The Mental Representation of Visual Information

Joel Robitaille

Psychology (Behavioural and Cognitive Neuroscience)

Submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

Faculty of Social Sciences, Brock University
St. Catharines, Ontario

General Abstract

Despite working in relative independence, the working memory and imagery literatures investigate the mental representation of visual information. Recent reports investigating the neural structure and their associated functional activity responsible for the creation and maintenance of these cognitive representations suggest a significant overlap between these fields of study. Because each field has adopted methodologies that does not allow for a direct comparison of the mental representation described by their respective literatures, it is difficult to determine whether imagery and working memory representations are related. Hence, the current thesis further investigates the properties of the visual representation of visual information to bridge between the imagery and working memory fields.

In a first study, I compare the psychophysical properties of simple stimuli commonly used in working memory reports with more complex objects adopted by the imagery field. In the course of three experiments, I demonstrate that the cost of stimulus complexity predominantly affects the quality of the mental representation while still providing evidence of a shared cognitive mechanism driving the formation and maintenance of these representations. In a second study, I evaluate the impact of mental rotation on these mental representations as well as whether the adoption of different paradigms, along with different performance metrics, assess the same cognitive construct. Here again, I show strong evidence in support of a common cognitive mechanism driving the performance across mental manipulation and through assessment methods. Finally, the last study attempted to track the manipulation of these visual representations by applying an encoding model to raw EEG activity. While I show evidence of the orientation-relevant activity during perception, the encoding model does not detect reliable enough activity to allow for tracking the orientation of the stimulus during retention and mental rotation. Together, this thesis provides evidence of a shared cognitive mechanism that drives visual working memory and imagery representation, but tracking these mental representation using EEG activity during manipulation remains unclear.

Keywords: visual working memory, imagery, visual representation, complex objects, paradigm, forward encoding model

Acknowledgements

First and foremost, I would like to thank my supervisor, Dr. Stephen Emrich, for his mentorship, enthusiasm, endless patience, and constant support throughout my graduate work. I greatly appreciate all your help, expertise, and motivation, which allowed for the completion of this thesis. I wish to extend my thanks to my committee members, Dr. Karen Arnell and Dr. Sidney Segalowitz, for their genuine interest, valuable feedback, and support during this journey.

To my friends and colleagues from the Brock University Visual Cognitive Neuroscience Laboratory – Christine Salahub, Holly Lockhart, Tom Nelson, Joseph Capozza, Brenda deWit – as well as my fellow graduate friends – Kari Lustig, Brent Pitchford, Andrew Chung, Kristen Baker, Kevin MacDonald, Kevin Mulvihill – I am grateful for your assistance and support.

Finally, I wish to express my gratitude to my family – André Bédard, Claude Robitaille, and David Robitaille – and my wife, Katharyn Stanclik, for believing in me and their endless encouragement and support through my ups and downs.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

%acc.	percent accuracy
%inacc.	percent inaccuracy
ANOVA	analysis of variance
AUC	area under the curve
BF	Bayesian factor
BMS	Bayesian model selection
BOLD	blood-oxygenation level-dependent
CC	counter-clockwise
CDA	contralateral delay activity
CS	composite score
CTFs	channel tuning functions
CW	clockwise
DIC	deviance information criterion
EEG	electroencephalography
fa	false alarms
fMRI	functional magnetic resonance imaging
h	hits
HEOG	horizontal electrooculogram
ICA	independent component analysis
ICs	independent components
IPS	intraparietal sulcus
ITI	inter-trial interval
k	capacity
M	mean
MEG	magnetoencephalography
MVPA	multivariate pattern analysis
OSIVQ	object-spatial and verbal questionnaire
P_g	proportion of guessed response
P_T	proportion of target response
PC	principal component
PCA	principal component analysis

PET	positron emission tomography
RC	rotated component
RE	raw error
ROC	receiver operating characteristics
RT	reaction time
rTMS	repetitive transcranial magnetic stimulation
SD	standard deviation
TLL	total log likelihood
VEOG	vertical electrooculogram
VSTM	visual short-term memory
VVIQ	vividness of visual imagery questionnaire
VVIQ-rv	vividness of visual imagery questionnaire revised version
WM	working memory

Chapter 1

General Introduction

"Try to pose for yourself this task: not to think of a polar bear, and you will see that the cursed thing will come to mind every minute" (Dostoyevsky, 1997). While this observation by Dostoyevsky highlights the difficulty in inhibiting intrusive thoughts, this quote overshadows a more intricate phenomenon: how are we experiencing this polar bear image in the first place? The mental representation of visual information has been extensively studied since the 1970s by both the imagery and working memory literatures. Interestingly enough, these fields have studied this phenomenon independently with very few references to one another (Tong, 2013). The lack of overlap between these fields of research is mainly due to divergent research questions and methodologies. However, recent discoveries in the neural correlates associated with working memory and imagery suggest that the mental representation of these processes are strongly overlapping (Albers et al., 2013; Ganis et al., 2004; Harrison & Tong, 2009 to name a few). While these findings are promising, we are left with two problems: (1) there is no clear indication whether the behavioural and cognitive properties of mental representations are the same, and (2) the neural processes mitigating the manipulation of these representations are still unknown. This chapter will address what we currently know about the mental representations of visual information as well as how this thesis proposes to address these two issues.

Conceptual and Behavioural properties of Mental Representations

Mental imagery and visual working memory describe a series of processes that includes the creation of a mental representation of visual information that is maintained and that can be manipulated at will (Tong, 2013). One particular distinction that can differentiate these mental representations from one another pertains to how they are generated. Indeed, while working memory typically comes from perception, imagery representations can be drawn from long-term memory storage. A recent review suggests that these processes can be differentiated by the extent to which they adopt bottom-up and top-down processes (Dijkstra et al., 2019). The imagery literature, however, is not very diligent about this distinction and will often provide a controlled visual stimulus for their experimental manipulation, which resemble the bottom-up process that would define working memory. In some cases where both working memory and imagery are reported, the formation and maintenance of a mental representation of visual information are sometimes referred to as the working memory aspect of the task whereas the manipulation of the information is defined as the imagery component (e.g., Albers et al., 2013). This distinction allows for the adherence to the conservative definitions, which maximizes the bottom-up processes associated with working memory as well as the top-down manipulation of imagery. Even though there may be some fundamental distinctions between working memory and imagery representations, both literatures investigate the representation, maintenance, and manipulation of visual information.

The theoretical framework that explains imagery and working memory processes has also been argued to be the same, where both representations can be described according to Baddeley and Hitch's (1974) model of working memory. This model was put forward as an attempt to address the shortcomings of the short-term memory model. One central innovation of the working memory model is that it allows for the manipulation of the mental representation maintained in memory. Therefore, this addition accounts for the flexibility and fragility of the mental representation, which applies to both imagery (Kosslyn, 1975, 1980) and working memory processes (Baddeley & Hitch, 1974). According to Baddeley and Hitch's (1974) three-component model, a central executive system is responsible for task inhibition, updating, and shifting (Miyake et al., 2000; Wongupparaj et al., 2015) as well as coordinating and binding the information coming from the slave systems, referred to as the phonological loop and the visuospatial sketchpad (Baddeley, 2010). While the phonological loop describes the encoding process associated with language, the visuospatial sketchpad component of the model is of particular interest for visual working memory and imagery representations as it describes how visual information is bound together into a mental representation. This slave system is further subdivided into the visual cache, the visual buffer, and the inner scribe (Cocchini et al., 2002; Logie, 1995; Pearson & Keogh, 2019), which provide different properties to the visual representation, while being interconnected. Some have argued that the earliest form of working memory representation begins in the visual buffer where low-level information is maintained (Cocchini et al., 2002; Pearson & Keogh, 2019). Likewise, imagery representations – particularly from the depictive format hypothesis, which will be discussed later – are believed to be maintained in topographically organized areas, recruiting low-level features (Kosslyn, 1980; Kosslyn et al., 2006; Kosslyn & Thompson, 2003). The visual cache serves as a temporary storage of a more abstract version of the representation (Cocchini et al., 2002; Pearson & Keogh, 2019). While adding spatial and movement information to the visual representation, the inner scribe is also responsible for rehearsing the information from the visual cache and the buffer to serve the goals of the central executive (Cocchini et al., 2002; Logie, 1995).

The working memory model (Baddeley & Hitch, 1974) proposes a type of architecture that can be investigated using neural imaging studies. Early evidence, however, reported greater neural activation, recorded using functional magnetic resonance imaging (fMRI) blood-oxygenation level-dependent (BOLD) signals, in the prefrontal and parietal cortices during working memory tasks (Curtis & D'Esposito, 2003; Todd & Marois, 2005). While these regions are often associated with high-level cortical – or task-specific – activity (Lee et al., 2013; Riggall & Postle, 2012), new methods investigating more subtle changes in neural activity (i.e., multivariate pattern analysis) demonstrated that low-level sensory activity found in the occipital cortex plays an integral part in working memory processes and, more particularly, in the formation and manipulation of the visual information (Albers et al., 2013; Emrich et al., 2013; Harrison & Tong, 2009; Serences et al., 2009). These results gave rise to the sensory recruitment hypothesis, which suggests that modality-sensitive information maintained in memory recruits the same sensory areas that encoded them during perception (Postle, 2006; Serences, 2016; Serences et al., 2009). Therefore, the primary visual cortex may be a viable candidate for the visual buffer component of the visuospatial sketchpad. Furthermore, a recent study (Ester et al., 2015) demonstrated that the activity detected in sub-parts of the frontal and parietal cortices also track some feature-sensitive information associated with working memory representations. Bettencourt and Xu (2016) showed that the activity found in intra-parietal sulcus that is associated with working memory representation is more closely related to behavioural performance than the more precise feature-specific information recorded in the early visual cortex. The feature-specific activity associated with the intra-parietal sulcus is also found to be of lower precision than the activity detected in the visual cortex (Galeano Weber et al., 2016). Although these results are still preliminary, they suggest that the occipital and parietal cortices respectively are possible candidates for the visual buffer and visual cache sub-components of the visuospatial sketchpad and that mental representations are likely maintained in these areas.

Interestingly, Kosslyn's quasi-pictorial model of imagery (Kosslyn, 1980) describes several components that map on to the working memory model. Indeed, this model is composed of higher-order components (e.g., associative memory, attention shifting, object or spatial properties processing, etc.) that feed to a visual buffer, which is believed to correspond to the primary visual cortex (Kosslyn, 1980; Kosslyn et al., 2006; Kosslyn & Thompson, 2003). The visual representation found in the visual buffer can be generated not only from perception but also from long-term memory, but is in a fragile state that needs constant re-activation (Kosslyn, 1980; Kosslyn et al., 2006) and is reliant on attentional resources (Logie & Salway, 1990, 1990; Pearson et al., 1996).

Although there is no explicit evidence or report of the overlap between this model with Baddeley and Hitch's (1974), both models describe a visual buffer component that relies on sensory processes to maintain relevant information for the formation and maintenance of mental representations as well as higher-cortical components that allows for the manipulation of these representations. Yet, there is still very little convergence in the study of the psychometrical properties of these representations. Over several decades, these fields have made significant progress in investigating mental representation behaviourally. Because they have asked different questions and used different methods, however, it is difficult to determine how these representations overlap. Hence, the next section will discuss the psychometrical properties of visual mental representations and the methods adopted by the working memory and imagery fields.

Working memory representations

Following a psychophysical approach, the working memory field has focused primarily on the encoding and storage of information coming from perception, as well as its particularly limited capacity, which distinguishes it from long-term or iconic memory. One of the earliest accounts of memory capacity comes from Miller's (1956) work who reported that participants can on average remember up to seven bits of information regardless of the modality. Philips (1974) demonstrated a distinction in the memory process between the iconic – or sensory – memory, and short-term memory. Through a series of experiments, Philips (1974) showed that participants have the ability to remember complex patterns with great precision as long as the retention period was short (ideally <100ms) and no movement or masking is present. This iconic memory is shown to have near infinite capacity, but it is fast-decaying (Dick, 1974; Sperling, 1960). Using alphanumerical characters, Pashler (1988) was able to replicate the large capacity associated with iconic memory as well as demonstrate that short-term memory shows a significantly more restrictive capacity-limit, in that participants can only remember about 4-5 alphanumerical character at once..

More recent work on working memory demonstrated that capacity is more limited than previously thought. Indeed, a seminal study by Luck and Vogel (1997) conducted a series of dichotomous change-detection tasks where participants were asked whether a probed square, selected from a briefly presented set of 1, 2, 3, 4, 8, or 12 coloured squares, changed colour after a short delay. Participants were able to accurately detect colour changes when a maximum of 3-4 items were presented, after which the presentation of more items was met with a sharp decline in performance. These results were reliable across a number of variations of the critical feature (i.e., colour, orientation, size, etc.) as well as with a conjunction of different features and varying memory retention delays. The authors concluded that, once the item’s features are bound into a single object, only a limited number of working memory representations can be stored, and that capacity is not subject to the number of features one object possesses. These results reported by Luck and Vogel (1997), suggesting a more limited capacity than Miller’s (1956) and Pashler’s (1988) accounts, began a debate on the representation of information in visual memory.

While investigating this limited working memory capacity, Zhang and Luck (2008) noticed that when participants are required to report the exact feature of a probed item on a continuous scale (i.e., continuous recall paradigm), the distribution of participants’ responses over several trials, although predictably normally distributed, exhibited long tails, indicating that participants are occasionally reporting some of the farthest possible expression of a feature. Zhang and Luck (2008) theorized that participants can either encode the probed item successfully and report a feature close to the target – which produces a normal distribution over several trials – or fail to encode the item and provide a random guess response – which produces a uniform distribution. These guessed responses would ultimately be responsible for the observed long tails. Therefore, by modelling participants’ response distribution according to a mixture of these two distributions, one can estimate the overall proportion of guesses and the precision of one’s responses, derived from the width of the normal distribution. Using this technique, they demonstrated that increasing the number of items beyond the putative 3 – 4 item capacity limit did not affect the precision in participants' responses when they successfully encoded the item, but participants would increase their guess rate. These results demonstrate that items are less likely to be encoded successfully as cognitive load increases. Therefore, the mental representation of these items can be described according to the discrete representation model of working memory, which suggests that the mental representations can be stored with high precision into a memory "slot" but items that don't make it to a "slot" will be lost.

The discrete view of mental representation, however, is not without criticism and is mainly challenged by resource-based models of working memory. Indeed, Wilken and Ma (2004) describes a series of experiments that adopt both dichotomous and continuous response paradigms, showing that participants' performance is mainly limited by internal noise generated by increasing number of items encoded to memory. Therefore, performance doesn't appear to rely on discrete "slots" or an upper capacity limit, as participants show a more progressive decrease in performance than what Luck and Vogel (1997) originally reported. Furthermore, Bays and Husain (2008) showed response precision decreases progressively as memory load increases, following a simple power law. Instead of a perfectly accurate mental representation of the memorized item, as suggested by the discrete representation model, these results propose that memory representations vary in quality depending on the working memory resource allocated during the encoding process. When all these working memory models are tested against each other (for a review of the models, see Ma et al., 2014; van den Berg et al., 2014), the best-fitting model is attributed to a variant of the resources models known as the variable precision model (van den Berg et al., 2012, 2014). This model posits that memory resources are not only flexible and limited across items presented at once but also vary on a trial-by-trial basis. Therefore, the mental representation of these items would be blurred to a variable degree across items and trials.

