have limited the potential for the temporal structure of the sequence to have any effect. No overt cue regarding the end of the sample sequence was provided beyond the passage of time: after the final target in the sequence was presented there was a 2,500ms delay.

### *Response*

After the delay period, visual cues marked the beginning of the response period. The most prominent among these cues was the disappearance of the central fixation cross. In addition, all possible target locations within the framework took on a subtle glow to indicate they were active and available for selection. Finally, the mouse cursor, invisible during all other times of the trial, became visible. Participants had a total of 5s to input their response and select the locations of each of the three remembered targets. Small

changes in the brightness of target locations indicated that the response had been registered. At the end of the response window, a feedback message was displayed for 2s to confirm a complete response of three locations had been provided or to indicate that time had run out.

### 5.2.3 Design

Half of trials in the experiment were defined by their Rhythmic temporal structure, the other half were Arrhythmic. All trials with Rhythmic temporal structure adhered to the following parameters of the Rhythmic condition (RHY). In RHY trials, all intervals, beginning with the *cue sequence* and continuing through the *sample sequence* presentation, had a fixed length of 750ms. In all conditions, target stimuli were presented for a duration of 250ms. Hence, in the RHY condition, stimulus onset asynchrony (SOA) was fixed at 1s. Two Arrhythmic conditions were defined, within which all intervals fell within the range of 250ms-1,250ms, varying in steps of 50ms. Intervals in the Arrhythmic conditions were predefined such that, altogether, their sequential presentation did not form a discernible temporal pattern. The order of the final two intervals – those occurring *between* the three memory targets of the terminal target set – distinguished the two Arrhythmic conditions. In the Arrhythmic Short-Long (ASL) condition, the terminal intervals were 250ms and 1,250ms. In the Arrhythmic Long-Short (ALS) condition, the same interval values were utilized with their order reversed: 1,250ms then 250ms. In Arrhythmic conditions, the minimum SOA was 500ms, the maximum 1,500ms. Across conditions, then, while the total duration of the sample sequence varied, the sum

duration of intervals between target stimuli of the terminal target sequence was held constant at 1,500ms (RHY: 750ms, 750ms, ALS: 1,250ms, 250ms, ASL: 250ms, 1,250ms). The length of the Delay period immediately following the terminal sequence presentation (see timeline in **Figure 9B**) was also held constant across conditions at 2,500ms.

Eighteen terminal target sets were predefined. Within half of the sets, all targets were spatially “distant” from one another – separated by one or more spaces belonging to other targets. The other half of the terminal target sets were labeled spatially “adjacent.” In these sets two, but not all three, of the targets were nearest radial neighbors. Examples of these configurations are shown in **Figure 10F**. It has been previously demonstrated that spatial “chunking” or grouping strategies based upon spatial proximity are often employed in spatial working memory tasks (Bor et al., 2003; Gmeindl et al., 2011). Further, individuals readily form predictions regarding upcoming target locations when spatial patterns emerge within a sequence (Amalric et al., 2017; Wang et al., 2019). Predefining target sets in this manner was performed with the aim of limiting the possible contribution of space-based prediction, while allowing for comparison between spatial configurations based upon proximity. Target sets were distributed across three broader groups to ensure that within a given experimental block, the same target sets were not presented in both Rhythmic and Arrhythmic conditions or repeated within conditions in successive blocks. The order of specific target presentation within each set was randomized.

#### 5.2.4 Procedure

After obtaining consent, the experiment was conducted with participants seated in front of a computer screen in a private booth, positioned using a chinrest, with headphones on. Participants received on-screen instructions and training, followed by the primary task. Participants completed six blocks of 14 trial each. Trial order was pseudorandomized within each block for each participant. Feedback on block-level task performance (proportion of correct target location responses) was provided after the completion of each block. Participants were given the opportunity for self-timed breaks between blocks. On completion of the experiment, Participants completed questionnaires and were then debriefed by the Experimenter. The entire study session lasted approximately one hour.

#### 5.2.5 Analysis

Different outcome variables of interest related to task performance were defined. *Decision Time* (DT) was operationalized as the time lapse between the start of the response window to the Participant's first target selection (e.g. van Ede et al., 2017). To facilitate across-participant comparisons in Experiments 1 and 2, DT was normalized ( $DT_{norm}$ ) for each Participant per the methods of (Golob & Mock, 2019) as follows. Across all trials for each participant, their maximum ( $DT_{max}$ ) and minimum ( $DT_{min}$ ) was identified. For each trial, the following formula was applied:  $DT_{norm} = (DT_{trial} - DT_{min}) / (DT_{max} - DT_{min})$ . DT normalized per this method results in a value ranging from 0 to 1.0. The value 0 represents the fastest DT for that participant, the value 1.0 the slowest.

Target reproduction was analyzed in two ways. Reproduction of target location was deemed correct for a single trial if the report included the correct locations of all three terminal targets presented, independent of order. Reproduction of target order was judged based upon whether targets were reported in the same, or different, serial order compared to sample sequence presentation.

Parametric group-level comparisons of performance across conditions both within, and between, Experiments 1 and 2 were performed using omnibus ANOVA tests. Normalized Decision Time and Response Proportions (Location and Order) were the outcome variables of interest. To examine within-subjects effects separately for Experiments 1 and 2, 1-way ANOVA tests were performed. To examine between-subjects effects across Experiments, 2-way ANOVA tests were performed. ANOVA tests were followed, when relevant, with paired *t*-tests to examine underlying effects. Multiple comparisons correction was performed via the False Discovery Rate method (Benjamini & Hochberg, 1995).

Statistical analyses were carried out in R (R Core Team, 2018) with the *dplyr* (Wickham et al., 2021) and *ez* (Lawrence, 2016) packages. Figures were produced in R with the *ggplot2* (Wickham, 2016), *viridis* (Garnier et al., 2021), *hrbrthemes* (Rudis, 2018), and *cowplot* (Wilke et al., 2019) packages.

## 5.3 Results

The results of within-experiment analyses are presented first, followed by those of analyses performed across Experiments. Catch trials were removed from the data prior to analysis. Statistics reported below reflect task performance averaged over trials with sample sequences in which 6, 7, or 8 targets were presented.

### 5.3.1 Experiment 1

Participants were well able to perform the task. Across conditions, the mean proportion of trials in which Participants correctly identified all three terminal targets exceeded 80%. With proportion of correct-identity report as the outcome variable, no within-participant differences were present across Rhythmic and Arrhythmic conditions:  $F(2,54) = 0.16$ ,  $p > 0.5$ . The proportion of correct-location trials in which targets were reported in the same order ranged between 45-70%. With proportion of correct-order report as the outcome variable, a small but insignificant difference was present across conditions:  $F(2,54) = 1.96$ ,  $p > 0.1$ .

To determine if memory target configuration had an effect on performance, separate analyses were performed based upon the presence of spatial grouping (adjacent or distant) within target sets. Results are shown in **Figure 10**. A main effect of spatial grouping on the report of target order was present:  $F(1,27) = 5.97$ ,  $p < 0.03$ ,  $\eta^2_G < 0.02$ , but did not interact with Condition:  $F(2,54) = 0.49$ ,  $p > 0.5$ . In RHY and ALS conditions in

particular, the proportion of targets reported in the same order as presented was lower when targets were spatially adjacent versus distant (see **Table 4**).

Within spatially adjacent target sets, two (but not three) targets neighbored one another. The order of individual target presentation was randomized across trials, therefore, in some trials adjacent targets were presented sequentially as the first and second targets (early sequential grouping), in others they were the second and third targets (late sequential grouping). Response patterns for trials in which spatially adjacent targets were presented in sequence early, compared to late, are shown in **Figure 10E**. Targets that were adjacent in space and time more frequently elicited responses that reflected order-based grouping. This was the case across conditions.

When analysis was performed with normalized Decision Time as the outcome variable, a main effect of condition was found:  $F(2,54) = 5.50$ ,  $p < 0.01$ ,  $\eta_G^2 = 0.05$ . This analysis was repeated with unscaled Decision Time to the same effect. Direct contrasts indicated that Decision Time was significantly faster in the Rhythmic condition than in both Arrhythmic conditions. RHY vs. ALS:  $t = 2.36$ ,  $p < 0.03$ , RHY vs ASL:  $t = 3.21$ ,  $p < 0.01$ . These results are shown in **Figure 11**.

### 5.3.2 Experiment 2

Despite the difference in instruction from Experiment 1, participants in Experiment 2 were also well able to perform the task. As in Experiment 1, on average, participants correctly reproduced target identities in more than 80% of trials across

conditions, with no significant effect of condition:  $F(2,52) = 0.34, p > 0.5$ . Experiment 2 instructed reproduction of targets in the same order as presented, and indeed participants were able to do so in more than 80% of trials, with no significant effect of condition:  $F(2,52) = 0.38, p > 0.5$ . With the increased reproduction in correct order, the effect of spatial grouping that was present in Experiment 1 disappeared in Experiment 2:  $F(1,26) = 1.85, p > 0.1$ . Interestingly, the effect of Condition upon normalized Decision Time that was present in Experiment 1, also disappeared in Experiment 2:  $F(2,52) = 1.31, p > 0.2$ . Due to the high proportion of same-order reproduction trials, too few trials remained for conclusive analysis of spatial and sequential grouping sample patterns on reproduction patterns. Detailed task performance results for both Experiments 1 and 2 are shown in **Table 4**.

### 5.3.3 Between-Experiment Outcomes

Data from both Experiments 1 and 2 was combined for between-subjects analyses assessing the effect of Experiment. There was no main effect of Experiment upon reproduction of target location:  $F(1,53) = 1.89, p > 0.5$ . As expected due to the difference in task instruction, a large effect of Experiment was present in reproduction of order:  $F(1,53) = 36.4, p < 0.0001, \eta_G^2 = 0.37$ . No main effect of Experiment was present on normalized Decision Time:  $F(1,53) = 0.61, p > 0.4$ . A trending interaction between Condition and Experiment,  $F(2,106) = 2.41, p < 0.1$ , was present due to the comparatively slower Decision Times in Arrhythmic conditions within Experiment 1. Direct comparisons

of Decision Time in Arrhythmic conditions between Experiments 1 and 2, however, were not significant:  $t < 1.5$ ,  $p > 0.1$  for all contrasts.

## 5.4 Discussion

In an experimental context, it is commonplace for stimuli to be presented so that they are consistently separated – in time or in space. Arranging visual stimuli in spatial configuration with equidistant concentric spacing, for example (e.g. paradigm in deBettencourt et al., 2019). In the well-known *Corsi Blocks* (Berch et al., 1998; Corsi, 1972) visuospatial working memory paradigm, while targets are randomly positioned in space, during presentation they are identified at a rate of one per second – in an isochronous rhythm. One motivation in defining these presentation parameters, is the desire to minimize potentially confounding effects or limit the amount of “noise” in a dependent variable. In applying this approach to temporal presentation parameters, however, *rhythm*, is introduced, even if unintentionally. In the present study, we hypothesized that rhythmic temporal structure within a stimulus stream is a factor which impacts visuospatial working memory performance. So, we tested this explicitly.

Rhythm, particularly as a parameter of music, is well studied in the auditory domain but less so in the visual domain. In the auditory domain, the presence of rhythm is relatively quickly perceived and extracted from the structure of a sequence (Jacoby & McDermott, 2017). Indeed people are biased towards reorganizing auditory sequences

so that they conform to a fundamentally rhythmic temporal structure (Ravignani et al., 2018). Though there is a bias for specifically temporal processing in the auditory domain (Michalka et al., 2015), as reviewed earlier rhythm can nevertheless impact the processing of visual stimuli.

By focusing upon working memory for visual information residing solely in the spatial domain, we aimed to isolate temporal effects specifically arising from the processing of rhythmic temporal structure. The ability to perceive *complex* rhythmic structure is thought to be fundamental to our capacity for processing language, but this capacity is not decoupled from the processing of other acoustic features and cues (Haegens & Zion Golumbic, 2018). In the present study we avoided specific inclusion of language – using spatial rather than verbal memory targets and a complex visual framework designed to minimize the usefulness of verbal strategies such as naming specific targets or associating their concentric arrangement with numbers on a clockface.

A concern in contrasting performance following Rhythmically versus Arrhythmically presented stimulus streams, is the possibility that temporal effects specific to Rhythm might be introduced. Serial dependence in vision, for example, causes discrete samples of past and present instances to be biased towards one another, in facilitation of perception as a continual experience (Fischer & Whitney, 2014). Serial dependence has been shown to cause errors of “intrusion” in short-term memory, attributed to *proactive interference*, by which one encoded stimulus interferes with another (Kiyonaga et al.,

2017; Makovski & Jiang, 2008). Serial dependence has been demonstrated to greatest effect in the biasing of encoding fundamental stimulus properties such as color or orientation (Fischer & Whitney, 2014). In both Experiments 1 and 2, however, across conditions, similar performance was present in the report of target location. As each target was both distinct in shape and discrete in location, it may be less likely that intrusion effects or errors of degree resulting from this type of blending in serial perception were possible.

We found that across experiments and temporal structure conditions, participants were well, and similarly, able to encode and reproduce the locations of the three terminal targets. This was the case across sequence lengths, and target arrangements, as shown in **Table 4**. The critical effect of temporal structure in these experiments was upon the *temporal* measure of behavioral performance assessed: Decision Time. Across both experiments, Decision Time was fastest in the case of Rhythmic compared to Arrhythmic temporal structure. However, with the additional requirement for memory of target order imposed (Experiment 2), Decision Times in Arrhythmic conditions sped up when compared to when only identity report was required (Experiment 1).

Despite the seeming increase in task demand of Experiment 2, and the faster Decision Times in Arrhythmic conditions, no speed-accuracy trade-off was present: proportions of trials in which target identities were correctly reported were similar across conditions in both Experiments. Further, when participants were free from task

constraints regarding target order in Experiment 1, they spontaneously reported targets in sequential order on nearly half of trials – far above chance (~17%). Among remaining trials with two of the targets spatially adjacent, response patterns indicate that organization of encoded items in working memory was also strongly driven by spatial parameters (**Figure 10E**). Utilization of spatial recoding strategies in working memory for visual streams of information has been demonstrated in the past work of our group (Gmeindl et al., 2011) as well as that of others (Bor et al., 2003). Further, it has been shown elsewhere that temporal and spatial parameters can interact, to additive perceptual benefit in a spatial task (Rohenkohl et al., 2014), and that “temporal segmentation” can alter conceptual grouping (Jones, 1976). In the present study, patterns of response indicate that spontaneous recoding occurs based upon both spatial (proximity) and temporal (order) factors. However, the lack of difference in response patterns between Arrhythmic conditions in particular, in which the short and long interval lengths might have encouraged opposite patterns of temporal segmentation, suggests temporal effects on recoding are limited to order.