There have also been investigations into a possible neural correlate of this working memory capacity limit. Vogel and Machizawa (2004) observed that a sustained event-related potential called the contralateral delay activity (CDA) tracked the number of items held in memory and reaches maximum amplitude at 3-4 items. Comparably, a similar pattern was reported in fMRI studies, where blood-oxygenation level-dependent (BOLD) activation was observed in the intraparietal sulcus (IPS) to increase as the number of items held in memory (Todd & Marois, 2004; Xu & Chun, 2006). These neural correlates were initially interpreted under the discrete representation model as they match the capacity limit. A recent reinterpretation of these reports, however, suggest that these results can be accommodated by the continuous resource model (Bays, 2018). A more recent investigation of the CDA component demonstrates that, when items are given different priorities, the CDA also tracks the amount of cognitive resources applied among items (Salahub et al., 2019), which further support the resource models of working memory representations.

While the mental representation of information stored in working memory is still debated, the psychophysical approach to the behavioural investigation of these representations has allowed for the development of methods and psychophysical measures that have given us greater insight into the properties of these mental representations. Despite some recent work that criticizes the validity of these models altogether (Schurgin et al., 2018), adopting this approach to the mental representation (e.g., adopting a continuous response paradigm, modelling of participants' responses) has allowed for the investigation of working memory properties which was not possible until the development of these methods and analytical techniques.

Imagery representations

In contrast to the psychophysical analysis of the capacity of visual working memory, which has focused on the psychophysical properties of its representations (i.e., capacity, precision), since the 1970s, the imagery field has been debating the format in which mental representation of visual information takes, where two major theories have emerged: depictive/pictorial and propositional formats. The depictive format describes mental representation of visual information as being kept in a state that preserves the features, as well as the spatial relations between features, of an object which corresponds closely to the physical object it represents. This description offers straightforward behavioural predictions about mental imagery as they share crucial properties with physical objects. For instance, in a seminal study, Shepard and Metzler (1971) presented pairs of 3D geometric figures where participants were required to evaluate the pair as identical objects other than the orientation, or as is a mirror image of the other. To make this judgement, participants are required to mentally rotate one (or both) objects and determine whether the features are corresponding. In the case where participants are pictorially representing these objects, the mental rotation should take longer when the angular difference between rotated objects is greater. Shepard and Metzler (1971) demonstrated a linear relationship between reaction time and the angular disparity between the objects, as predicted by the theory. These specific findings were replicated multiple times, particularly within the sex difference literature as men show a small, but significant, advantage in mental rotation ability (see Maeda & Yoon, 2013, for a review and meta-analysis). Similarly, Stephen Kosslyn (1973), who is an influential proponent of the depictive format of imagery, showed that participants require more time to mentally search for features of a memorized illustration when these features are more spatially distant to one another. Kosslyn (1973) concluded that, as someone needs more time to complete a visual search when targets are further, participants require more time to screen through the memorized map to find the target. Similarly, further evidence in favour of the depictive format was evidenced by studies requiring participants to mentally zoom (Kosslyn, 1980), rotate (Cooper & Shepard, 1973, 1973; Shepard & Cooper, 1986), and re-organize images (Reed, 1974; Reisberg & Chambers, 1991). It should be noted, however, that most of the behavioural evidence for the pictorial format is mainly based on a dichotomous response paradigm where chronometry is the critical measure, which gives limited insight into the properties of these representations.

In opposition to the pictorial hypothesis, the propositional format, pioneered by Zenon Pylyshyn (1973), describes imagery as a language-like or symbolic classification. Although this format does not deny the possibility of a visual experience, these “images” are believed to be an epiphenomenal experience, which doesn’t provide any functional purpose. Unfortunately, this possible format in imagery representation is rather difficult to assess as it does not offer clear predictions that can be tested behaviourally. Therefore, most of the evidence brought forth is often presented against the pictorial format. For instance, several reports demonstrated that some imagery tasks can be performed without requiring visual processes. Marmor and Saback (1976) replicated the results reported by

Shepard and Metzler (1971) with congenitally blind individuals. Similarly, in a task that requires participants to mentally arrange the relative object placement of 5 items based on a set of descriptions, Fleming et al. (2006) demonstrated no significant performance or chronometry difference between blind, blindfolded and sighted individuals. Furthermore, Pylyshyn argues that the mental rotation process is not performed holistically on a given figure. Indeed, using eye-tracking, Just and Carpenter (1976) showed that participants made several visual checks between the two figures in an attempt to match the features on each display. These results suggest that participants are rotating different parts of the image one at a time and not the whole object at once. Moreover, contrary to the pictorial format assumption that rotations are completed at a steady rate, Pylyshyn (1979) demonstrated that the rate of rotations is highly dependent on the complexity of the object as well as the complexity of comparison task. Although these studies are not directly presenting evidence in favour of the propositional format of mental representations, they demonstrate that the depictive format hypothesis is not without flaws.

While the imagery debate continued throughout several decades, the evaluation of these processes using behavioural methods has significantly diminished for a few reasons. First, as mentioned earlier, the sex difference discovered in the Sheppard & Metzler (1971) study generated an exciting new branch of literature investigating sex differences in cognitive abilities, which is mainly tangential to the understanding of the nature of mental images. Then, the development of dual-code theory (Paivio, 1978) opened another field of research with important implications to the study of cognitive processes, which includes thinking styles (Paivio, 1975; Paivio & Harshman, 1983), bilingualism (Paivio & Lambert, 1981), and education (Clark & Paivio, 1991) to name a few. This theory proposed by Paivio (1978) suggests that verbal and imagery representations that are functionally distinguishable systems. Interestingly, these systems seem to map on to the phonological loop and the visuospatial sketchpad slave systems described in Baddeley and Hitch's (1974) working memory model. Pylyshyn (1973) criticized this theory when applied to the imagery debate as it has limited explanatory power concerning the format of mental representations, as it mainly focuses on the separation of these processes, and the literature has not addressed the mechanisms responsible for each code. Finally, the investigation of imagery format through neuro-imaging methods appears to favour the depictive format and possibly ended the debate (Pearson & Kosslyn, 2015). Therefore, the emergence of the sex difference in imagery abilities and the dual-code theory as well as the evaluation of neural correlates of imagery representation has changed the focus away from the investigation of the behavioural properties of imagery representations in the past several decades.

Neural Basis for Mental Representations

The investigation into the neural mechanisms of mental representations was initiated by the imagery literature in the 80s. Indeed, neurophysiological evidence from brain lesion literature provided a new line of evidence in support of the depictive format of imagery representations. In a review by Farah (1984), individuals with visual impairments (e.g., achromatopsia, agnosia, etc.) often report the same impairments in their experience of mental imagery. In some cases, however, individuals report the ability to generate visual representation despite suffering from visual blindness (Chatterjee & Southwood, 1995). Similarly, some patients afflicted with visual agnosia can recall and identify objects from memory, while they are unable to recognize physical objects from perception (Behrmann et al., 1992; Jankowiak et al., 1992). Although there is a significant body of evidence supporting the overlap in the generation of visual representation coming from perception and imagery, these last reports suggest that the overlap is far from being perfect. Recent reports suggest that these "inconsistencies" may be explained as a function of where in the bottom-up and top-down processes these impairments occur (Kok et al., 2016; Roelfsema & de Lange, 2016).

The development of fMRI technology and new analysis methods has allowed for a more accurate evaluation of the neural correlates associated with the formation of mental representations. Here again, the depictive format appears to be favoured over the propositional format, mostly because a large portion of the visual system is also recruited when participants are performing imagery manipulation. Indeed, in univariate analyses of BOLD signals from fMRI, two-thirds of the neural structures activated during perception were also recruited during an imagery task (Ganis et al., 2004; see also Kosslyn et al., 1997 for comparable PET results). Moreover, disrupting the functional activity found in the visual cortex, using repetitive transcranial magnetic stimulation (rTMS), slowed down participants' response time and decreased behavioural accuracy in the imagery task (Kosslyn et al., 1999). Thus, these early findings suggest that the visual cortex plays a central role in the processes mediating mental representations.

As mentioned earlier, several brain areas were identified as neural correlates of working memory processes, mainly within the prefrontal and parietal cortices (Curtis & D'Esposito, 2003; Todd & Marois, 2004; for a comprehensive review see D'Esposito & Postle, 2015). While several overlapping brain areas are recruited during the generation of imagery and working memory representation of visual information, univariate analyses cannot differentiate feature-specific information about the mental representation. Indeed, the functional purpose of these increases in BOLD activity may be the result of task-demands, attentional demands, error correction, etc. Although it is uncertain whether imagery and working memory are represented similarly, the recruitment of similar areas suggests that they could be.

With the development of multivariate pattern analysis (MVPA), it became possible to investigate the representational nature of the functional activity associated with BOLD signals. Indeed, Harrison and Tong (2009) used this method to train a pattern classifier to recognize the activity found in the visual cortex that was most predictive of the perception of two orientation gratings. In a working memory task where participants were required to hold both gratings in memory, the pattern classifier was able to detect – above chance level – the presence of orientation-specific activity for both orientations during the retention phase as well as disappearance of the activity associated with the un-cued orientation. More interestingly, the authors demonstrated that the pattern classification performance decreased up the visual hierarchy (i.e., V1, V2, V3, and V3A/V4). These results demonstrate that subtle activation changes in the visual cortex are more telling of the orientation feature of the stimulus held in memory than the stronger BOLD activation detected in the prefrontal and parietal cortices (see Bettencourt & Xu, 2016; and E. F. Ester et al., 2015 in particular). Also, not unlike the imagery literature, it appears that the visual cortex is also recruited during the generation of working memory representation (Serences et al., 2009).

Although there is a significant body of evidence suggesting the involvement of the occipital lobe during imagery manipulation, these reports were similarly based on univariate analyses. As mentioned earlier, this approach cannot comment on the functional contribution of early visual cortex to specific representations in mental imagery. Based on Harrison and Tong's (2009) study, Albers et al. (2013) trained a pattern classifier on the perception and memory of three orientation gradients, and added imagery trials, where participants were cued to mentally rotate the pattern gradient in the direction that matched either the un-cued pattern gradients or the pattern that was not presented during the trial. The pattern classifier was able to recognize the activity as participants dropped one of the items and transformed the gradient into another orientation. Therefore, these results suggest that the visual cortex contains enough feature-sensitive activity to track the mental representation while it is manipulated.

While fMRI allows for great spatial resolution, the application of this classification method to electroencephalography (EEG) and magnetoencephalography (MEG) has allowed for evaluation for the temporal generalization of the activity associated with mental representations. Indeed, by training a classifier at a given time point and testing its accuracy on all possible time points of a given trial, it is possible to evaluate how generalizable this activity is across time. Reports using EEG and MEG activity from a working memory task – which generally consist of presenting an object for a brief moment (e.g., < 500ms) and a retention period – reveal that the activity associated with the mental representation recorded during the perception of the object is still present during the retention period (Carlson et al., 2011; Cichy et al., 2014; Isik et al., 2013). These results suggest some activity associated with the mental representation of a stimulus is temporally generalizable. Therefore, it may be possible to use this method to track these mental representations while they are manipulated.

Current Thesis

Given the strong overlap in the conceptualization and neural correlates between the mental representation of visual information described in the imagery and working memory literatures, the first objective of this thesis is to address the lack of behavioural specificity of the mental representation of visual information. Because the behavioural assessment of mental representations conducted by the working memory and imagery fields have been drastically different, there is little understanding of whether the behavioural properties of mental representation coming from either literature are the same. Indeed, both literatures include the formation and manipulation of visual information in their definition. The working memory field, however focused more on the static memory representation while the imagery field has historically been more motivated in investigating the manipulation of visual information. The differences in approach in the behavioural investigation of mental representation by either field has led to several methodological decisions that makes the direct comparison of psychophysical properties difficult. The methodological differences addressed in this thesis are threefold: (1) simple versus complex stimuli, (2) dichotomous versus continuous paradigms, and (3) choice of performance metrics (i.e., percent and absolute errors). As such, Chapter 2 and 3 directly evaluate these three methodological differences and directly compare the performance when participants are required to form and retain a memory representation (Chapter 2) as well as to mentally rotate it in an imagery manipulation (Chapter 3). Since the cortical structures as well as functional activity when participants engage in a working memory or imagery task show significant overlap, the literature has recently developed methods that may allow for the evaluation of the manipulation of mental representations.

As such, a second goal of this thesis is to determine whether it is possible to track the manipulation of visual information from EEG activity (Chapter 4).

Chapter 2 focuses on the psychophysical properties of stimuli commonly used in each literature. While the working memory literature focuses on simple stimuli that vary in one key feature (e.g., coloured squares, line orientation, etc.), the imagery field typically uses significantly more complex objects (e.g., 3D block, common household objects, maps, etc.), which introduces several variables (e.g., complexity, dimensionality, familiarity, etc.) that impact behavioural performance, but are not fully understood. In a series of three experiments, Chapter 2 examines the mental representations of simple and complex stimuli and their effect on memory capacity, stimulus complexity, and the axis of rotation. All experiments demonstrate a performance cost associated with the complexity of the item memorized, where response precision is primarily affected.

After investigating the static mental representation of simple and complex objects, Chapter 3 describes a behavioural investigation of the psychophysical properties of mental representations when they are mentally rotated. In a first experiment, the psychophysical properties of simple and complex stimuli were evaluated using the same paradigm as in Chapter 2, but with the addition of a mental rotation manipulation task. The performance cost associated complexity was replicated, as there were more errors in reporting the orientation of complex items compared to simple stimuli. In the second experiment, participants were presented with the dichotomous imagery task inspired by the seminal Shepard & Metzler (1971) experiment as well as a continuous report version of the task. After standardizing participants' performance scores, no significant differences in performance were found between the continuous and dichotomous paradigms across the mental rotations. While these results indicate that the effects of mental rotation on performance are captured by both report methods, Principal Component Analysis indicates that a single component can account for large portion of the variance present in both tasks and across imagery conditions. Therefore, both methods appear to assess the same mental representation regardless of whether they are static or manipulated.

Finally, Chapter 4 shows an exploratory analysis to determine whether it is possible to track imagery manipulation from EEG activity. It builds on the findings of a previous thesis by Robitaille (2016) where an encoding model was applied to induced and evoked alpha band activity to identify the orientation feature of the mental representation of a line during perception, memory retention, and imagery rotation. Although results showed that induced alpha activity can track the orientation of the stimulus during the presentation and retention phase of the experiment, it was not possible to detect the specific orientation of the mental representation after being transformed (Robitaille, 2016). Chapter 4 describes an experiment that adopts the same encoding method applied to raw EEG activity as an attempt to track orientation beyond memory retention. At first glance, the encoding model appears to be able to detect the orientation of the stimulus while being maintained and after it was manipulated. The statistical analysis of the activity recovered, however, is not reliable enough to clearly identify the orientation of the mental representation beyond the presentation.

Together, these findings show evidence that imagery and working memory are largely relying on similar processes. Chapter 2 and 3 also show that there is, however, some evidence of more independent factors at play in that affect the psychometrical properties of mental representations that are linked to the methodology employed as well as individual differences. Despite the strong overlap in the mechanisms driving these mental representations, Chapter 4 suggests that memory and imagery representations may not rely on the same neural activity.

References

Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & de Lange, F. P. (2013). Shared Representations for Working Memory and Mental Imagery in Early Visual Cortex. Current Biology, 23(15), 1427–1431. https://doi.org/10.1016/j.cub.2013.05.065

Baddeley, A. D. (2010). Working memory. Current Biology, 20(4), R136–R140. https://doi.org/10.1016/j.cub.2009.12.014

Baddeley, A. D., & Hitch, G. (1974). Working Memory. In G. H. Bower (Ed.), Psychology of Learning and Motivation (Vol. 8, pp. 47–89). Academic Press. https://doi.org/10.1016/S0079-7421(08)60452-1

Bays, P. M. (2018). Reassessing the Evidence for Capacity Limits in Neural Signals Related to Working Memory. Cerebral Cortex (New York, N.Y.: 1991), 28(4), 1432–1438. https://doi.org/10.1093/cercor/bhx351

Bays, P. M., & Husain, M. (2008). Dynamic Shifts of Limited Working Memory Resources in Human Vision. Science (New York, N.Y.), 321(5890), 851–854. https://doi.org/10.1126/science.1158023

Behrmann, M., Winocur, G., & Moscovitch, M. (1992). Dissociation between mental imagery and object recognition in a brain-damaged patient. Nature, 359(6396), 636–637. https://doi.org/10.1038/359636a0

Bettencourt, K. C., & Xu, Y. (2016). Decoding the content of visual short-term memory under distraction in occipital and parietal areas. Nature Neuroscience, 19(1), 150–157. https://doi.org/10.1038/nn.4174

Carlson, T. A., Hogendoorn, H., Kanai, R., Mesik, J., & Turret, J. (2011). High temporal resolution decoding of object position and category. Journal of Vision, 11(10), 9–9. https://doi.org/10.1167/11.10.9

Chatterjee, A., & Southwood, M. H. (1995). Cortical blindness and visual imagery. Neurology, 45(12), 2189–2195.

Cichy, R. M., Pantazis, D., & Oliva, A. (2014). Resolving human object recognition in space and time. Nature Neuroscience, 17(3), 455–462. https://doi.org/10.1038/nn.3635

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. Educational Psychology Review, 3(3), 149–210. https://doi.org/10.1007/BF01320076

Cocchini, G., Logie, R. H., Sala, S. D., MacPherson, S. E., & Baddeley, A. D. (2002). Concurrent performance of two memory tasks: Evidence for domain-specific working memory systems. Memory & Cognition, 30(7), 1086–1095. https://doi.org/10.3758/BF03194326

Cooper, L. A., & Shepard, R. N. (1973). Chronometric studies of the rotation of mental images. In W. G. Chase (Ed.), Visual Information Processing (pp. 75–176). Academic Press. https://doi.org/10.1016/B978-0-12-170150-5.50009-3

Curtis, C. E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. Trends in Cognitive Sciences, 7(9), 415–423. https://doi.org/10.1016/S1364-6613(03)00197-9

D'Esposito, M., & Postle, B. R. (2015). The cognitive neuroscience of working memory. Annual Review of Psychology, 66, 115–142. https://doi.org/10.1146/annurev-psych-010814-015031

Dick, A. O. (1974). Iconic memory and its relation to perceptual processing and other memory mechanisms. Perception & Psychophysics, 16(3), 575–596. https://doi.org/10.3758/BF03198590

Dijkstra, N., Bosch, S. E., & van Gerven, M. A. J. (2019). Shared Neural Mechanisms of Visual Perception and Imagery. Trends in Cognitive Sciences, 23(5), 423–434. https://doi.org/10.1016/j.tics.2019.02.004

Dostoyevsky, F. (1997). Winter Notes on Summer Impressions. Northwestern University Press.

Emrich, S. M., Riggall, A. C., Larocque, J. J., & Postle, B. R. (2013). Distributed patterns of activity in sensory cortex reflect the precision of multiple items maintained in visual short-term memory. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 33(15), 6516–6523. https://doi.org/10.1523/JNEUROSCI.5732-12.2013

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

Farah, M. J. (1984). The neurological basis of mental imagery: A componential analysis. Cognition, 18(1), 245–272. https://doi.org/10.1016/0010-0277(84)90026-X

Fleming, P., Ball, L. J., Ormerod, T. C., & Collins, A. F. (2006). Analogue versus propositional representation in congenitally blind individuals. Psychonomic Bulletin & Review, 13(6), 1049–1055. https://doi.org/10.3758/bf03213924

Galeano Weber, E. M., Peters, B., Hahn, T., Bledowski, C., & Fiebach, C. J. (2016). Superior Intraparietal Sulcus Controls the Variability of Visual Working Memory Precision. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 36(20), 5623–5635. https://doi.org/10.1523/JNEUROSCI.1596-15.2016

Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2004). Brain areas underlying visual mental imagery and visual perception: An fMRI study. Cognitive Brain Research, 20(2), 226–241. https://doi.org/10.1016/j.cogbrainres.2004.02.012

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. Nature, 458(7238), 632–635. https://doi.org/10.1038/nature07832

Isik, L., Meyers, E. M., Leibo, J. Z., & Poggio, T. (2013). The dynamics of invariant object recognition in the human visual system. Journal of Neurophysiology, 111(1), 91–102. https://doi.org/10.1152/jn.00394.2013

Jankowiak, J., Kinsbourne, M., Shalev, R. S., & Bachman, D. L. (1992). Preserved visual imagery and categorization in a case of associative visual agnosia. Journal of Cognitive Neuroscience, 4(2), 119–131. https://doi.org/10.1162/jocn.1992.4.2.119

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. Cognitive Psychology, 8(4), 441–480. https://doi.org/10.1016/0010-0285(76)90015-3

Kok, P., Bains, L. J., van Mourik, T., Norris, D. G., & de Lange, F. P. (2016). Selective Activation of the Deep Layers of the Human Primary Visual Cortex by Top-Down Feedback. Current Biology: CB, 26(3), 371–376. https://doi.org/10.1016/j.cub.2015.12.038

Kosslyn, S. M. (1973). Scanning visual images:Some structural implications. Perception & Psychophysics, 14(1), 90–94. https://doi.org/10.3758/BF03198621

Kosslyn, S. M. (1975). Information representation in visual images. Cognitive Psychology, 7(3), 341–370. https://doi.org/10.1016/0010-0285(75)90015-8

Kosslyn, S. M. (1980). Image and Mind. Harvard University Press.

Kosslyn, S. M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J. P., Thompson, W. L., Ganis, G., Sukel, K. E., & Alpert, N. M. (1999). The role of area 17 in visual imagery: Convergent evidence from PET and rTMS. Science (New York, N.Y.), 284(5411), 167–170.

Kosslyn, S. M., & Thompson, W. L. (2003). When is early visual cortex activated during visual mental imagery? Psychological Bulletin, 129(5), 723–746. https://doi.org/10.1037/0033-2909.129.5.723

Kosslyn, S. M., Thompson, W. L., & Alpert, N. M. (1997). Neural Systems Shared by Visual Imagery and Visual Perception: A Positron Emission Tomography Study. NeuroImage, 6(4), 320–334. https://doi.org/10.1006/nimg.1997.0295

Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). The Case for Mental Imagery (1 edition). Oxford University Press.

Lee, S.-H., Kravitz, D. J., & Baker, C. I. (2013). Goal-dependent dissociation of visual and prefrontal cortices during working memory. Nature Neuroscience, 16(8), 997. https://doi.org/10.1038/nn.3452

Logie, R. H. (1995). Visuo-spatial Working Memory. Psychology Press.

Logie, R. H., & Salway, A. F. S. (1990). Working memory and modes of thinking: A secondary task approach. In Lines of thinking: Reflections on the psychology of thought, Vol. 2: Skills, emotion, creative processes, individual differences and teaching thinking (pp. 99–113). John Wiley & Sons.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. Nature, 390(6657), 279–281. https://doi.org/10.1038/36846

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

Maeda, Y., & Yoon, S. Y. (2013). A Meta-Analysis on Gender Differences in Mental Rotation Ability Measured by the Purdue Spatial Visualization Tests: Visualization of Rotations (PSVT:R). Educational Psychology Review, 25(1), 69–94. https://doi.org/10.1007/s10648-012-9215-x

Marmor, G. S., & Zaback, L. A. (1976). Mental rotation by the blind: Does mental rotation depend on visual imagery? Journal of Experimental Psychology: Human Perception and Performance, 2(4), 515–521. https://doi.org/10.1037/0096-1523.2.4.515

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63(2), 81–97. https://doi.org/10.1037/h0043158

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “Frontal Lobe” tasks: A latent variable analysis. Cognitive Psychology, 41(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Paivio, A. (1975). Imagery and synchronic thinking. Canadian Psychological Review/Psychologie Canadienne, 16(3), 147–163. https://doi.org/10.1037/h0081801

Paivio, A. (1978). Imagery and Verbal Processes (1 edition). Psychology Press.

Paivio, A., & Harshman, R. (1983). Factor Analysis of a Questionnaire on Imagery and Verbal Habits and Skills*. Canadian Journal of Psychology, 37(4), 461–483. https://doi.org/10.1037/h0080749

Paivio, A., & Lambert, W. (1981). Dual coding and bilingual memory. Journal of Verbal Learning and Verbal Behavior, 20(5), 532–539. https://doi.org/10.1016/S0022-5371(81)90156-0

Pashler, H. (1988). Familiarity and visual change detection. Perception & Psychophysics, 44(4), 369–378. https://doi.org/10.3758/BF03210419

Pearson, D. G., Logie, R. H., & Green, C. (1996). Mental manipulation, visual working memory, and executive processes. Psychologische Beitrage, 38(3–4), 324–342.

Pearson, J., & Keogh, R. (2019). Redefining Visual Working Memory: A Cognitive-Strategy, Brain-Region Approach. Current Directions in Psychological Science, 0963721419835210. https://doi.org/10.1177/0963721419835210

Pearson, J., & Kosslyn, S. M. (2015). The heterogeneity of mental representation: Ending the imagery debate. Proceedings of the National Academy of Sciences, 112(33), 10089–10092. https://doi.org/10.1073/pnas.1504933112

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. Perception & Psychophysics, 16(2), 283–290. https://doi.org/10.3758/BF03203943

Postle, B. R. (2006). Working Memory as an Emergent Property of the Mind and Brain. Neuroscience, 139(1), 23–38. https://doi.org/10.1016/j.neuroscience.2005.06.005

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. Psychological Bulletin, 80(1), 1–24. https://doi.org/10.1037/h0034650

Pylyshyn, Z. W. (1979). The rate of "mental rotation" of images: A test of a holistic analogue hypothesis. Memory & Cognition, 7(1), 19–28. https://doi.org/10.3758/BF03196930

Reed, S. K. (1974). Structural descriptions and the limitations of visual images*. Memory & Cognition, 2(2), 329–336. https://doi.org/10.3758/BF03209004

Reisberg, D., & Chambers, D. (1991). Neither pictures nor propositions: What can we learn from a mental image? Canadian Journal of Psychology, 45(3), 336–352. https://doi.org/10.1037/h0084297

Riggall, A. C., & Postle, B. R. (2012). The relationship between working memory storage and elevated activity as measured with functional magnetic resonance imaging. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 32(38), 12990–12998. https://doi.org/10.1523/JNEUROSCI.1892-12.2012

Robitaille, J. (2016). Encoding Induced Alpha EEG Activity Reveals Visual Working Memory Representations and their Manipulations. http://hdl.handle.net/10464/10688

Roelfsema, P. R., & de Lange, F. P. (2016). Early Visual Cortex as a Multiscale Cognitive Blackboard. Annual Review of Vision Science, 2, 131–151. https://doi.org/10.1146/annurev-vision-111815-114443

Salahub, C., Lockhart, H. A., Dube, B., Al-Aidroos, N., & Emrich, S. M. (2019). Electrophysiological correlates of the flexible allocation of visual working memory resources. Scientific Reports, 9(1), 19428. https://doi.org/10.1038/s41598-019-55948-4

Schurgin, M. W., Wixted, J. T., & Brady, T. (2018). Psychophysical Scaling Reveals a Unified Theory of Visual Memory Strength. BioRxiv, 325472. https://doi.org/10.1101/325472

Serences, J. T. (2016). Neural mechanisms of information storage in visual short-term memory. Vision Research, 128, 53–67. https://doi.org/10.1016/j.visres.2016.09.010

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

Shepard, R. N., & Cooper, L. A. (1986). Mental Images and Their Transformations (New edition edition). The MIT Press.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. Science (New York, N.Y.), 171(3972), 701–703.

Sperling, G. (1960). The information available in brief visual presentations. Psychological Monographs: General and Applied, 74(11), 1–29. https://doi.org/10.1037/h0093759

Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. Nature, 428(6984), 751–754. https://doi.org/10.1038/nature02466

Tong, F. (2013). Imagery and visual working memory: One and the same? Trends in Cognitive Sciences, 17(10), 489–490. https://doi.org/10.1016/j.tics.2013.08.005

van den Berg, R., Awh, E., & Ma, W. J. (2014). Factorial Comparison of Working Memory Models. Psychological Review, 121(1), 124–149. https://doi.org/10.1037/a0035234

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

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. Nature, 428(6984), 748–751. https://doi.org/10.1038/nature02447

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. Journal of Vision, 4(12), 11. https://doi.org/10.1167/4.12.11

Wongupparaj, P., Kumari, V., & Morris, R. G. (2015). The relation between a multicomponent working memory and intelligence: The roles of central executive and short-term storage functions. Intelligence, 53, 166–180. https://doi.org/10.1016/j.intell.2015.10.007

Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. Nature, 440(7080), 91–95. https://doi.org/10.1038/nature04262

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. Nature, 453(7192), 233–235. https://doi.org/10.1038/nature06860

Chapter 2

Examining Effects of Dimensionality and Complexity in Visual Short-Term Memory using Continuous Recall

Abstract

While significant advances have been made to evaluate the psychophysical properties of short-term memory storage, most studies focus on the memoranda of simple 2D objects. When complexity is evaluated, most studies rely on familiar or naturalistic stimuli (e.g., faces) which have been shown to increase memory performance. Here, I investigate the role of complexity (i.e., the number of features), specifically on the dimensionality (i.e., spatial extent based on the presence of monocular depth cues) of the stimulus, on the psychophysical properties of visual short-term memory (VSTM) for 2.5D objects that are unfamiliar. Using delayed-response continuous report tasks, I demonstrate decrease memory recall associated with complexity across load (Experiment 1), dimensionality (Experiment 2), and axis of rotation (Experiment 3) manipulations. By applying a two-components mixture model, I also reveal that complexity affects participants' memory encodability or accessibility to a greater extent than on the quality of the memoranda. I provide evidence for a shared memory mechanism across all experiments but also some specific processes that distinguish feature dimension and axis of rotation as separate properties of VSTM.

Introduction

The temporary storage and manipulation of information, known as short-term memory, has been studied extensively for the past few decades. One particular characteristic of short-term memory that distinguishes it from other types of memory, such as episodic and long-term memory, is its limited capacity to store information (Baddeley, 2003; Cowan, 2001), where increasing visual memory load using simple stimuli (e.g., colored squares, line orientation, etc.) results in a decrease in short-term memory performance (Luck & Vogel, 1997; Vogel, Woodman, & Luck, 2001; Pashler, 1988). Although there are still debates surrounding which model best characterizes VSTM representations (Suchow, Fougnie, Brady, & Alvarez, 2014; M. Schurgin & Brady, 2018; van den Berg & Ma, 2018), these models tend to agree that the amount of information that can be represented in memory decreases with increased memory load, particularly for the first few items (Bays & Husain, 2008; Emrich, Lockhart, & Al-Aidroos, 2017; van den Berg & Ma, 2014).

Recently, the effect of memory load on visual short-term memory (VSTM) performance has been frequently measured using a continuous report paradigm (Wilken and Ma, 2004), which allows participants to report the specific feature of a memorized item continuously along its feature dimension (e.g., colour wheel or orientation). This method not only provides a measure of error (i.e., the angular difference between the reported feature value and the target), but also provides a way to model these responses and estimate other psychophysical measures of memory. For instance, Zhang and Luck (2008) created a mixture model based upon two principles. First, given the chance to freely report the exact feature of the memorized item, participants reporting an item that is successfully encoded and stored in memory should report a value close to the target feature (i.e., with low error or high precision). Over several trials, these responses will be normally distributed around the target value where the width of this distribution is inversely related to the precision of one's memory. Second, when an item is not stored in memory, participants will pick a random feature value, which will generate a uniformly distributed pattern of responses over the course of an experiment. Consequently, Zhang and Luck (2008) applied a mixture model (i.e., circular normal and uniform) to the participants' responses to estimate the precision of their memory representations separately from the number of items stored in memory. This modeling method has been used to investigate additional aspects of VSTM, such as one's propensity to report distractors (Bays, Catalao, & Husain, 2009) and the variability in VSTM representations on a trial-by-trial basis (van den Berg et al., 2012; Fougnie et al., 2012), for example.