The differences in Decision Time based upon different temporal structures can be interpreted as either Rhythmic *facilitation*, or Arrhythmic *trapping*. It has previously been argued that people don’t begin reproduction of a list of items until all items are in an activated state in memory (Anderson et al., 1998). One possibility is that the perception of a rhythmically structured stream speeds up processing of it. Prior work in the reproduction of rhythmic auditory sequences has found that people when people hear

simple rhythms, they reproduce them at a faster rate than presented (Parncutt, 1994). Because rhythms generate strong temporal expectation (Nobre & van Ede, 2018) and certainty regarding the timing of an upcoming event (Grabenhorst et al., 2021), the path may be cleared, so to speak, for neural resources to be more rapidly dedicated to encoding items into working memory.

On the other hand, a stream with arrhythmic temporal structure may have an effect of “trapping” by temporarily halting further processing. Whereas rhythm promotes certainty, arrhythmicity may increase uncertainty - promoting instead, a reluctance to “commit” to encoding until uncertainty is reduced. In Experiment 1 participants exhibited a propensity to reproduce targets in serial order, despite the lack of explicit requirement to do so, but they employed spatial recoding strategies a large proportion of the time as well. In Experiment 2, the specific serial order requirement may have increased the overall demand of the task, requiring a greater degree of focused attention. In past studies, higher compared with lower attentional loads have been shown to reduce the impact of distractors (Konstantinou et al., 2014; Lavie & De Fockert, 2005). Further, control processes play a strong role in input gating – restricting what items become encoded in working memory (Lorenc et al., 2021). Though *uncertainty* is not necessarily an object encoded from the sensorium, it is possible that higher demands upon a limited pool of attentional resources limit the amount of internal attention that can be allocated to focusing upon that uncertainty.

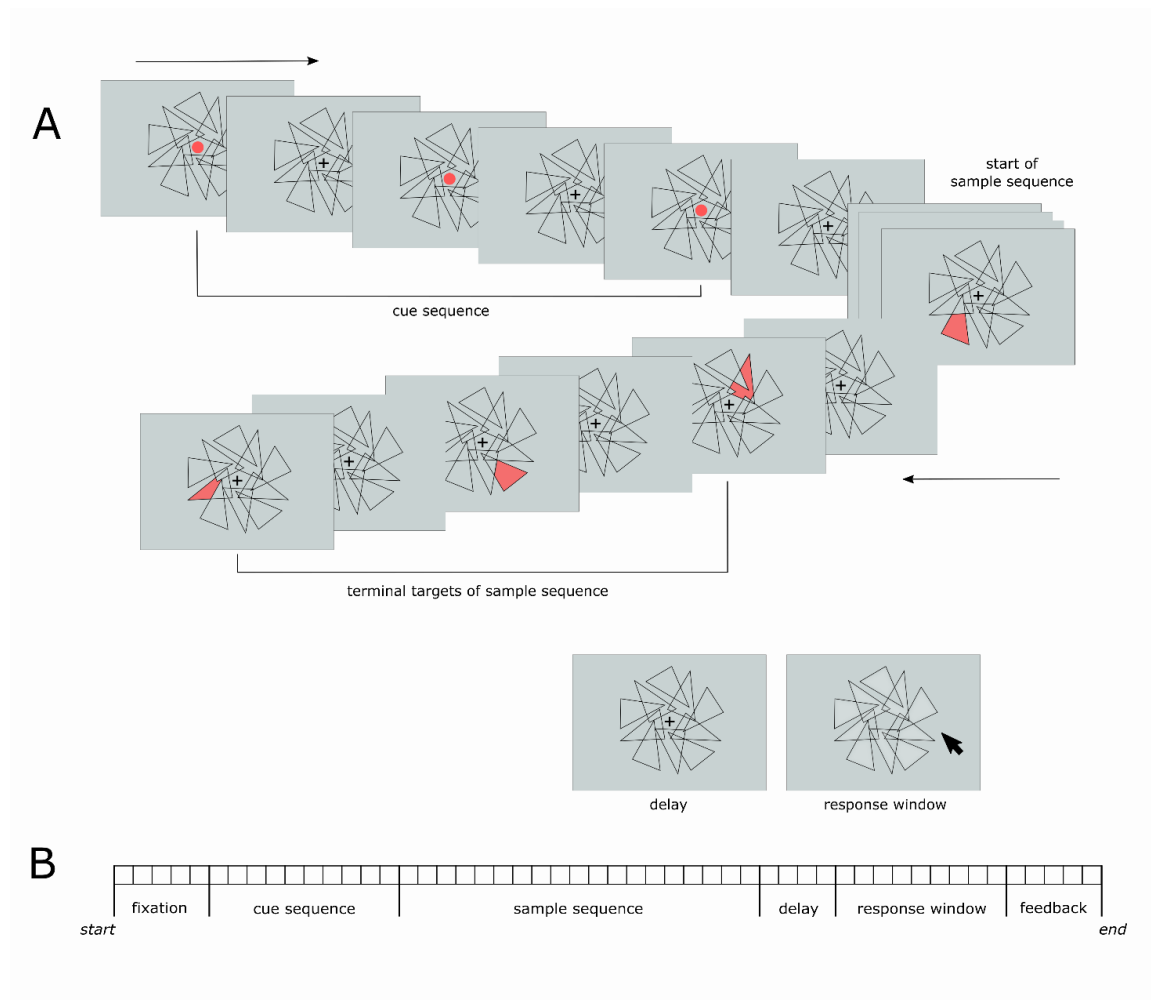
Though the present study is agnostic in ascribing to any particular model of temporal cognition, it has been pointed out that the cognitive and neural underpinnings of interval processing may differ from those underlying the processing of other temporal structures (Ivry & Richardson, 2002). In the present results, though the difference between Decision Time in ALS versus ASL conditions is not significant, there is a consistent pattern across Experiments 1 and 2 that may point towards an effect of the final interval duration: Decision Time was slowest when the final interval was longest. It has been argued that the perception of time is subject to “autoscaling,” whereby the same computations are carried out in the representation of different periods of time by scalar adjustments in reference frame (Gallistel, 2011; Gibbon et al., 1984). Early work on temporal perception in animals, describes “scalar-timing” as a property of time perception which enables prediction of when an impending, meaningful event will occur (Gibbon, 1977). Accordingly, the longer compared with shorter final interval in the ASL condition, may have generated a relatively longer period of uncertainty, reflected in the slower Decision Times specific to this condition.