While numerous studies and models have attempted to assess the properties of VSTM representations from continuous recall paradigms, these studies have tended to rely on the use of tightly controlled, surface features such as line orientation and color. By contrast, studies investigating the memory representation of more complex, or naturalistic stimuli, generally rely only on dichotomous-choice tasks. For instance, using a change detection task, Luck and Vogel (1997) examined the effect of complexity on VSTM performance, and demonstrated that the capacity of memory was similar regardless of whether the stimuli were composed of a single feature or a conjuncture of multiple features. However, these results have not been reliably replicated (see Delvenne & Bruyer, 2004, and Wheeler & Treisman, 2002). Indeed, in a series of experiments, Wheeler and Treisman (2002) demonstrated that items held in memory would only compete along their common feature dimension (e.g., orientation, shape, colour), suggesting that separate memory capacity exists in parallel for different feature dimensions. In a systematic investigation of stimulus complexity, Alvarez & Cavanagh (2004) ranked the complexity of a number of different stimuli (e.g. squares, letters, random polygons, shaded cubes) based on a visual search processing speed where they showed a linear inverse relationship between complexity and memory capacity as assessed by change detection. Thus, much like the traditional load manipulation carried out by increasing the number of distinct items to be remembered, these studies suggest that the integration of multiple visual features into an object increases the total memory load.

Although dichotomous-choice tasks have revealed that complexity affects working memory performance, few studies have examined memory performance for complex stimuli using continuous-recall tasks; thus, it is unclear whether this impairment is due to a reduction in memory precision, or due to an increase in the likelihood of forgetting complex objects. Studies that have examined continuous recall for complex stimuli have tended to use more naturalistic objects such as faces. For example, Lorenc et al. (2014) used memory recall for faces in a continuous response task and found that upright faces were recalled with greater precision than inverted faces. Similar effects are observed for own-race compared to other-race faces (Zhou, Mondloch, & Emrich, 2018), suggesting that the holistic integration of face stimuli, rather than complexity per se, may confer this particular advantage (Curby & Gauthier, 2007). Although these stimuli are substantially more complex than single-feature items, these studies are complicated by the fact that memory performance is typically better for naturalistic stimuli than for objects of similar complexity (Brady et al., 2016; Brady & Störmer, 2020a). For instance, more faces can be held in short-term memory than other complex non-face stimuli (Curby & Gauthier, 2007) or in comparison to more common memory stimuli such as line orientations (Jiang, Shim, & Makovski, 2008). Comparable findings are observed with other objects with which individuals show expertise. Xie and Zhang (2017) demonstrated that individuals who are familiar with a certain generation of Pokémon characters have a clear memory advantage over unfamiliar (from a more recent generation) characters. Therefore, it seems that there is evidence for a memory advantage for naturalistic stimuli over abstract simple-featured items – specifically for those with which we have extensive familiarity – despite the apparent requirement for more cognitive resources when holding complex stimuli in memory.

Making stimuli more complex and more naturalistic can also introduce additional features to stimuli which may also affect VSTM performance. For example, one feature of objects – particularly naturalistic objects – that is underexplored is dimensionality (i.e., spatial extent of an object suggested by monocular depth cues). There is some evidence suggesting that the closer stimuli are to real physical and tangible objects, the better the memory performance. Indeed, although the cause is not well understood, visual agnosia patients express greater difficulty recognizing an object from a drawing or a picture than when they are faced with the physical object (Riddoch and Humphreys, 1987; Young and Ellis, 1989; Servos et al., 1993; Humphrey et al., 1994; Chainay and Humphreys, 2001; Hiraoka et al., 2009). It is speculated that the presence of depth cues confers an advantage to recognition and memorization (Servos et al., 1993; Chainay and Humphreys, 2001). Snow et al. (2014) evaluated this "real-world" advantage and demonstrated that participants show greater memory performance when presented with real objects than photographs or drawings of the same objects. Thus, like naturalistic memory items, evidence suggests that, although increasing the dimensionality and realism of objects may increase the complexity of those stimuli, rather than increasing memory load, it ultimately confers a benefit to memory performance.

To summarize, although it is clear that increasing stimulus complexity (i.e., augmenting the number of features) tends to impair VSTM performance, it is unclear from dichotomous-choice tasks which aspects of memory performance are affected – whether complex items are remembered with lower resolution, or whether they are more easily forgotten. Moreover, studies that have examined continuous-recall performance for more complex stimuli have tended to use more naturalistic stimuli, making it difficult to compare performance to the more typically used surface features. Consequently, in the current study, I aimed to examine the effect of complexity on VSTM by comparing memory for orientation of simple lines and complex objects. To examine complexity using stimuli that could be easily controlled, while still remaining unfamiliar, I compared continuous recall performance for 2D lines to that of novel 2.5D[1] shapes consisting of an assembly of cubes. In Experiment 1, I compared VSTM for simple and complex objects while also manipulating VSTM load. By using a continuous recall paradigm, I was able to evaluate the accuracy of participants' responses (i.e., raw error), as well as model these responses using a mixture model devised by Zhang & Luck (2008) in order to evaluate

---

[1] Given that these objects are 2D representations of 3D stimuli that projects monocular depth cues (i.e., lightning and shading) without stereoscopic visual cues, they are not real 3D representation of objects.

participants' propensity to guess independently from the precision of their memory. Experiment 2 aimed to separate the effects of dimensionality from complexity by adding complex 2D objects that resembled the complex 2.5D shape outlines, but without the suggested depth cues. I also compared these stimuli on a dichotomous-choice change-detection task. Experiment 3 examined whether the effect of dimensionality on VSTM depended on the direction axes of rotation in which stimuli could be manipulated.

Experiment 1

Methods

Participants. Forty-four (12 males; normal or corrected-to normal vision, 17-32 years of age, M = 19.54) Brock university students participated in the experiment and were compensated with course credits or $10/hr. Participants were screened upon arrival for colour-blindness using Ishihara's 38 plates test. All procedures were cleared by the Brock University Research Ethics Board.

Apparatus. Participants completed the tasks on a 20-in. flat LCD with a resolution of 1600-1200 ppi and a refresh rate of 60 Hz. Stimuli were generated and presented using MATLAB (mathworks.com) software along with the Psychophysics Toolbox extension (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997).

Stimuli. The simple stimuli used in our short-term memory task consisted of a line 4.5° of visual angle long and 0.45° wide. The orientation of this simple stimulus was manipulated from the center point of the stimulus and could span the full 180° of possible positions. The 2.5D complex shape version of the short-term memory task used a selection of 25 objects constructed from an arrangement of 9 cubes with 4-5 90° bends. To ensure cohesion and create one solid object with no floating parts in a given shape, at least one face of every cube faced another cube. These shapes were rendered using OpenGL functions of the Psychophysics Toolbox MATLAB extension. Following creation, they were visually inspected to ensure that they possessed unique and distinguishable features from one another. Depending on their orientation, these 2.5D complex shapes were between 1.32 – 2.6 visual degrees in any direction. At the beginning of each trial, one of the 25 2.5D objects was randomly selected (without replacement), each of which was assigned a random orientation, and were rotated only over the z-axis across set sizes. In a given multi-item trial, simple or complex stimuli were oriented at least 30° from one another.

Procedure. Participants performed a continuous-recall task in which the memory sample varied in set size, with simple or complex objects presented in separate tasks (see Figure 2-1A). A trial would begin with the serial presentation of 1, 2 or 4 stimuli (simple line or 2.5D shape) where each stimulus would be displayed for 500 ms, with a 500 ms inter-stimuli interval (ISI) based on the sequential set size design described by Gorgoraptis, Catalao, Bays, & Husain (2011). Each object's orientation varied along the z-axis and the colour of each object randomly selected between red, blue, green, or yellow. Following a 500 ms delay, one of the colored objects in the sequence reappeared during the probe. Using the mouse, participants reported the orientation of the probed item by rotating the object along the z-axis and clicking to report their response. Each trial was separated by a 1,000 ms inter-trial interval (ITI). The order of the colors and the probe position in the sequence were counterbalanced. Complex 2.5D objects were assigned a random orientation over the x-, y-, and z-axis at the beginning of a given trial and it would only vary across the z-axis between sample and probe. Prior to beginning the simple and complex continuous report tasks, participants were trained on a series of practice trials until they felt comfortable with the task. Participants completed 80 trials for each set size (i.e. 1, 2, or 4), for a total of 240 trials each of simple 2D and complex 2.5D object orientations. To reduce the primacy effect, 23 participants began the experiment with line orientation task first, while the rest began with the complex 2.5D shapes task. All 2.5D stimuli and stimulus presentation paradigm for Experiment all 3 experiments described in this article are available online in an Open Science Framework repository (https://osf.io/9cdqz/).

After the short-term memory tasks were completed, participants were then required to complete a computerized version of the Vividness of Visual Imagery Questionnaire (VVIQ-rv) (Marks, 1973; McKelvie, 1995) as well as the Object-Spatial Imagery and Verbal Questionnaire (OSVIQ) (Blajenkova et al., 2006). As part of a different research question, the score on these questionnaires were used to correlate between memory performance and subjective measures of imagery quality and strategy. The scores from these questionnaires are not reported here and these analyses did not yield any significant correlations with performance.

Analysis. One measure used to assess participants' performance was raw response error, which is calculated by taking the circular standard deviation of the absolute angular distance between the target and reported orientation. This measure was computed for each condition and participant. Raw error is inversely related to response precision, where a lower value corresponds to a response closer to the target orientation. As this measure has been demonstrated as the most reliable measure of performance (see van den

Berg & Ma, 2018, about summary statistics; also Ma, 2018) all reported correlations used raw error as a measure of participants' performance.

Performance was also modeled according to the two-component Mixture Model proposed by Zhang and Luck (see Zhang & Luck 2008), for a full description of the model). The model assumes that the overall distribution of participants' responses come from two types of responses: in trials where participants fail to recall a probed item, participants will guess by picking a random location on the wheel, which in turn will produce a uniform distribution – the greater the proportion of guesses, the greater the height of the uniform distribution. Conversely, in trials where participants successfully encode and store an item in memory, their responses will cluster around the target value and will produce a Von Mises distribution (i.e., circular analog to the Gaussian distribution) centered at the target item. The standard deviation (SD) of the latter distribution is inversely related to the precision of participant's memorized item (i.e., SD = precision^{-1}), that is the smaller SD indicates greater response precision. Therefore, a participant's pattern of responses is produced from the contribution of these two types of trials: guesses and target responses. Using the maximum likelihood estimation (MLE), the distribution of raw error is decomposed into the guess rates (P_g), target rate (P_T) and precision^{-1} (SD) for each condition. Since the guess and target rate are complementary (i.e., $P_g + P_T = 1$), only the proportion of guesses will be reported, along with precision^{-1} (or SD). The parameters of the mixture models were computed using the Memtoolbox Matlab extension (see Suchow, Brady, Fougnie, & Alvarez, 2013; memtoolbox.org).

For statistical comparisons, a 2 (simple vs. complex) x 3 (set sizes of 1, 2, or 4) within-subject frequentist and Bayesian ANOVA was performed on the measures of raw error, mixture model guess rate, and mixture model precision. All statistical analyses were computed using JASP (JASP Team, 2018). Note that, in the frequentist analysis, degrees of freedom were corrected using Greenhouse-Geisser estimate whenever the assumption of sphericity was violated. To explore the interaction between object complexity and set size, two one-way repeated measures ANOVAs were carried to evaluate the effect of set size on each object complexity separately as well as three paired t-test comparisons to evaluate the effect of complexity within each set sizes. Bayesian repeated measures ANOVA were also computed to provide more information about the strength of the evidence in favour of the null or alternative hypotheses. For simplicity's sake, I will only be reporting the Bayesian Factor as it pertains to the alternative hypothesis (i.e., BF_{10}). Note that Bayesian ANOVA compares the strength of the evidence in support of all possible models (i.e., null, main effect 1, main effect 2, both main effects, and both main effects with an interaction). As such, I will be reporting the BF_M which indicates the how many times a given model fits the data compared to the other Bayesian models. For each Bayesian ANOVA, the prior term was set using JASP's default Cauchy width and I report the best-supported models only.

Finally, to determine the extent of one's ability to perform this short-term memory task with simple or complex stimuli, I performed a repeated-measures correlation. Repeated-measures correlations is a measure of a within-subject shared variance designed to investigate the linear relationship between paired repeated data and across participants (Bakdash & Marusich, 2017).

Model comparison. To evaluate the type of strategy that participants adopted for the working memory of simple 2D line and complex 2.5D shapes across our set size manipulation, I adopted a hypothesis-driven factorial comparison as proposed by Dowd, Kiyonaga, Beck, & Egner (2015). This method consists of fitting the data for each participant to a number of mixture models that describes a specific strategy. I created a total of twelve mixture models and adjusted the model's prior distribution ("priors") for the precision (SD) and guess rate (G) parameters (see Table 2-1). The priors were adjusted so that the parameters of the model are either fixed/constant, variable/independent, or ordered across the set sizes. Note that the ordered priors always followed set size 1 < set size 2 < set size 4 for either parameters. The *a priori* assumption for this particular order of priors is that performance is likely to be worse as memory load increases, as demonstrated in the literature. Of course, it is a possibility that a different order would present a better fit. In such a case, I expect that models with variable/independent priors will offer a better fit and would require further investigation. I also included three one-component models that assumes all items are remembered (i.e., no guess rate parameter: ∅G) and only the precision parameter of the von Mises distribution is modeled. These latter models are assessing the possibility that I am forcing the presence of guessed responses to the data when there may be none. In sum, all combinations of our three types of priors applied to both parameters resulted in our twelve models.

Model comparison was conducted by evaluating the Deviant Information Criterion (DIC: Spiegelhalter, Best, Carlin, & van der Linde, 2002), which is a measure of goodness of fit and model complexity. The DIC values are used as a mean to approximate marginal likelihoods, since exact marginal likelihoods typically required for model comparison are computationally difficult to obtain. For each participant and each stimulus type, all twelve models were evaluated by obtaining DIC values using the *MemFit.m* function of the MemToolbox Matlab extension (Suchow, Brady, Fougnie, & Alvarez, 2013; memtoolbox.org). The obtained DIC values were used to compute the total log likelihood ($TLL = -DIC/2$) and submitted to a Bayesian Model Selection (BMS) routine SPM12 (*spm_BMS.m*; http://www.fil.ion.ucl.ac.uk/spm/software/spm12/), which treats models as random effects and returns the expected posterior probability of a model being selected given a random subject as well as the probability of a model over the others (exceedance probability: P^*). It is important to note that P^* should not be interpreted as a p-value despite the fact that it denotes the confidence in model selection, where the bigger the value the greater the confidence, range between 0 and 1 (Rigoux, Stephan, Friston, & Daunizeau, 2014).

Table 2-1.
Description of Mixture Model Prior Probability Distributions for the Precision (SD) and Guess Rate (G) Parameters.

Model number	Guess rate (G)	Precision^{-1} (SD)	Abbreviation
Model 1	-	ss1 = ss2 = ss4	∅G-*f*SD
Model 2	-	ss1 | ss2 | ss4	∅G-*v*SD
Model 3	-	ss1 < ss2 < ss4	∅G-*o*SD
Model 4	ss1 = ss2 = ss4	ss1 = ss2 = ss4	*f*G-*f*SD
Model 5	ss1 | ss2 | ss4	ss1 = ss2 = ss4	*v*G-*f*SD
Model 6	ss1 < ss2 < ss4	ss1 = ss2 = ss4	*o*G-*f*SD
Model 7	ss1 = ss2 = ss4	ss1 | ss2 | ss4	*f*G-*v*SD
Model 8	ss1 = ss2 = ss4	ss1 < ss2 < ss4	*f*G-*o*SD
Model 9	ss1 | ss2 | ss4	ss1 | ss2 | ss4	*v*G-*v*SD
Model 10	ss1 | ss2 | ss4	ss1 < ss2 < ss4	*v*G-*o*SD
Model 11	ss1 < ss2 < ss4	ss1 | ss2 | ss4	*o*G-*v*SD
Model 12	ss1 < ss2 < ss4	ss1 < ss2 < ss4	*o*G-*o*SD

Note: A fixed (or constant) prior relationship between conditions is represented with an equal sign (=) in the parameter columns (i.e., G and SD) and is abbreviated with a lower-case *f* next to its designated parameter in the Abbreviation column. Similarly, ordered relationship is indicated with lesser than sign (<) and abbreviated with the lowercase *o*, whereas a variable (or independent) relationship is depicted by a vertical line (|) and the lowercase *v*. The first three models are one-parameter models with precision as their only parameter and do not include a guess rate parameter (∅G).

Results

The aim of the first experiment was to determine how memory load affects the properties of VSTM in a continuous recall task when participants are required to remember simple or complex stimuli. To do so, participants were presented serially with 1, 2, or 4 stimuli that varied in orientation (180° for simple line stimuli or 360° complex 2.5D objects) and were required to report the orientation of the probed item. I first examined participants' raw error (derived from the angular difference between the target orientation and the reported orientation). Participants responses were also modeled according to Zhang and Luck's (2008) two-component mixture model to obtain participants' proportion of guesses and precision. Raw errors, guess rate, and precision were analyzed independently using repeated ANOVA with within subject factors of stimulus complexity (simple 2D and complex 2.5D) and set size (1, 2, or 4 items). Finally, I also performed a Bayesian model comparison of hypothesis-driven models to investigate the type of strategy adopted by participants for each stimulus.

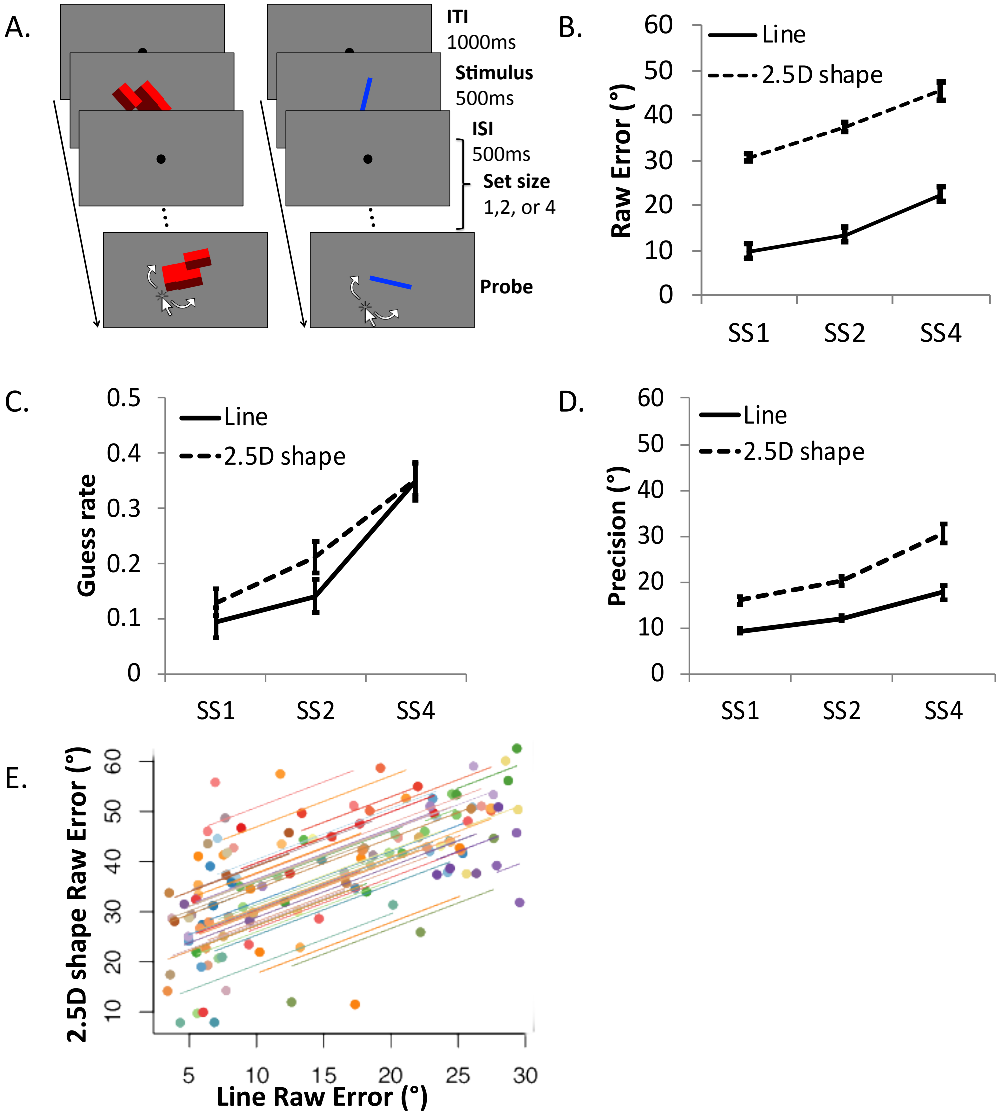

Figure 2-1. Experiment 1 task schematic and behavioural results. (A.) A schematic representation of the delayed-response task used in experiment 1. Participants were serially presented with 1, 2, or 4 objects (simple line or complex 2.5D) and were required to report orientation of a probed object after a short delay. (B.) Participants' mean raw error, (C.) guess rate and (D.) precision^{-1} *(SD) comparison between the report of simple or complex object orientation across set size. Error bars represents standard error of the mean. (E.) Scatterplot of the mean raw error between simple and complex objects across set size for a repeated-measures correlation, where each observation for a participant is given by the same color and line represent and their individual line of best fit,* $r_{rm} = .7179, p = 5.054e^{-15}$.

Raw error. As our first examination of participants' performance, 2 (simple line vs. complex shape stimuli) x 3 (set size of 1, 2, or 4 items) repeated measure frequentist and Bayesian ANOVAs were conducted to evaluate the variation in the standard deviation of the absolute angular difference between the target and reported stimuli orientation (i.e. raw error) (see Figure 2-1B). Consistent with the literature, I demonstrate that performance declines as the number of items held in memory increases regardless of the complexity of the object remembered, $F(2{,}84) = 84.40, p < .001, \eta^2 = 0.134$, where post-hoc comparisons using Bonferroni correction reveals that participants report the target item with more error at larger set sizes for all comparisons, $ps < .001, \log BF_{10} > 9$. Moreover, I found a significant main effect of stimulus complexity, $F(1, 42) = 413.8, p < .001, \eta^2 = 0.54$, where there is a significant increase in raw errors when participant reported complex items, $M_{diff} = 22.63°,\ SE = 1.11°$. These results support the idea that complexity increases the amount of error in VSTM. The absence of an interaction effect between shape complexity and set size ($F(2{,}84) = 1.990, p = .143, \eta^2 = 0.002$) suggests that the memory of simple and complex objects is encoded and stored in a similar fashion. The Bayesian analysis confirms that including complexity of the object and set size without an interaction term offers the best explanation of the data, $\log BF_M = 3.072$. Altogether, these results support the idea that the memorization of complex stimuli requires greater cognitive resources, but that the effect of load and complexity are independent. Indeed, it appears that there is a fixed cognitive cost in memorizing complex 2.5D object that works independently from increases in set size.

To investigate whether short-term memory for line orientation is associated with working memory of the orientation of complex items, I calculated simple-complex object repeated-measures correlations on raw error and resulted in a strongly significant correlation, $r_{rm} = .7179, p = 5.054e^{-15}$. This correlation shows that over 50% of the within-subject variance between simple and complex object for each set size and across each subject is shared.

Mixture Model.

Guess rate. Next, I examined the effect of stimulus type and set size on the proportion of guesses derived for the two-component mixture model (Figure 2-1C). Repeated measures frequentist and Bayesian ANOVA revealed that only the main effect of set size is significant, $F(1.58{,}66.30) = 68.43, p < .001, \eta^2 = 0.2160$, whereas complexity and the interaction are not significant, $F(1{,}42) = 1.86, p = .180, \eta^2 = 0.007$ and $F(1.44{,}60.39) = 2.438, p = .112, \eta^2 = 0.004$, respectively. A Bayesian analysis confirmed that the model of a main effect of set size alone was the preferred model, $\log BF_M = 1.34$, although a model with both effects of set size and complexity showed similar evidence, $\log BF_M = 1.096$. Post-hoc t-test comparisons with a Bonferroni correction between the different set sizes reveal a greater proportion of guesses when 4 objects were presented than 2, $M_{diff} = 17.4\%, SE = 2.2\%, p < .001, \log BF_{10} = 24.99$, or 1, $M_{diff} = 23.9\%,\ SD = 2.5\%, p < .001, \log BF_{10} = 34.15$, and between 2 objects compared to 1, $M_{diff} = 6.5\%, SD = 1.5\%,\ p < .001, \log BF_{10} = 8.72$. Thus, the results suggest that guesses are not significantly greater for complex 2.5 D objects than simple 2D lines.

Target precision^{-1} (SD). The second component of the mixture model consists of the SD (1/precision) in participants' responses. There is a main effect of set size, $F(1.23, 51.61) = 73.52, p < .001, \eta^2 = 0.209$, and complexity, $F(1, 42) = 74.21, p < .001, \eta^2 = 0.198$. Repeated measures frequentist revealed a significant interaction at $p < .05$) between set size and stimulus type, $F(1.23, 51.55) = 4.94, p =$

$.024, \eta^2 = 0.016$). A Bayesian analysis confirmed that the model with two main effects and an interaction was the preferred model, $\log BF_M = 2.6$. Investigating the interaction further revealed significant effects between all pairwise comparisons within stimulus type ($p < .05$), with the exception of between loads 1 and 2 for the line stimuli. Thus, while overall there is lower precision for complex 2.5D objects than simple lines, both stimuli are encoded with more error as set size increases, although the rate of increase is greater for complex objects.

Model comparison. One issue with comparing the results of the standard mixture model is that comparisons are made across parameters, rather than comparing the effects of both parameters simultaneously. Further, comparing across group level results may average out important individual differences that affect the pattern at the aggregate level. Thus, to better compare the ways in which participants encode and maintain the different stimuli, I employed the model comparison technique developed by Dowd et al. (2015). For each participant, raw errors were fitted to our twelve hypothesis-driven mixture models (see Table 2-1) for both the simple line and complex 2.5D stimuli. Figure 2-2 shows DIC values relative to each participant's best fitting model for simple and complex stimuli separately, where the lower values depict a better fit of the participant's data. Using the Bayesian Model Selection (BMS) method, I obtained the posterior probability of selecting each model when a participant is selected at random and the exceedance probability (P^*), which represents each model's overall favourability (see Table 2-2). For simple objects (see Figure 2-2A), most participants showed a better fit for the ordered guess rate and fixed precision model (Model 6: *o*G-*f*SD), which is supported by the results of the BMS analysis as posterior chances of being selected by a random participant is 60.3% with and exceedance probability (P^*) of >99.9%. The second-best model fit for simple items is characterized by ordered guesses and ordered precision (Model 12: oG-oSD), however, the fit is considerably reduced in comparison to Model 6, $posterior = 0.170, P^* = 0.003$. The summed posterior probabilities for all models excluding the best fitting model (i.e., model 6: *o*G-*f*SD) do not exceed 40% (summed: $posterior = 0.397, P^* < .001$).

Similarly, for complex items, the BMS routine shows greater posterior probabilities for Model 6 (*o*G-*f*SD), $posterior = 0.629, P^* > 0.999$, followed by Model 8 ($o$G-$o$SD), $posterior = 0.144, P^* = 4.37e^{-4}$. The posterior probabilities for the models excluding the best model do not exceed 40% (summed: $posterior = 0.371, P^* < .001$). Both simple and complex objects appear to show greater probabilities for all models that have an ordered guess rate parameter (summed *o*G models respectively: $posterior = 0.798, P^* > .999$; $posterior = 0.832, P^* > .999$) and fixed precision (summed *f*SD models respectively: $posterior = 0.642, P^* > .999$; $posterior = 0.667, P^* > .999$), with a small group favoring ordered precision (summed *f*SD models respectively: $posterior = 0.238, P^* < .001$; $posterior = 0.222, P^* < .001$). Although these results are somewhat different from the comparison of the standard mixture model analysis above, they are consistent with the finding that VSTM for lines and 2.5D shapes show a similar pattern of behaviour.

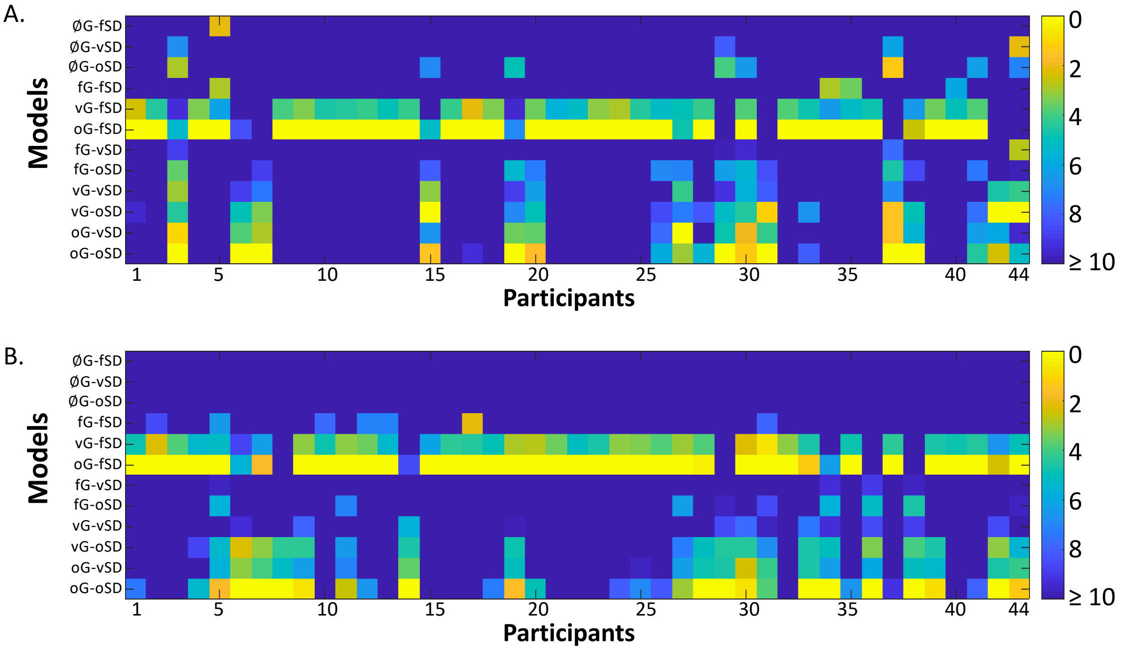

Figure 2-2. Bayesian model comparison between participants (columns) and models (rows) for (A.) line and (B.) 2.5D shapes stimulus. For each participant, values are obtained from the difference in Deviance Information Criteria (DIC) between their best fitting model and a given model, where the lower the DIC difference indicates a better fit of the data depicted in yellow. Each model represents a mixture model with a fixed (f), variable (v), or ordered (o; where set size 1 < set size 2 < set size 4) prior probability distributions for the proportion of guesses (G) and precision (SD) model parameters (see Table 2-1 for details).