## 5.5 Conclusion

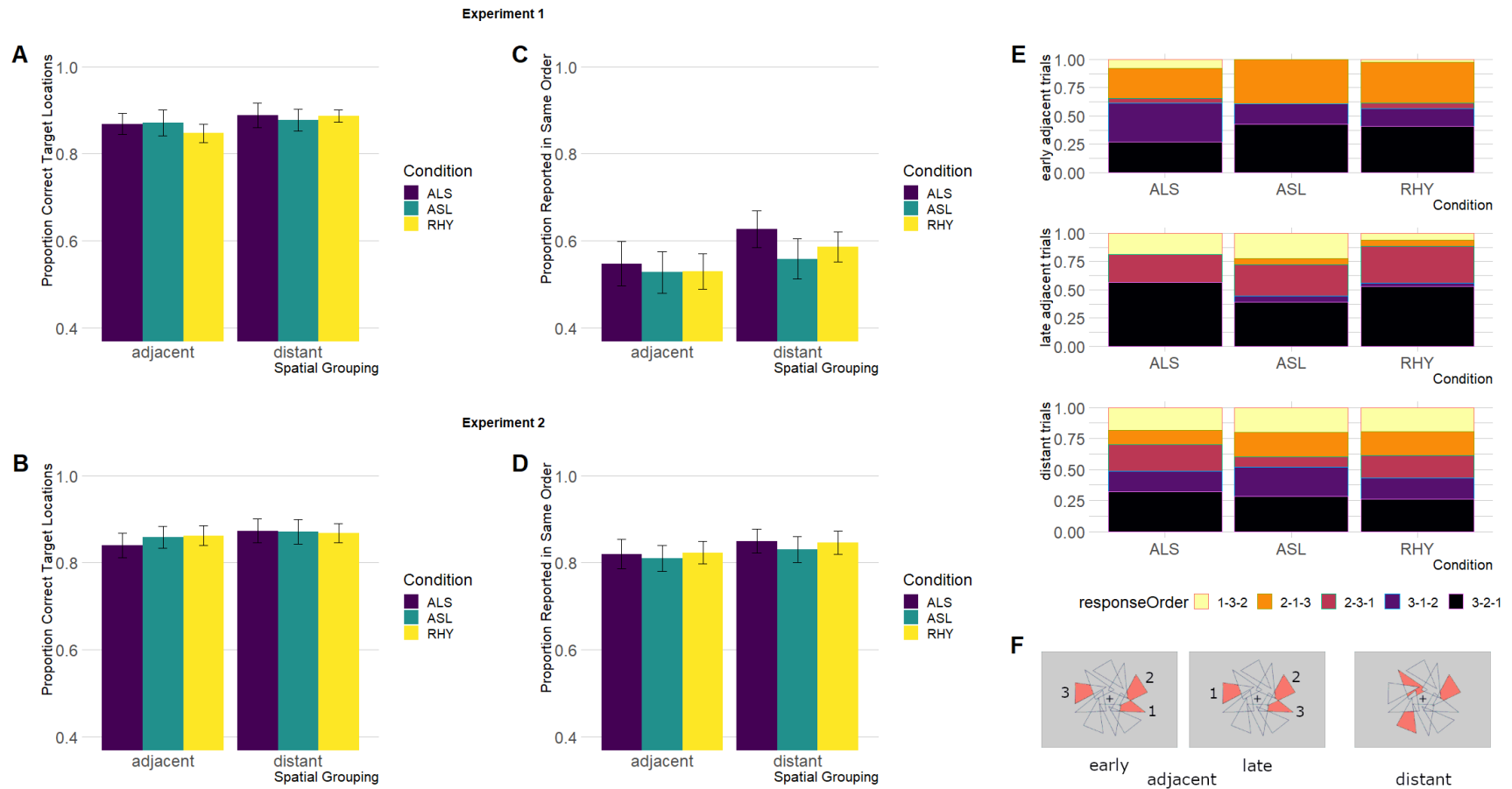
Temporal structures that may impact cognitive processes can be classified in a number of ways (Nobre & van Ede, 2018). While previous research in the field of visual working memory has examined the effects of some types of temporal structure: those of

relative, discrete intervals for example, the impact of rhythmic temporal structure on visual working memory has remained an open question. In the present study, we contrasted working memory for visually presented spatial information presented such that it had rhythmic, or arrhythmic, temporal structure. While individuals were successful in performing the task regardless of temporal structure, they were consistently faster when presented with Rhythmically versus Arrhythmically structured information. Arrhythmic temporal structure may impair consolidation of order and identity information into working memory, but that impairment may be alleviated when attentional demand is increased. Future studies might test the effect of increasing task demands to explore the bounds of beneficial impacts on performance that rhythmic structures yield.

## 5.6 Figures



**Figure 9:** Task and general Timeline of a single trial. **A)** Sequence of events in a trial. At the start of each trial (upper left) an auditory and visual cue sequence provided valid context regarding the temporal structure within the trial. The cue sequence was immediately followed by presentation of the sample sequence (upper right). The sample sequence varied in length. The final three targets of the sample sequence comprised the *terminal target set* – whose locations were to be remembered and reproduced during the response window. All cue and target stimuli were presented for a duration of 250ms. During all intervals between cue or target stimulus presentation, the geometric framework and fixation were displayed for a duration defined by Rhythmic (750ms) or Arrhythmic (range between 250ms-1,250ms) condition parameters. The final two intervals occurring within the terminal target set summed to 1,500ms across all conditions. Following the 2,500ms delay period, participants had 5s to reproduce the three terminal target locations. **B)** Relative timeline of events in each trial.

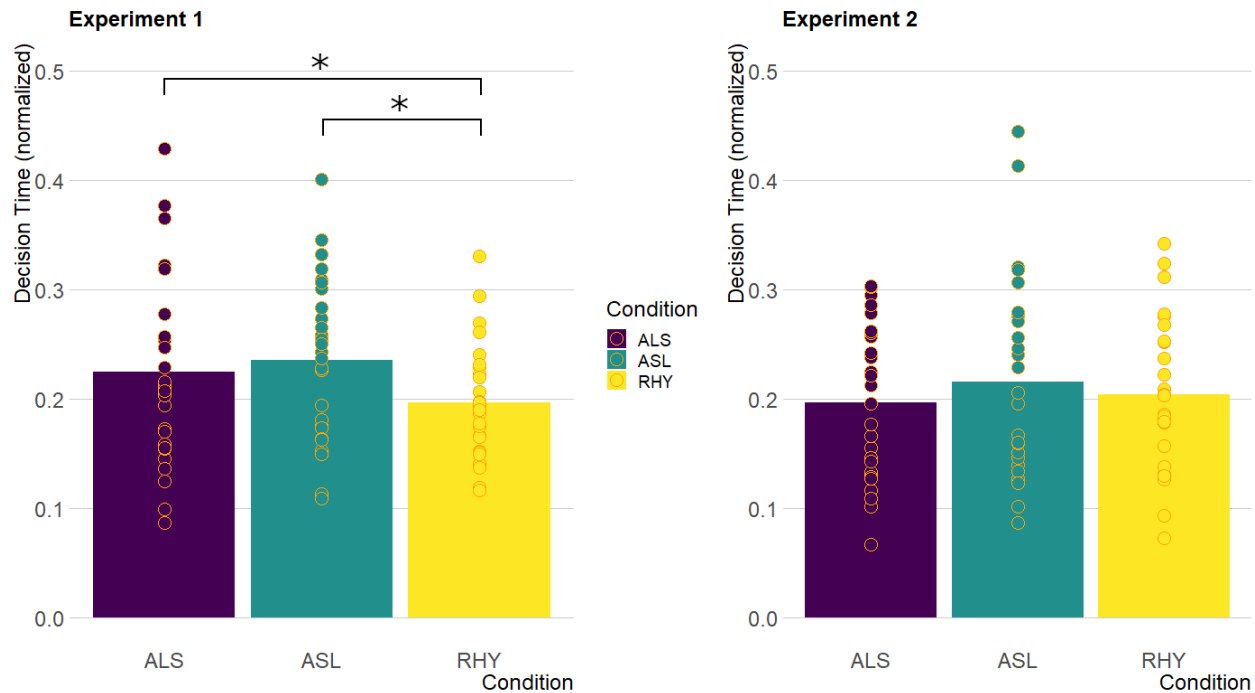


**Figure 10:** Behavioral Performance by terminal target set Grouping across Rhythmic (RHY), and Arrhythmic (ALS, ASL) Conditions for Experiments 1 and 2. **A, B:** Proportion of trials in which the correct Location of all three terminal targets was reported. **C, D:** Proportion of correct Location trials in which all three terminal target locations were reproduced in the same Order as presented. Chance order of report = 0.17. Note that in Experiment 1, a main effect of spatial grouping is present,  $F(1,27)=5.97$ ,  $p < 0.025$  (**C**). Participants reproduced target locations in the same order as presented more so when targets were spatially distant compared to when adjacent.

**E:** Response order patterns in Experiment 1 for correct target location trials not reproduced in the same order as presented, separated by the presence of

spatial and order grouping among targets. Response Order numbers indicate the actual serial position of a particular target as presented (see examples in **F**), compared to the serial position of targets selected by the Participant during response. The upper and middle bar plots are specific to spatially adjacent target sets which were also adjacent in order, the lower plot depicts response patterns within trials with spatially distant target sets. When spatially adjacent terminal targets were presented in adjacent order early in the sequence (Top), response patterns in which the first and second targets were grouped together (**2-1-3**, **3-2-1**, **3-1-2**) predominated. When spatially adjacent terminal targets were in adjacent order late in the sequence (Middle), response patterns in which the second and third targets were grouped together (**2-3-1**, **3-2-1**, **1-3-2**) predominated. These can be compared with response order among spatially distant target sets (bottom), in which all possible response pattern is well represented.

**F:** Visual examples of spatially adjacent and distant terminal targets arrays. Distant target sets were always separated radially by one or more target areas. Within adjacent sets, two (but not all) of the three targets were radial nearest neighbors. Note that during sequence presentation within the experiment, only a single target was ever present at a time (as shown in **Figure 9**) – single arrays of the terminal target set are shown here for ease of visualization. Numbering within the adjacent target sets demonstrates two possible sequence orders of target presentation. In the *early* example, adjacent targets are presented in the first and second serial positions. In the *late* example, adjacent targets are presented in the second and third positions. Hence, grouping within working memory could occur based upon spatial proximity and/or presentation order, in addition to temporal proximity as determined by Condition.



**Figure 11:** Decision Time by Rhythmic (RHY), and Arrhythmic (ALS, ASL) Conditions for Experiments 1 and 2. A main effect of Condition,  $F(2,106) = 4.61$ ,  $p < 0.02$ , is present across experiments, with the largest differences resulting from Experiment 1. Decision Time is faster when working memory was employed for stimuli with rhythmic compared to arrhythmic temporal structure. In Experiment 2, when attending to and encoding target order was required, compared to Experiment 1, Decision Time was faster in Arrhythmic conditions, pointing towards the distracting effect of arrhythmic structure having been ameliorated when task demands included the report of order.

**Table 4:** Behavioral measures of central tendency across all analyzed factors for Experiments 1 and 2.

Condition	Terminal Intervals	Spatial Grouping	Sequence Length	Experiment 1				Experiment 2			
				Decision Time		Proportion Same Report <i>mean (sd)</i>		Decision Time		Proportion Same Report <i>mean (sd)</i>	
				<i>ms</i>	<i>norm.</i>	Order & ID	ID only	<i>ms</i>	<i>norm.</i>	Order & ID	ID only
<b>Rhythmic</b>	Equal	distant		1081.9 (95.98)	0.20 (0.07)	0.59 (0.18)	0.89 (0.07)	1222.2 (223.0)	0.20 (0.07)	0.85 (0.14)	0.87 (0.11)
		adjacent		1074.0 (111.0)	0.19 (0.05)	0.53 (0.21)	0.85 (0.11)	1259.3 (224.9)	0.21 (0.08)	0.82 (0.14)	0.86 (0.12)
			6	1095.0 (134.8)	0.20 (0.06)	0.59 (0.25)	0.86 (0.14)	1247.3 (253.3)	0.20 (0.08)	0.84 (0.16)	0.87 (0.16)
			7	1060.7 (101.1)	0.19 (0.08)	0.53 (0.19)	0.87 (0.09)	1241.7 (201.0)	0.21 (0.08)	0.83 (0.15)	0.87 (0.13)
			8	1078.4 (139.5)	0.20 (0.06)	0.55 (0.22)	0.87 (0.10)	1233.3 (236.4)	0.20 (0.08)	0.82 (0.17)	0.85 (0.15)
<b>Arrhythmic</b>	S-L	distant		1153.5 (177.9)	0.24 (0.10)	0.56 (0.25)	0.88 (0.13)	1311.6 (362.7)	0.23 (0.11)	0.83 (0.16)	0.87 (0.15)
		adjacent		1139.5 (193.4)	0.23 (0.09)	0.53 (0.25)	0.87 (0.16)	1248.6 (300.1)	0.20 (0.10)	0.81 (0.15)	0.86 (0.13)
			6	1121.6 (181.1)	0.23 (0.12)	0.57 (0.27)	0.88 (0.16)	1274.4 (340.1)	0.22 (0.12)	0.81 (0.14)	0.85 (0.13)
			7	1171.4 (196.9)	0.24 (0.09)	0.54 (0.28)	0.85 (0.19)	1282.8 (392.8)	0.21 (0.12)	0.86 (0.17)	0.88 (0.17)
			8	1146.3 (216.36)	0.24 (0.12)	0.52 (0.28)	0.88 (0.16)	1269.6 (304.6)	0.22 (0.10)	0.79 (0.17)	0.85 (0.15)
	L-S	distant		1105.1 (173.3)	0.22 (0.12)	0.63 (0.22)	0.89 (0.15)	1217.9 (225.8)	0.19 (0.08)	0.85 (0.14)	0.87 (0.14)
		adjacent		1145.5 (178.1)	0.23 (0.09)	0.55 (0.27)	0.87 (0.13)	1250.9 (300.0)	0.20 (0.08)	0.82 (0.17)	0.84 (0.15)
			6	1143.1 (236.9)	0.23 (0.13)	0.67 (0.29)	0.92 (0.12)	1198.7 (272.1)	0.18 (0.09)	0.85 (0.19)	0.87 (0.16)
			7	1086.4 (138.2)	0.20 (0.09)	0.60 (0.24)	0.89 (0.14)	1265.4 (265.7)	0.21 (0.08)	0.81 (0.19)	0.83 (0.18)
			8	1148.6 (170.8)	0.24 (0.09)	0.49 (0.25)	0.83 (0.19)	1238.0 (314.4)	0.20 (0.10)	0.83 (0.18)	0.85 (0.17)