Table 2-2.
Posterior and Exceedance probabilities (P^*) across all models and stimulus type

	Simple Line		Complex 2.5D Shape	
	Posterior	P^*	Posterior	P^*
Models				
Model 1: ØG-*f*SD	.018	$6.01e^{-21}$	.018	$5.57e^{-24}$
Model 2: ØG-*v*SD	.020	$6.01e^{-21}$	.018	$5.57e^{-24}$
Model 3: ØG-*o*SD	.019	$6.01e^{-21}$	.018	$5.57e^{-24}$
Model 4: *f*G-*f*SD	.018	$6.01e^{-21}$	.018	$5.57e^{-24}$
Model 5: *v*G-*f*SD	.020	$6.01e^{-21}$	.020	$5.57e^{-24}$
Model 6: *o*G-*f*SD	.603 **	> .999 **	.629 **	> .999 **
Model 7: *f*G-*v*SD	.019	$6.01e^{-21}$	.018	$5.57e^{-24}$
Model 8: *f*G-*o*SD	.019	$6.01e^{-21}$	.018	$5.57e^{-24}$
Model 9: *v*G-*v*SD	.020	$6.01e^{-21}$	.018	$5.57e^{-24}$
Model 10: *v*G-*o*SD	.049	$6.01e^{-21}$	.020	$5.57e^{-24}$
Model 11: *o*G-*v*SD	.024	$6.01e^{-21}$	.019	$5.57e^{-24}$
Model 12: *o*G-*o*SD	.170 *	$7.30\ e^{-5}$ *	.184 *	$7.60\ e^{-5}$ *
Sum excluding:				
Best model	.397	< .001	.371	> .001
Best and second best	.226	< .001	.187	> .001
Sum of all models with:				
No guess rate (ØG)	.057	$1.80e^{-20}$	.055	$1.67e^{-23}$
Fixed guess rate (*f*G)	.056	$1.80e^{-20}$	.055	$1.67e^{-23}$
Variable guess rate (*v*G)	.089	$1.80e^{-20}$	.058	$1.67e^{-23}$
Ordered guess rate (*o*G)	.798	> .999	.832	> .999
Fixed precision (*f*SD)	.642	> .999	.667	> .999
Variable precision (*v*SD)	.063	$1.80e^{-20}$	.056	$1.67e^{-23}$
Ordered precision (*o*SD)	.238	$7.30e^{-5}$	.222	$7.60\ e^{-5}$

Note. **Best model, * second best model, the *posterior* term refers to the probability of a model being preferred when a random participant is selected, the exceedance probability (P^*) refers to the confidence of a model's overall favourability.

Discussion

In this first experiment, I investigated the effect of object complexity on raw error, proportion of guesses and memory precision using a standard memory load manipulation with the two-component mixture model (Zhang & Luck, 2008). I also compared hypothesis-driven models with prior assumptions about the mixture model parameters across set sizes for both kinds of stimuli. These results revealed very similar patterns of results across all measures.

Examining raw error revealed that memory performance was similarly affected by load for both simple and complex objects. Critically, memory performance was worse overall for the complex objects. This finding is consistent with previous studies which have demonstrated that object complexity affects VSTM performance (Alvarez & Cavanagh, 2004; Delvenne & Bruyer, 2004; Rademaker et al., 2018; Wheeler & Treisman, 2002). I believe this is the first study to directly compare the effects of complexity in a recall task, without the stimuli being affected by familiarity and experience. Thus, I establish here that the effect of complexity on VSTM performance is not restricted to change-detection tasks, which could be affected by differences at the comparison stage (Awh et al., 2007; Barton et al., 2009; Brady & Störmer, 2020a), or relying on high-threshold responses. Importantly, there was no observed interaction between object complexity and set size. Thus, while both manipulations appear to impact cognitive load, they appear to do so independently. I also show that the ability to perform this VSTM task is strongly correlated across stimulus type, demonstrating a reliance on a shared process.

Using the two-component mixture-model devised by Zhang and Luck (2008), I also examined the effect of complexity on guesses and memory precision (SD). Here, the results differ slightly from those of raw error, as the observed guess rates show no significant effect of complexity, although the Bayesian test revealed nearly identical evidence to a model with main effects of both complexity and load. In contrast, both complexity and load affected SD, with a small but significant interaction driven by a less substantial increase in SD for lines between loads 1 and 2. I also compared the changes in guess rates and target SD simultaneously using Bayesian Model Comparison, which revealed that almost all participants exhibited an increase in guess rate from smaller to larger memory loads for both stimuli types. Thus, consistent with the standard mixture model analysis, guess rates increased across all tested loads for both simple and complex stimuli. In contrast, while the results from the standard mixture model revealed a significant main effect of load on response precision, the Bayesian Model Comparison revealed that most participants favored a model with a fixed precision across set sizes, whereas only a small group preferred an ordered precision parameter.

Although the mixture model results combined with the Bayesian Model Comparison are not unequivocal, they do reveal some important insights into VSTM for simple and complex objects. Overall, the results suggest that, although orientation memory is worse for complex 2.5D shapes than simple lines, this is not due to participants having a lower "capacity" for complex shapes, as the mixture model analysis revealed there was negligible evidence for an effect of complexity on guess rates. Similarly, a model comparison method revealed that for nearly all subjects, guess rates consistently increased across set sizes for both stimulus types, including from load one to two. By contrast, most fixed-capacity models suggest that guess rates should only increase beyond capacity (Zhang & Luck, 2008). Thus, although guess rates may overall be greater for the complex shapes, this is likely due to a reduction in encoding rather than a reduced capacity. Similarly, the mixture model analysis suggested that complex objects were remembered with less precision (more error), particularly at intermediate set sizes. While the Bayesian Model Comparison suggests that for most subjects' target SD remained relatively constant across loads, this was true for both simple and complex objects. In other words, although complex objects are recalled with less detail, precision is similarly affected by load for both simple and complex shapes. These results suggest that VSTM for complex 2.5D shapes is not categorically different from that of simple 2D lines.

It is important to note a few limitations from the current findings. First, caution must be taken in interpreting the mixture model results in general, as some have recently questioned the validity of the approach (Ma, 2018; van den Berg & Ma, 2018), in particular calling into question whether guesses are real responses. For example, these apparent responses could be the result of variability in response precision (Fougnie et al., 2012; van den Berg et al., 2012), or a result of simple signal-detection errors (Schurgin et al., 2018). Second, it is possible that with a more extensive set of set sizes different results may have been observed, as I only tested a limited range of set sizes here. Finally, it is possible that other models may be a better fit to the data than the ones tested here (Emrich et al., 2017; Hardman et al., 2017; Stengård & Berg, 2019; van den Berg et al., 2012, 2014). Thus, although these results do not definitively rule out potential differences in the manner in which simple and complex shapes are stored in memory, they overall suggest that VSTM for complex shapes shares many of the same properties as more simple surface features, albeit with worse overall encoding and recall.

Experiment 2

Experiment 1 demonstrated that although they are remembered with more error, short-term memory for complex shapes exhibits similar behaviour as compared with simple stimuli, although fewer complex objects may potentially be remembered with overall lower precision. However, it is possible that the difference in rotation span between the simple (i.e., 180°) and the complex stimuli (i.e., 360°) may create more room for error when reporting complex objects. Thus, it may be the case that the increased error rate for complex objects was mainly driven by this methodological difference. Moreover, there were two features that were confounded in our complex stimuli: namely, dimensionality and complexity. That is, while the 2.5D shapes contained more features, they also added depth cues, which may have affected performance. Finally, although memory performance between simple and complex objects was correlated, suggesting a shared underlying mechanism, it is unclear to what extent this relationship was observed due to the shared procedure (i.e., continuous recall for orientation) versus general shared memory resources. Consequently, the aim of Experiment 2 was to address these three issues.

Methods

Participants. A total of fifty-four (6 males; normal or corrected-to normal vision, 17-37 years of age, M = 21.49) Brock university students were recruited for experiment 2. Here again, participants were compensated with course credits or $10/hr and they were screened upon arrival for colour-blindness using Ishihara's 38 plates test. All procedures were approved by the Brock University Research Ethics Board.

Apparatus. The tasks were presented using MATLAB (mathworks.com) software along with the Psychophysics Toolbox extension (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) on a 20-in. flat LCD with a resolution of 1600-1200 ppi and a refresh rate of 60 Hz.

Stimuli and Procedure. In this second experiment, I used the same stimuli from experiment 1 (i.e., 2.5D shapes and simple line). The line stimulus was modified, however, so that one of the ends would always have a white square, which resembled a colored matchstick. This modification allowed for the line stimulus to span the full 360° instead of 180° and investigate whether the differences between the simple and complex stimuli from experiment 1 was due to this 180-360° span difference. I also included a 2D shape stimulus that was created from an arrangement of 9 squares with 2-4 bends. A total of 25 2D shapes were created and visually inspected to ensure that these objects were unique and distinguishable from one another. Participants first performed a continuous-recall task which is the same as set size 2 condition from Experiment 1 with the exception that participants can be presented with either the line/matchstick stimulus, 2D or 2.5D objects (see Figure 2-3A). Participants were presented with these stimuli in a block design, 40 trials per block, and participants completed 2 blocks for each stimulus type (i.e., total of 6 blocks, 80 trials per stimulus, and 240 trials in total). Participants were first trained on a short version of the task that cycled through 4 trials of each stimulus type and the order of the stimulus blocks was counter-balanced between participants.

After completion of the continuous-recall paradigm, participants were required to complete a change detection task that was designed to evaluate working memory capacity. Each trial presented a set of 2, 4, or 6 items for 250ms, which all varied in a key feature (e.g., colour or orientation) (see Figure 2-3B). After a retention period of 1,000ms, participants were presented with a single probe item, and they were required to determine whether the item was the same or different by pressing the “Z” or “?” keyboard keys respectively. The change-detection task consisted of 4 blocks of 16 trials per set size condition, where the probed item was changed on 8 trials, for a total of 192 trials (64 trials per set size). Participants were required to complete three change detection tasks for each of line, 2D and 2.5D stimulus taken from the previous task and the more seminal colored squares version (Luck & Vogel, 1997; Wheeler & Treisman, 2002). For a given trial, stimulus features were randomly selected without replacement from nine possible colors (i.e., red, blue, green, yellow, magenta, cyan, orange, white, and black) for the colored square version or orientations (i.e., 0°, 20°, 40°, …, 160°) for the line, 2D and 2.5D version. In the 2D and 2.5D versions, a single random stimulus was selected (without replacement) from their pool of 25 possible stimuli and a random orientation along the x- and y-axes was applied to the 2.5D object. Each stimulus in the display was also randomly assigned a position balancing the number of items across quadrants when possible. These changed-detection tasks were based on MATLAB code created by Adams (github.com/kcsa/change-detection-task). The order of the change detection tasks was counter-balanced across participants so that all twenty-four possible combinations were represented at least twice in the sample.

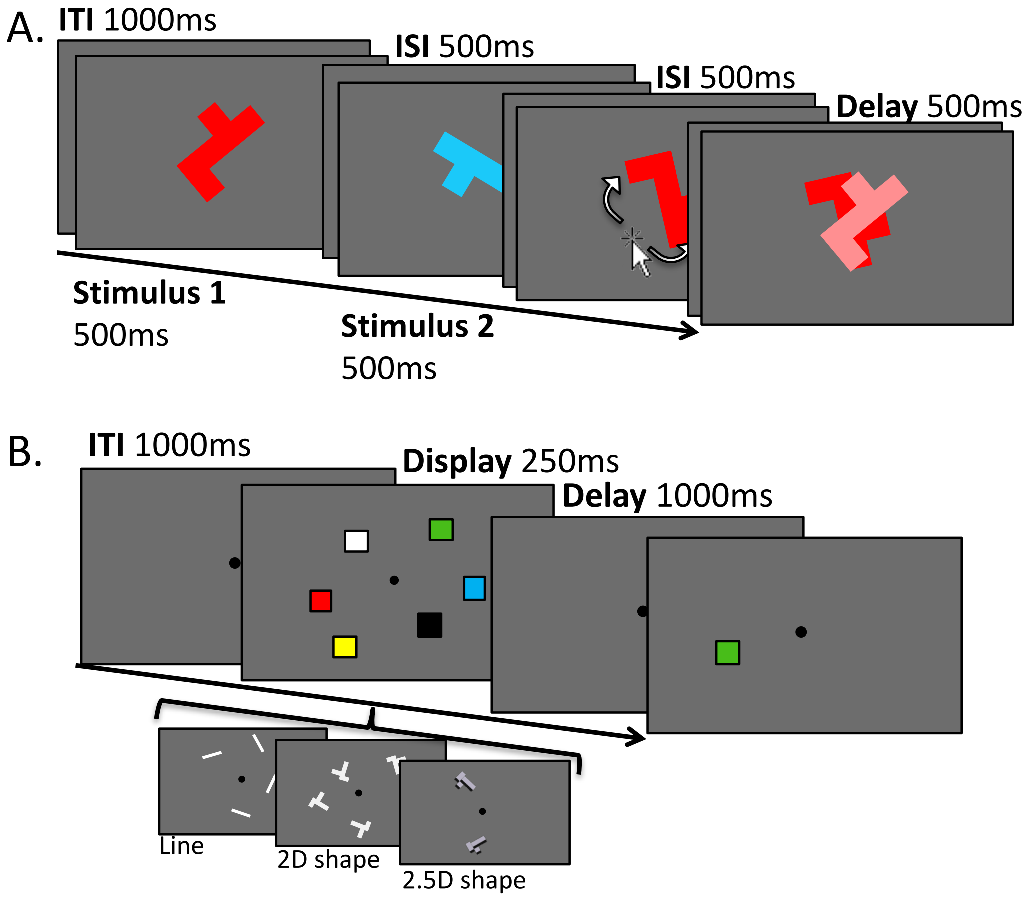

Figure 2-3. Task schematic of the working memory tasks used for Experiment 2. In the Continuous-report (A.) task, participants were presented two stimulus for 500ms each, separated by a 500 inter-stimulus interval. After a 500m rentention delay, one of the stimulus reappears and participants are required to report the orientation of the colored object as it was presented earlier in the trial. Participants were presented with either the matchstick, 2D shape (as in the schematic) or 2.5D objects. Note that, in each trial, particpants are presented with the same stimulus twice but they differ in orientation and color. (B.) In the change detection task, participants were presented with an array of 2, 4, or 6 items that varied in one feature (i.e., colors for the colored squares and orientation for the line, 2D and 2.5D objects.) for 250ms. After a 1000ms retention delay, participants were required to report whether one of the features of the stimuli that reappeared is the same or different from the array.

Analysis. For the continuous-response task, analyses of participants' raw error (SD) as well as the guess rate and precision parameters were derived from modeling the distribution of participants' responses using the two-components mixture model (Zhang & Luck, 2008), as described in Experiment 1. From data recorded from the change-detection tasks, I computed participants hits (h) and false alarms (f) each set size (N) and computed their maximum Cohen's capacity: $k = N(h - f)$, across set size. I then compared participants' performance for each stimulus type (i.e., line, 2D and 2.5D shapes) on raw error, guess rate, precision, and capacity using frequentist one-way repeated measures ANOVAs, with a Greenhouse-Geisser correction when the sphericity assumption is violated, Bonferroni corrected post-hoc t-tests comparisons, and repeated Bayesian ANOVA.