## 5.7 References

- Amalric, M., Wang, L., Pica, P., Figueira, S., Sigman, M., & Dehaene, S. (2017). The language of geometry: Fast comprehension of geometrical primitives and rules in human adults and preschoolers. In *PLoS Computational Biology* (Vol. 13, Issue 1). <https://doi.org/10.1371/journal.pcbi.1005273>
- Anderson, J. R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An Integrated Theory of List Memory. *Journal of Memory and Language*, 38(4), 341–380. <https://doi.org/10.1006/jmla.1997.2553>
- Barnes, R., & Johnston, H. (2010). The role of timing deviations and target position uncertainty on temporal attending in a serial auditory pitch discrimination task. *Quarterly Journal of Experimental Psychology*, 63(2), 341–355. <https://doi.org/10.1080/17470210902925312>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57, 289–300. <https://doi.org/10.2307/2346101>
- Berch, D. B., Krikorian, R., & Huha, E. M. (1998). The Corsi block-tapping task: Methodological and theoretical considerations. *Brain and Cognition*, 38(3), 317–338. <https://doi.org/10.1006/brcg.1998.1039>
- Bor, D., Duncan, J., Wiseman, R. J., & Owen, A. M. (2003). Encoding strategies dissociate prefrontal activity from working memory demand. *Neuron*, 37(2), 361–367. [https://doi.org/10.1016/S0896-6273\(02\)01171-6](https://doi.org/10.1016/S0896-6273(02)01171-6)
- Breska, A., & Deouell, L. Y. (2014). Automatic Bias of Temporal Expectations following Temporally Regular Input Independently of High-level Temporal Expectation. *Journal of Cognitive Neuroscience*, 26, 1555–1571. <https://doi.org/10.1162/jocn>
- Breska, A., & Deouell, L. Y. (2017). Dance to the rhythm, cautiously: Isolating unique indicators of oscillatory entrainment. *PLoS Biology*, 15(9), 2–6. <https://doi.org/10.1371/journal.pbio.2003534>
- Breska, A., & Ivry, R. B. (2018). Double dissociation of single-interval and rhythmic temporal prediction in cerebellar degeneration and Parkinson's disease. *Proceedings of the National Academy of Sciences of the United States of America*, 115(48), 12283–12288. <https://doi.org/10.1073/pnas.1810596115>
- Buhusi, C. V., & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews Neuroscience*, 6(10), 755–765. <https://doi.org/10.1038/nrn1764>
- Corsi, P. M. (1972). *Human memory and the medial temporal region of the brain*. McGill University, Montreal, QB.
- Cowan, N. (2010). The Magical Mystery Four: How Is Working Memory Capacity Limited, and Why? *Current Directions in Psychological Science*, 19(1), 51–57. <https://doi.org/10.1177/0963721409359277>

- Cowan, N., Elliott, E. M., Saults, S. J., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, 51(1), 42–100. <https://doi.org/10.1016/j.cogpsych.2004.12.001>
- deBettencourt, M. T., Keene, P. A., Awh, E., & Vogel, E. K. (2019). Real-time triggering reveals concurrent lapses of attention and working memory. *Nature Human Behaviour*, 3(August). <https://doi.org/10.1038/s41562-019-0606-6>
- Doelling, K. B., Assaneo, M. F., Bevilacqua, D., Pesaran, B., & Poeppel, D. (2019). An oscillator model better predicts cortical entrainment to music. *Proceedings of the National Academy of Sciences*, 201816414. <https://doi.org/10.1073/pnas.1816414116>
- Doelling, K. B., & Poeppel, D. (2015). Cortical entrainment to music and its modulation by expertise. *Proceedings of the National Academy of Sciences of the United States of America*, 112(45), E6233–E6242. <https://doi.org/10.1073/pnas.1508431112>
- Fiebelkorn, I. C., & Kastner, S. (2019). A Rhythmic Theory of Attention. *Trends in Cognitive Sciences*, 23(2), 87–101. <https://doi.org/10.1016/j.tics.2018.11.009>
- Fischer, J., & Whitney, D. (2014). Serial dependence in visual perception. *Nature Neuroscience*, 17(5), 738–743. <https://doi.org/10.1038/nn.3689>
- Fries, P. (2015). Rhythms for Cognition: Communication through Coherence. *Neuron*, 88(1), 220–235. <https://doi.org/10.1016/j.neuron.2015.09.034>
- Gallistel, C. R. (2011). Mental Magnitudes. In S. Dehaene & E. Brannon (Eds.), *Space, time and number in the brain: Searching for the foundations of mathematical thought* (pp. 3–12). Academic Press.
- Garnier, S., Ross, N., Rudis, R., Camargo, P. A., Sciaini, M., & Scherer, C. (2021). *viridis - Colorblind-Friendly Color Maps for R* (R package version 0.6.2). <https://doi.org/10.5281/zenodo.4679424>
- Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. *Annals of the New York Academy of Sciences*, 423(1), 52–77.
- Gibbon, John. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, 84(3), 279–325. <https://doi.org/10.1037/0033-295X.84.3.279>
- Gmeindl, L., Walsh, M., & Courtney, S. M. (2011). Binding serial order to representations in working memory: a spatial/verbal dissociation. *Memory & Cognition*, 39(1), 37–46. <https://doi.org/10.3758/s13421-010-0012-9>
- Golob, E. J., & Mock, J. R. (2019). Auditory spatial attention capture, disengagement, and response selection in normal aging. *Attention, Perception, and Psychophysics*, 81(1), 270–280. <https://doi.org/10.3758/s13414-018-1611-y>
- Grabenhorst, M., Maloney, L. T., Poeppel, D., & Michalareas, G. (2021). Two sources of uncertainty independently modulate temporal expectancy. *Proceedings of the National Academy of Sciences*, 118(16), 1–11. <https://doi.org/10.1073/pnas.2019342118>

- Gresch, D., Boettcher, S., Ede, F. van, & Nobre, A. C. (2021). Shielding working-memory representations from temporally predictable external interference. *PsyArXiv*.  
<https://psyarxiv.com/wtesm/>
- Haegens, S., & Zion Golumbic, E. (2018). Rhythmic facilitation of sensory processing: A critical review. *Neuroscience and Biobehavioral Reviews*, 86(December 2017), 150–165. <https://doi.org/10.1016/j.neubiorev.2017.12.002>
- Ivry, R. B. (1996). The representation of temporal information in perception and motor control. *Current Opinion in Neurobiology*, 6(6), 851–857.  
[https://doi.org/10.1016/S0959-4388\(96\)80037-7](https://doi.org/10.1016/S0959-4388(96)80037-7)
- Ivry, R. B., & Hazeltine, R. E. (1995). Perception and Production of Temporal Intervals Across a Range of Durations: Evidence for a Common Timing Mechanism. *Journal of Experimental Psychology: Human Perception and Performance*, 21(1), 13–18.
- Ivry, R. B., & Keele, S. W. (1989). Timing functions of the cerebellum. *Journal of Cognitive Neuroscience*, 1(2), 136–152. <https://doi.org/10.1162/jocn.1989.1.2.136>
- Ivry, R. B., & Richardson, T. C. (2002). Temporal control and coordination: The multiple timer model. *Brain and Cognition*, 48(1), 117–132.  
<https://doi.org/10.1006/brcg.2001.1308>
- Ivry, R. B., & Spencer, R. M. C. (2004). The neural representation of time. *Current Opinion in Neurobiology*, 14(2), 225–232. <https://doi.org/10.1016/j.conb.2004.03.013>
- Jacoby, N., & McDermott, J. H. (2017). Integer Ratio Priors on Musical Rhythm Revealed Cross-culturally by Iterated Reproduction. *Current Biology*, 27(3), 359–370.  
<https://doi.org/10.1016/j.cub.2016.12.031>
- Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, 83(5), 323–355.  
<https://doi.org/10.1037/0033-295X.83.5.323>
- Kiyonaga, A., Scimeca, J. M., Bliss, D. P., & Whitney, D. (2017). Serial Dependence across Perception, Attention, and Memory. *Trends in Cognitive Sciences*, 21(7), 493–497.  
<https://doi.org/10.1016/j.tics.2017.04.011>
- Konstantinou, N., Beal, E., King, J. R., & Lavie, N. (2014). Working memory load and distraction: Dissociable effects of visual maintenance and cognitive control. *Attention, Perception, and Psychophysics*, 76(7), 1985–1997.  
<https://doi.org/10.3758/s13414-014-0742-z>
- Lavie, N., & De Fockert, J. (2005). The role of working memory in attentional capture. *Psychonomic Bulletin and Review*, 12(4), 669–674.  
<https://doi.org/10.3758/BF03196756>
- Lawrence, M. A. (2016). Package ‘ez.’ In *R package 3.1 version* (4.4-0).  
<https://github.com/mike-lawrence/ez>
- Lorenc, E. S., Mallett, R., & Lewis-Peacock, J. A. (2021). Distraction in Visual Working Memory: Resistance is Not Futile. *Trends in Cognitive Sciences*, 25(3), 228–239.  
<https://doi.org/10.1016/j.tics.2020.12.004>

- Makovski, T., & Jiang, Y. V. (2008). Proactive interference from items previously stored in visual working memory. *Memory and Cognition*, 36(1), 43–52. <https://doi.org/10.3758/MC.36.1.43>
- Michalka, S. W., Kong, L., Rosen, M. L., Shinn-Cunningham, B. G., & Somers, D. C. (2015). Short-Term Memory for Space and Time Flexibly Recruit Complementary Sensory-Biased Frontal Lobe Attention Networks. *Neuron*, 87(4), 882–892. <https://doi.org/10.1016/j.neuron.2015.07.028>
- Miller, K. J., Foster, B. L., & Honey, C. J. (2012). Does rhythmic entrainment represent a generalized mechanism for organizing computation in the brain? *Frontiers in Computational Neuroscience*, 6(October), 1–3. <https://doi.org/10.3389/fncom.2012.00085>
- Nobre, A. C., & van Ede, F. (2018). Anticipated moments: Temporal structure in attention. *Nature Reviews Neuroscience*, 19(1), 34–48. <https://doi.org/10.1038/nrn.2017.141>
- Oberauer, K. (2002). Access to Information in Working Memory: Exploring the Focus of Attention. *Journal of Experimental Psychology: Learning Memory and Cognition*, 28(3), 411–421. <https://doi.org/10.1037/0278-7393.28.3.411>
- Oberauer, K. (2019). Working Memory and Attention – A Conceptual Analysis and Review. *Journal of Cognition*, 2(1), 1–23. <https://doi.org/10.5334/joc.58>
- Okazaki, Y. O., De Weerd, P., Haegens, S., & Jensen, O. (2014). Hemispheric lateralization of posterior alpha reduces distracter interference during face matching. *Brain Research*, 1590(1), 56–64. <https://doi.org/10.1016/j.brainres.2014.09.058>
- Palmer, C., & Krumhansl, C. L. (1990). Mental Representations for Musical Meter. *Journal of Experimental Psychology: Human Perception and Performance*, 16(4), 728–741. <https://doi.org/10.1037/0096-1523.16.4.728>
- Parncutt, R. (1994). A Perceptual Model of Pulse Salience and Metrical Accent in Musical Rhythms. *Music Perception: An Interdisciplinary Journal*, 11(4), 409–464.
- R Core Team. (2018). *R: A Language and Environment for Statistical Computing*. <https://www.r-project.org/>
- Ravignani, A., Thompson, B., Lumaca, M., & Grube, M. (2018). Why do durations in musical rhythms conform to small integer ratios? *Frontiers in Computational Neuroscience*, 12(November), 1–9. <https://doi.org/10.3389/fncom.2018.00086>
- Ricker, T. J. (2015). The role of short-term consolidation in memory persistence. *AIMS Neuroscience*, 2(4), 259–279. <https://doi.org/10.3934/Neuroscience.2015.4.259>
- Rohenkohl, G., Coull, J. T., & Nobre, A. C. (2011). Behavioural dissociation between exogenous and endogenous temporal orienting of attention. *PLoS ONE*, 6(1), 1–5. <https://doi.org/10.1371/journal.pone.0014620>
- Rubin, J. D., Atwood, S., & Olson, K. R. (2020). Studying Gender Diversity. *Trends in Cognitive Sciences*, 24(3), 163–165. <https://doi.org/10.1016/j.tics.2019.12.011>
- Rudis, B. (2018). Hrbthemes: Additional themes, theme components and utilities for

“ggplot2.” In *Hrbrthemes Documentation*.

- Shalev, N., Nobre, A. C., & van Ede, F. (2019). Time for What? Breaking Down Temporal Anticipation. *Trends in Neurosciences*, 42(6), 373–374. <https://doi.org/10.1016/j.tins.2019.03.002>
- van Ede, F., Niklaus, M., & Nobre, A. C. (2017). Temporal expectations guide dynamic prioritization in visual working memory through attenuated  $\alpha$  oscillations. *Journal of Neuroscience*, 37(2), 437–445. <https://doi.org/10.1523/JNEUROSCI.2272-16.2016>
- van Wassenhove, V. (2016). Temporal cognition and neural oscillations. *Current Opinion in Behavioral Sciences*, 8, 124–130. <https://doi.org/10.1016/j.cobeha.2016.02.012>
- Wang, L., Amalric, M., Fang, W., Jiang, X., Pallier, C., Figueira, S., Sigman, M., & Dehaene, S. (2019). Representation of spatial sequences using nested rules in human prefrontal cortex. *NeuroImage*, 186(July 2018), 245–255. <https://doi.org/10.1016/j.neuroimage.2018.10.061>
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>
- Wickham, H., François, R., Henry, L., & Müller, K. (2021). *dplyr: A Grammar of Data Manipulation* (1.0.7). <https://github.com/tidyverse/dplyr>
- Wilke, C. O., Wickham, H., & Wilke, M. C. O. (2019). Package ‘cowplot.’ In *Streamlined Plot Theme and Plot Annotations for ‘ggplot2’*.
- Winfree, A. T. (1967). Biological Rhythms and the Behavior of Populations of Coupled Oscillators. *Journal of Theoretical Biology*, 16, 15–42. [papers3://publication/uuid/1D1315DF-787F-4E67-9319-AE7E6F89F044](https://pubmed.ncbi.nlm.nih.gov/11111111/)
- Zalta, A., Petkoski, S., & Morillon, B. (2020). Natural rhythms of periodic temporal attention. *Nature Communications*, 11(1), 1–12. <https://doi.org/10.1038/s41467-020-14888-8>

## Chapter 6

# Conclusions and Future Work

### 6.1 Discussion

In the Cognitive Sciences, the “levels of understanding” as described by Marr (Marr, 1982) have provided a framework by which to approach study of the brain and cognition. These levels are threefold and interdependent — achieving some understanding at one level can facilitate study that leads to understanding at another. The *Computational* level is most abstract, describing the ultimate problem that a complex system has evolved to solve. Study at the *Algorithmic* level sheds light on the processes and constraints of the system that carries out said Computation. The final and most fundamental level, that of *Implementation* describes the physical system that gives rise to the other levels (Krakauer et al., 2017; Marr, 1982).