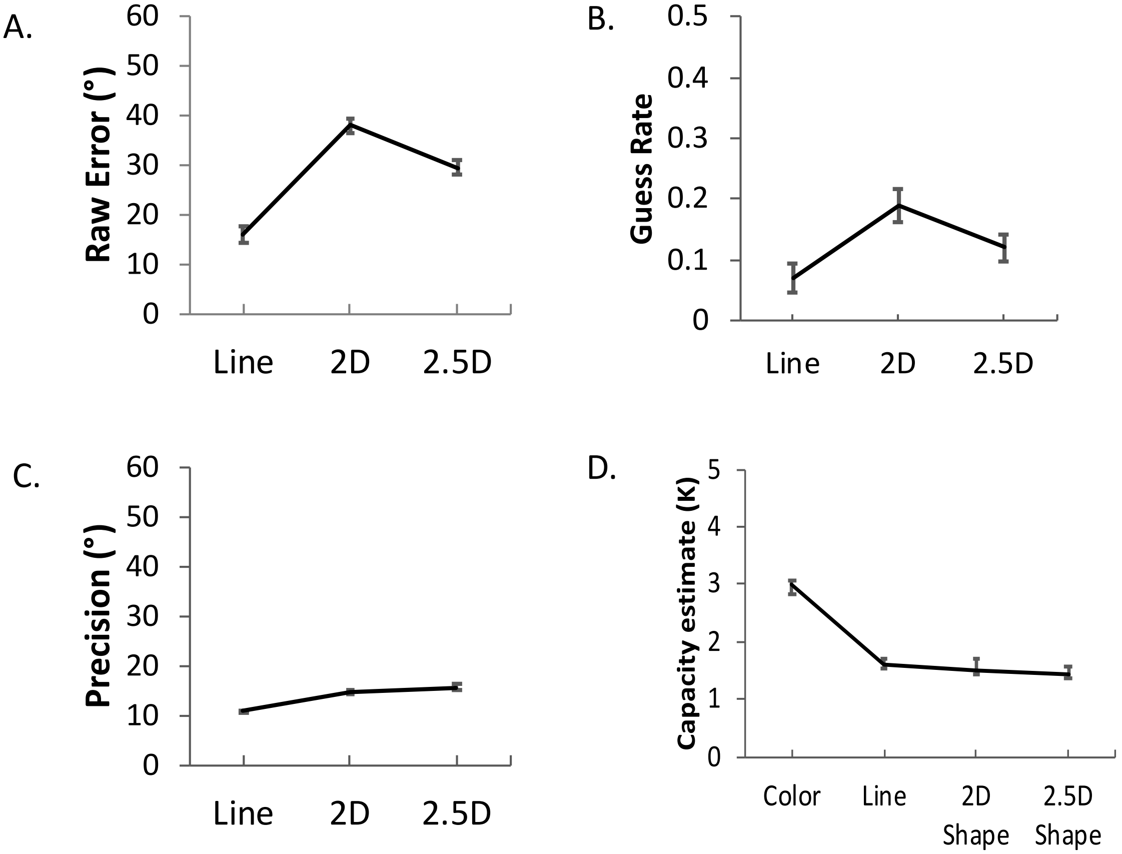

Figure 2-4. Experiment 2 behavioural results. Participants' (A.) mean raw error, (B.) guess rate, (C.) precision^{-1} (SD), and (D.) Cohen's capacity estimate (k) across stimulus type. Error bars represents standard error of the mean.

Results

The aim of the second Experiment was threefold: 1) to compare the 2.5D stimuli to line stimuli with a matched 360-degree stimulus space, 2) to separate the effect of dimensionality from complexity, and 3) to compare performance to capacity estimates on a change detection task. I used the same paradigm as set size 2 in Experiment 1 and presented a 360° spanning line stimulus, a 2D complex stimulus that exhibits no depth cue and the 2.5D objects from our previous experiment. Here again, I analyzed raw error and an estimate of guess rate and precision derived from Zhang and Luck's (2008) two-component mixture model. Finally, I also evaluated participants working memory capacity for colored squares as well as line, 2D and 2.5D stimulus orientation from a dichotomous changed-detection task.

Raw error. As a measure of behavioural performance, I evaluated participants' mean raw error (see Figure 2-3B) across stimulus type (i.e., line, 2D and 2.5D stimuli) using frequentist and Bayesian one-way repeated measures ANOVA, which revealed significant differences in stimulus type, $F(2, 102) = 106.75, p < .001, \eta^2 = 0.383; \log BF_M = 52.43$. Post-hoc t-test comparison with a Bonferroni correction shows a significant difference in reporting the orientation of line and 2D shapes, $M_{diff} = 22.11°, SE = 1.66°, p < .001; \log BF_{10} = 35.14$, and 2.5D shapes, $M_{diff} = 13.66°, SE = 1.41°, p < .001, ; \log BF_{10} = 23.8$, as well as between 2D and 2.5D shapes, $M_{diff} = 8.45°, SE = 1.51°, p < .001, ; \log BF_{10} = 9.88$.

Comparing 180° and 360° line performance. To determine whether there is a difference between the line stimulus from Experiment 1, which spans 180°, and the matchstick stimulus from Experiment 2, which spans 360°, I conducted a frequentist and

Bayesian independent sample t-test. Because the assumption of equal variance is violated, as indicated by Levene's test, $F = 4.546, p = .036$, a Welch's independent samples t-test was employed, and showed no statistical difference between the 180° ($M = 13.49, SD = 7.15$) and 360° span ($M = 15.88, SD = 12.19$) of the line orientation, $t(84.49) = 1.190, p = .237, d = 0.24$. A Bayesian independent sample t-test revealed moderate evidence in favor of the null hypothesis, $\log BF_{10} = -0.418$.

Mixture Model.

Guess rate. Next, I applied Zhang and Luck's (2008) two-components mixture model to each participant's raw error to obtain the estimated proportion of guesses and SD (precision^{-1}). Evaluating the proportion of guesses, a one-way repeated measures frequentist and Bayesian ANOVA reveals a significant effect of stimulus type, $F(1.44, 73.58) = 27.90, p < .001, \eta^2 = 0.072$; $\log BF_M = 17.48$. Frequentist and Bayesian post-hoc comparison revealed significant differences between stimulus type, where participants demonstrated fewer guesses when presented with lines in comparison to 2D shapes, $M_{diff} = 11.9\%, SE = 1.90\%,\ p < .001$; $\log BF_{10} = 11.62$, and 2.5D shapes, $M_{diff} = 5.0\%, SE = 1.0\%,\ p < .001; \log BF_{10} = 7.51$, as well as a greater proportion of guesses for 2D shapes than with 2.5D shapes, $M_{diff} = 6.90\%, SE = 1.70\%,\ p < .001,\ \log BF_{10} = 4.98$.

Target precision^{-1} (SD). Target precision, the second component of the mixture model, also demonstrated a significant difference between stimulus type as revealed by one-way repeated frequentist and Bayesian ANOVAs, $F(2, 102) = 46.84, p < .001, \eta^2 = 0.257$; $\log BF_{10} = 28.56$. Post-hoc frequentist t-tests with Bonferroni correction and Bayesian t-tests comparisons revealed better target precision when reporting lines in comparison to 2D shapes, $M_{diff} = 3.82°, SE = 0.502°,\ p < .001;\ \log BF_{10} = 16.78$, and 2.5D shapes, $M_{diff} = 4.83°, SE = 0.4881°,\ p < .001;\ \log BF_{10} = 24.50$, but no significant differences between the precision of 2D and 2.5D shapes, $M_{diff} = 1.01°, SE = 0.586°,\ p = .273,\ \log BF_{10} = -0.515$.

Working Memory Capacity. Finally, I evaluated participants working memory capacity (Cohen's k) for each of the stimulus type (i.e., line, 2D and 2.5D shapes) as well as a colored square stimulus using a single-probed design and set sizes 2, 4, and 6. Repeated measures frequentist and Bayesian ANOVAs revealed a significant capacity limit differences between stimulus type, $F(2.06,105.08) = 93.82, p < .001, \eta^2 = 0.443; \log BF_M = 17.46$. Post-hoc frequentist (Bonferroni corrected) and Bayesian t-test comparisons revealed that working memory capacity for colours is significantly (all $ps < .001$) greater than that of lines, $M_{diff} = 1.37, SE = 0.124;\ \log BF_{10} = 28.20$, 2D shapes, $M_{diff} = 1.48, SE = 0.125;\ \log BF_{10} = 30.51$, and 2.5D shapes, $M_{diff} = 1.54\%, SE = 0.139; \log BF_{10} = 28.33$. Conversely, there is no significant difference in memory capacity between line and 2D shapes, $M_{diff} = 0.109\ SE = 0.070,\ p = .760; \log BF_{10} = -0.766$, line and 2.5D shapes, $M_{diff} = 0.175, SE = 0.084,\ p = .245$, $\log BF_{10} = 0.125$, and between 2D and 2.5D shapes, $M_{diff} = 0.066, SE = 0.082,\ p > .999,\ \log BF_{10} = -1.580$.

Correlations. To investigate the association between working memory capacity and continuous recall performance across stimuli, I calculated Pearson's correlations between participants' raw error for line, 2D and 2.5D stimulus as well as capacity estimates for the same stimuli and color. Table 2-2 reveals that memory performance was moderately to strongly correlated across all measures, regardless of task or stimulus type.

In order to determine how much performance was related independent of baseline working memory capacity (K), I also ran a partial correlation controlling for K estimates from for colored squares. A meaningful amount of variance in participants' performance seems to be associated with their short-term memory capacity. However, I also observed that a greater portion of the variance in raw error is accounted for when capacity estimates for lines was accounted for, relative to colored squares. For example, while the partial correlation in recall error between 2D lines and 2D shapes remains significant when accounting for color capacity estimates, it is no longer significant once accounting for line capacity. To evaluate whether these changes in variance accounted for were significant, I conducted post-hoc comparisons of the correlation coefficients for dependent samples using an online calculator (psychometrica.de; Lenhard & Lenhard, 2014) based on Eid, Gollwitzer, & Schmitt (2011) procedure. I first applied a Fisher's Z transformation to the coefficients, and I evaluated whether there are significant differences between the Pearson's correlation and the two partial correlations (i.e., controlling for colour K and line K) for each type of stimulus pair (i.e., line-2D, line-2.5D, and 2D -2.5D). There was a marginal but non-significant difference between the correlations for the line and 2.5D, $z = 1.6, p = .055$, but no differences for line-2D, $z = 0.86, p = .195$, or 2D -2.5D, $z = 0.79, p = .213$. Despite showing a possible trend, I cannot conclusively demonstrate a significant decrease in variance accounted by orientation memory capacity.

Table 2-3.

Correlations and Partial Correlations between Raw Error (RE) and Memory Capacity (K).

	1. Line RE	**2. 2D RE**	**3. 2.5D RE**	**4. Color K**	**5. Line K**	**6. 2D K**
Correlation						
2. 2D RE	0.493***					
3. 2.5D RE	0.628***	0.548***				
4. Color K	-0.586***	-0.326*	-0.419**			
5. Line K	-0.708***	-0.491***	-0.599***	0.613***		
6. 2D K	-0.606***	-0.344*	-0.430**	0.540***	0.562***	
7. 2.5D K	-0.665***	-0.429**	-0.589***	0.608***	0.726***	0.716***
Partial correlation – controlling for Color K						
2. 2D RE	0.395**					
3. 2.5D RE	0.520***	0.479***				
Partial correlation – controlling for Line K						
2. 2D RE	0.238					
3. 2.5D RE	0.361**	0.364**				

Note: * $p < .05$, ** $p < .01$, *** $p < .001$, RE = Raw Error, K = Cohen's capacity estimate

Discussion

In this second experiment, I again investigated the effect of complexity on continuous recall performance, by comparing performance for stimuli matched in their degree of feature space (i.e., 360°). I also examined performance for a complex 2D stimuli that do not present monocular depth cues, thus disentangling the effects of complexity and dimensionality. Finally, I also examined performance for all stimuli in a change detection task, examining performance across measures. Our results generally replicate our findings from Experiment 1, where short-term memory performance is reduced for complex objects. In contrast to Experiment 1, however, I observed statistical differences in both guess rates and precision (1/SD) between lines and 2.5D shapes. Interestingly, participants' recall error was worse for complex 2D stimuli than 2.5D objects. This pattern suggests that, although more complex, the dimensionality cues present in the 2.5D stimuli facilitated participants recall performance in comparison to similar 2D stimuli. The same pattern is also true for guess rate estimates, suggesting a decreased ability to encode these complex objects without dimensionality cues, while retaining similar recall precision. These results support the idea that dimensionality facilitates the encoding process and makes complex objects easier to remember, while still negatively affecting the quality of those stored representations.

I also examined performance in a change-detection task, obtaining capacity estimates (K) for all stimuli, as well as the frequently used colored squares. Although participants generally demonstrate higher capacity estimates for colored squares, there were no significant capacity difference for the orientation of line, 2D and 2.5D stimuli. The absence of differences in change-detection *K* estimates despite differences in continuous-recall performance suggests that continuous recall may be a more sensitive measure of assessing VSTM performance. Moreover, the difference in capacity estimates between colored squares and the three other stimuli suggest that there may be some feature-specific differences in VSTM abilities. That is, memory for orientations may have some shared processes that are distinct from those of color VSTM.

The correlational analysis supports this conclusion. Overall, the measures underlying short-term memory performance and capacity across all stimuli were strongly inter-correlated. Interestingly, correlations remained significant when controlling for short-term memory capacity for colored squares, suggesting that the variance accounted for here is likely to be related to feature-specific working memory mechanisms. In contrast, there was a substantial decrease in the strength of the correlations in raw error between stimulus type when controlling for capacity estimates for lines, albeit not significantly different from one another. These results support the finding by Wheeler and Treisman (2002) that short-term memory competes along feature-specific dimensions, namely, where orientation is the critical feature common to these stimuli. Thus, the strong overlap between these measures may be a result of orientation-related variance, rather than stimulus-related. One way to investigate this potential source of variance is to investigate the effect of the axis of rotation since thus far, all of our stimuli varied in orientation with respect to the z-axis.

Experiment 3

The strong correlation between one's ability to report the orientation of our simple, 2d and 2.5D stimuli is perhaps not surprising given that our tasks involved rotating different objects over the same dimension (i.e., around the z-axis). Working with

2.5D objects allows for rotation over additional planes. To our knowledge, however, it has not been previously investigated whether memory for the features of visual objects is represented with similar precision across all dimensions. Consequently, the aim of Experiment 3 was to examine VSTM for complex objects in a continuous response task over multiple dimensions.

Methods

Participants. Forty-eight (7 males; normal or corrected-to normal vision, 18-32 years of age, M = 19.62) Brock university students were recruited for the third experiment. As in Experiment 1 and 2, participants were compensated with course credits or $10/hr, and they were screened for colour-blindness upon arrival.

Stimuli & procedure. The third experiment is the same as Experiment 1 with the following exceptions: participants are required to remember the orientation of a single item per trial and the complex object could be manipulated over one of three different axes of rotation (i.e., x-, y-, or z-axis; see Figure 2-3A). Participants completed 80 trials for each 3D orientation (i.e., x-, y-, or z-axis), for a total of 240 intermixed trials. Since a 2D line can only be manipulated over the z-axis, participants were required to complete only 80 trials. To reduce the primacy effect, 24 participants began with simple 2D line short-term memory task first followed by the complex object while the other 24 participants began with the complex shapes first.