In this dissertation work, I have considered *working memory* to be naturally positioned at Marr’s level of Computation. As reviewed in Chapter 2, working memory has been modeled in various ways and from different angles. Each model

captures at least some portion of the emergent properties that result from numerous processes, continually interacting across the brain, which together give rise to goal-driven control over representations in the here and now. The work I present here, however, is agnostic to any specific model. Though this work may inform the work of others, and may be built upon in the future, it in itself does not necessarily increase insight or understanding at the level of Computation.

This work, rather, has been guided by more pragmatic aims. Chief among them, is the aim to increase the specificity with which working memory as a cognitive construct is examined. Toward this aim, I have placed emphasis on identifying and discretizing among different structures of information that may serve as inputs to working memory, in an effort to identify and reveal instances in which processes and behavior may differ based upon them. Further, I have targeted sex, a biologically determined factor which differs across individuals, and presented findings that indicate understanding the brain activity underlying behavior may in some instances require more specificity regarding the particular human being studied.

In Chapter 3, I presented evidence for systematic and predictable differences in the neural processes underlying working memory for spatial information between females and males. It is important to note that the ultimate outcome of behavior, did not differ between these groups: females and males perform equally well. The neural activity that enabled them to do so is what differed. Among the prior research that informed and motivated this study, were

studies having found a strong relationship between cyclical sex-hormone fluctuations and cyclical peak alpha frequency fluctuations (e.g., Bazanova et al., 2014). In follow up, then, a future study in which an assay of key sex-hormones is conducted along with EEG and behavioral measures might provide evidence as to whether hormones are causally linked to the sex-based differences that were found.

In the laboratory it is commonplace to place constraints on the number of factors that are independently varied. In daily life, however, humans contend with and adapt to continuously changing environments and unexpected events. In Chapter 4, I took a step towards the study of young adult human behavior in an environment with a modest level of complexity by presenting multimodal stimuli while individuals performed a visual task. The effect on behavior resulting from an interaction between spatial and temporal structures across modalities in this study, shows that we attend to and integrate multimodal information present in our environment even when doing so is detrimental.

In other research that has focused upon how the brain changes as we age, it has been shown that in later adulthood as we become more senior, we become less adept at filtering out irrelevant information during early stages of processing (Cabeza et al., 2018; Reuter-Lorenz & Park, 2010). In follow-up to the study presented here, a study explicitly including participants across the aging spectrum from young to later adulthood, performing a similar multimodal task while undergoing neuroimaging, could reveal whether irrelevant information is

subjected to different degrees of processing across the aging spectrum, to different effects upon behavior. Such a study may shed further light on the findings of (Gmeindl & Courtney, 2012), in which differential effects on visuospatial working memory performance arose from the presence of sound, perhaps based upon aging-related factors.

The focus of the work I have presented here has been on a specific classification of working memory — that which represents information which is both visual and spatial in nature. It is necessary to highlight this classification because goal-driven behaviors that require visuospatial information as inputs may differ from those which are reliant upon or require other types of informational forms. For example, it has been shown previously (Gmeindl et al., 2011), that different methods of consolidating and reorganizing information may be automatically employed when working memory is engaged to handle spatial versus verbal information.

In Chapter 5, I present the design of a novel visuospatial working memory task, utilized to study the effects of rhythmic temporal structure upon working memory. I show that rhythmic and arrhythmic temporal structures have particular impact upon temporal aspects of performance. Further, I presented behavioral evidence indicating that spatial information may be reorganized differently based upon both temporal structure and the specific goals an individual has. These findings have real life implications, particularly among individuals who regularly perform visuospatial tasks. In the field of radiology, for example, a core function of

the medical professional is the identification of suspicious visual targets within images. Understanding how visuospatial content may be reorganized, and in result have an impact on behavior, may be key in developing methods of image presentation or best practices for the review of images to best support the most desirable outcomes on performance.

The dissertation work I have presented increases understanding primarily at Marr's level of *Algorithm*. As such, it serves as one single contribution in answer to a larger call that has been made of the neuroscience community: for improved characterization of behavior that will better guide study at the level of Implementation and inform broader understanding at the level of Computation (Krakauer et al., 2017). Importantly, I have shown that some differences between individuals should not be overlooked. Further, I have demonstrated that while we are adept at representing various types of information, our working memory for spatial information in particular, is susceptible to and influenced by the structures present within it.

## 6.2 References

- Bazanova, O. M., Kondratenko, A. V., Kuzminova, O. I., Muravlyova, K. B., & Petrova, S. E. (2014). EEG alpha indices depending on the menstrual cycle phase and salivary progesterone level. *Human Physiology*, 40(2), 140–148. <https://doi.org/10.1134/S0362119714020030>
- Cabeza, R., Albert, M., Belleville, S., Craik, F. I. M., Duarte, A., Grady, C. L., Lindenberger, U., Nyberg, L., Park, D. C., Reuter-Lorenz, P. A., Rugg, M. D., Steffener, J., & Rajah, M. N. (2018). Maintenance, reserve and compensation: the cognitive neuroscience of healthy ageing. *Nature Reviews Neuroscience*, 19(11), 701–710. <https://doi.org/10.1038/s41583-018-0068-2>
- Gmeindl, L., & Courtney, S. M. (2012). Deconstructing spatial working memory and attention deficits in multiple sclerosis. *Neuropsychology*, 26(1), 57–70. <https://doi.org/10.1037/a0026213>
- Gmeindl, L., Walsh, M., & Courtney, S. M. (2011). Binding serial order to representations in working memory: a spatial/verbal dissociation. *Memory & Cognition*, 39(1), 37–46. <https://doi.org/10.3758/s13421-010-0012-9>
- Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A., & Poeppel, D. (2017). Neuroscience Needs Behavior: Correcting a Reductionist Bias. *Neuron*, 93(3), 480–490. <https://doi.org/10.1016/j.neuron.2016.12.041>
- Marr, D. (1982). The Philosophy and the Approach. In *Vision : a computational investigation into the human representation and processing of visual information* (pp. 8–38). San Francisco: W.H. Freeman.
- Reuter-Lorenz, P. A., & Park, D. C. (2010). Human Neuroscience and the Aging Mind: at Old Problems A New Look. *Journals of Gerontology: Psychological Sciences*, 65B(4), 405–415. <https://doi.org/10.1093/geronb/gbq035>