Following the short-term memory tasks, participants were required to complete a computerized version of Sheppard and Metzler (2008) imagery task and the VVIQ-rv and OSVIQ. This task and two questionnaires were part of a different project and will not be reported here.

Analysis. Here again, I analyzed participants' raw error and modelled response error using Zhang & Luck's two components mixture model to obtain an estimated proportion of guesses and target precision. In the first set of comparisons, I evaluate the error between simple lines rotated over the z-axis against all axis rotations of the complex shapes using frequentist and Bayesian t-tests. Then, I compared the different axes of rotations within the complex conditions using a one-way repeated measure frequentist ANOVA, where Greenhouse-Geisser correction is applied when the sphericity assumption is violated. Here again, all statistical analyses were computed using JASP (JASP Team, 2018) with the same statistical parameters (e.g., corrections, prior distribution, etc.) as in our previous experiments.

Results

In this experiment, I investigated the properties of VSTM representations of complex objects rotated over different 3D axes (x-, y-, and z-axes) and compared them with a simple 2D line rotation. Here again, I used raw error as a measure of participant's performance. In addition, I applied the two-component mixture model (Zhang & Luck, 2008) to obtain the proportion of guesses separately from the precision of correct responses. Note that the manipulation over the z-axis for both the simple and complex shape is a direct replication of the set size 1 condition in Experiment 1.

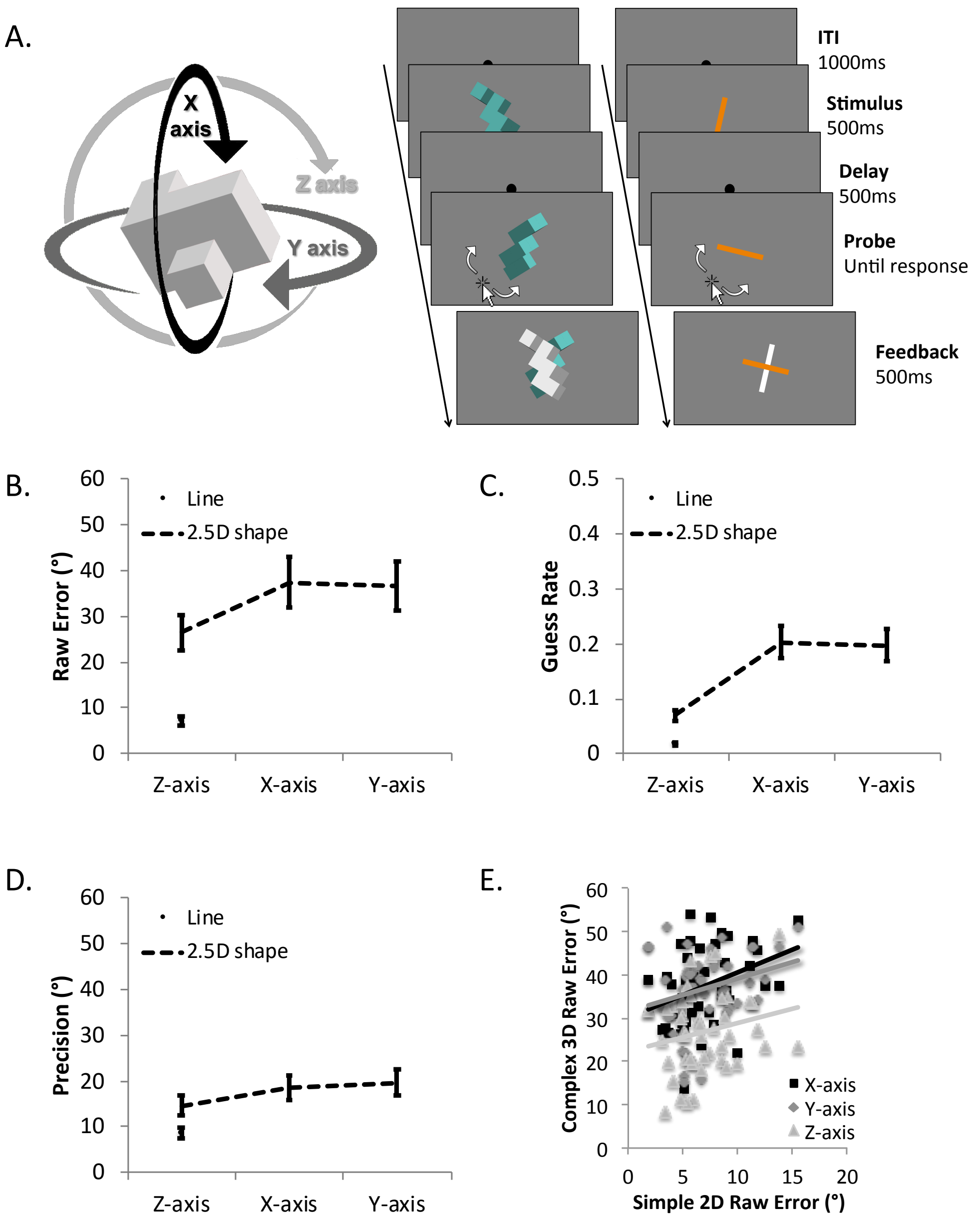

Figure 2-5. Task schematic and behavioural results of Experiment 3. (A) A schematic representation of all three axes of 3D rotations and the delayed-response task used in experiment 3, which was exactly the same as the set size 1 condition in experiment 1 with the exception that the complex 3D object can be rotated over the x-, y-, or z-axes. Simple 2D shape orientation was only rotated over the z-axis. (B) Participants' mean raw error, (C) guess rate, and (D) precision^{-1} (SD) comparison between the report of simple line and complex 3D object orientation across axis rotation. Error bars represent the standard error of the mean.

Raw error. I first examined raw error as a measure of behavioural performance (see Figure 2-5B). To evaluate the memory performance for simple objects in comparison to complex objects, I computed frequentist and Bayesian paired t-tests. Significant differences were found across all three comparisons demonstrating a lower SD of responses for line orientations than complex objects rotated over the same axis (i.e., z-axis), $M_{diff} = 19.94°, SE = 1.33°, t(47) = 14.58, p < .001, d = 2.13, \log BF_{10} = 36.46$, as well as over the x-axis, $M_{diff} = 30.31°, SE = 1.22°, t(47) = 24.83, p < .001, d = 3.62, \log BF_{10} = 57.60$, and y-axis, $M_{diff} = 29.54°, SE = 1.19°, t(47) = 24.90, p < .001, d = 3.63, \log BF_{10} = 57.72$. These results replicate the cost associated with complexity from Experiments 1 and 2.

To compare the raw error in reporting the orientation the complex shapes over the x-, y-, and z-axes, one-way-within subject frequentist and Bayesian ANOVAs were conducted, revealing a significant difference between conditions, $F(1.68, 77.48) = 39.51, p < .001, \eta^2 = 0.240, \log BF_M = 23.91$. Post-hoc comparison using frequentist t-tests with a Bonferonni correction and Bayesian t-test revealed significant differences between complex objects rotated over the z-axis and both the x-axis, $M_{diff} = 10.93°, SE = 1.54°, p < .001, \log BF_{10} = 14.47$, and the y-axis, $M_{diff} = 10.16°, SE = 1.49°, p < .001, \log BF_{10} = 13.60$. No significant difference between the shape rotated over the x- and y-axes was found, $M_{diff} = 0.773°, SE = 1.04°, p > .999, \log BF_{10} = -1.581$.

In order to determine whether there is a relationship between one's ability to remember the simple object orientation and complex shapes across the z-axis (see results from Experiment 1 and Figure 2-1E for x-axis aggregated correlation) and the x- and y-axes (see Figure 2-5E), pairwise Pearson's correlations and Bayes correlation factors were computed on participant's mean raw error. Although the relationship between the raw error for simple and complex items over the x-axes shows a significant but small correlation, $r = .397, n = 48, p = .006$, where 15.8% of the variance is accounted for, Bayesian factor shows that there is only anecdotal evidence for the alternative hypothesis over the null, $\log BF_{10} = 2.00$. There is also a weak but significant relationship between the item and shape rotated over the y-axis, $r = .340, n = 48, p = .019$, and anecdotal evidence for either alternative hypothesis, $\log BF_{10} = 0.952$.

Turning to the inter-relationship between 3D rotations, significant correlations were found between the z-axis and both the x-axis, $r = .331, n = 48, p = .023$, and the y-axis (see Figure 2-2G), $r = .345, n = 48, p = .018$, and anecdotal evidence for the alternative hypothesis, respectively: $\log BF_{10} = 0.797$; $\log BF_{10} = 1.026$. A moderate correlation was found between the in the report of the orientation of complex shapes rotated across the x- and y-axes, $r = .681, n = 48, p < .001$, and Bayesian factor reveals very strong evidence towards the alternative hypothesis ($\log BF_{10} = 11.72$). These results demonstrate that one's ability to remember the orientation of complex shapes over 3D axes are strongly interrelated, particularly when the object is affected by the same occlusion constraints (i.e., rotation over the x- and y- axes).

Mixture Model

Proportion of guesses. Next, I examined the effect of 3D rotation on the proportion of guesses (see Figure 2-5C). Frequentist and Bayesian t-test comparisons revealed a decreased proportion of guesses in reporting simple object orientation compared to complex objects rotated over the z-, x-, and y-axes, Z-axis: $M_{diff} = 5.1\%, SE = 0.8\%, t(47) = 6.21, p < .001, d = 0.906, \log BF_{10} = 11.56$; X-axis: $M_{diff} = 18.4\%, SE = 1.5\%, t(47) = 12.11, p < .001, d = 1.77, \log BF_{10} = 29.92$; Y-axis: $M_{diff} = 17.9\%, SE = 1.4\%, t(47) = 12.61, p < .001, d = 1.84, \log BF_{10} =$

31.31. Here again, I have evidence that of a decrease in memory performance related to the complexity of the object memorized. Frequentist and Bayesian repeated measures ANOVA evaluating the proportion of guesses between 3D rotations for complex shape reveal significant differences, $F(2{,}92) = 65.53, p < .001, \eta^2 = 0.314;\ \log BF_{10} = 35.75$, where post-hoc comparisons reveal lower guess rate for the complex shape rotated over the z-axis compared to the x- and y- axes, X-axis: $M_{diff} = 13.3\%, SE = 1.5\%, p < .001; \log BF_{10} = 20.56$; Y-axis: $M_{diff} = 12.6\%, SE = 1.3\%,\ p < .001; \log BF_{10} = 22.27$, but not between the x- and y-axes, $M_{diff} = 0.60\%, SE = 1.1\%, p > .999;\ \log BF_{10} = -1.72$.

Target precision-1 (SD). Finally, I evaluated the differences in response precision between the VSTM of simple object as it compares to complex object rotated over the three possible axes using frequentist and Bayesian t-test (see Figure 2-5D). Participants showed greater precision^{-1} (SD) reporting simple objects compared to all 3 rotations of the complex objects, Z-axis: $M_{diff} = 6.08°, SE = 0.38°, t(47) = 15.87, p < .001, d = 2.31, \log BF_{10}\, 39.66$; X-axis: $M_{diff} = 9.95°, SE = 0.72°, t(47) = 13.84, p < .001, d = 2.02, \log BF_{10}\, 34.60$; Y-axis: $M_{diff} = 10.98°, SE = 0.70°, t(47) = 15.68, p < .001, d = 2.39, \log BF_{10} = 39.19$. Here again, I reveal the same pattern exhibited by the previous measures, where a strong cost associated with complexity is present and decreases performance on the precision of the memorized item.

Repeated measures frequentist and Bayesian ANOVA revealed significant precision differences in reporting the orientation of the complex object across the three axes of rotations, $F(2{,}92) = 28.32, p < .001, \eta^2 = 0.153;\ \log BF_{10} = 17.40$. Post-hoc comparison using frequentist and Bayesian t-tests reveals greater precision for the complex shape rotated over the z-axis compared to the x- and y- axes, X-axis: $M_{diff} = 3.87°, SE = 0.68°, p < .001, d = .827;\ \log BF_{10} = 9.82$; Y-axis: $M_{diff} = 4.90°, SE = 0.72°, p < .001, d = 0.988;\ \log BF_{10} = 13.39$, but not between the x- and y-axes, $M_{diff} = 1.03, SE = 0.65°, p = .363, d = 0.23;\ \log BF_{10} = -0.689$. The precision of the memorized items rotated over the x- and y-axes are greatly reduced in comparison to rotation over the z-axis. This pattern mirrors those of other two behavioural measures of raw error and Mixture Model guess rates.

Discussion

Experiment 3 takes advantage of the nature of the 2.5D objects to investigate the effect of axis of rotation on dimensionality. Here, complex stimuli are manipulated over all three axes and compared to line stimuli rotated over the z-axis only. Looking at the performance over the z-axis, participants exhibited lower raw error when reporting the orientation of simple stimuli compared to the complex shapes. As with Experiment 2, this effect of complexity on performance is also consistently extended to the mixture model parameters of precision and guess rate. Here again, these results indicate a cognitive cost associated with complex stimuli when performing a short-term memory recall task. Participants also exhibited a systematic reduced performance on raw error, proportion of guesses and precision when the complex object was rotated over the x- and y-axes relative to the z-axis, whereas there is no difference between rotation over the x- and y-axes across all three measures. A potential explanation for these results comes from the partial occlusion of the object, which may render the task of recognizing the object as a whole object more difficult. A shape rotated over the z-axis will present all of its relevant features regardless of its orientation, whereas objects rotated over the x- or y-axes will have a significant portion of features occluded by the most forward features. Thus, the recall task enables participants to eventually view all sides of the figure, these results suggest greater difficulty to encode and report accurately the orientation of an object when one can encode a small portion of its key features, and when those features change as a function of the recall angle.

The pattern of correlations found across conditions in Experiment 3, along with the psychophysical similarities, again highlights the involvement of shared processes for VSTM across the simple and complex shapes. The strong correlation between raw error in the rotation of complex objects over the x- and y-axes suggests a considerable overlap between the underlying mechanisms relative to the weaker relationship with the simple line or 2.5D shapes rotated over the z-axis. However, the finding that the effect of shape complexity (i.e., simple vs. complex rotated over the z-axis) has a greater effect on VSTM performance than partial occlusion (i.e., complex rotated over the z-axis vs. x- and y-axes) suggests that, while the axis of orientation may fundamentally affect the underlying VSTM abilities, the nature of the stimuli remains of fundamental importance to VSTM performance, even when all the features are visible.

General Discussion

The primary aim of this study was to evaluate the generalizability of VSTM from simple features (i.e., orientations) to more complex objects using a continuous-recall paradigm. As such I evaluated the accuracy of reporting simple line compared to a 2.5D objects, manipulating set size in Experiment 1, dimensionality in Experiment 2, and the axis of rotation in Experiment 3. The results of all three experiments demonstrated costs in participants' recall performance (i.e. reduced raw error) associated with the memory for complex objects relative to simple 2D lines, a finding that replicates the effects previously reported mainly from dichotomous paradigms (Alvarez & Cavanagh, 2004; Delvenne & Bruyer, 2004; Rademaker et al., 2018; Wheeler & Treisman, 2002). Thus, I provide evidence that this effect is not limited to change-detection tasks. Importantly, this method also revealed differences in recall performance between different complex stimuli that was not observed in a change detection task (Experiment 2). Thus, consistent with previous findings that continuous-recall tasks provide a more sensitive measure of VSTM performance (Adam et al., 2019; Zokaei et al., 2014), our results confirm that the continuous recall approach may reveal novel insights into the effects of complexity on VSTM.

By modeling participants' responses using a two-component mixture model, and using Bayesian Model Comparison, Experiment 1 also demonstrated that memory load has a similar effect on guess rates and target precision (SD) for both simple and complex objects. Importantly, although there was moderate evidence that guess rates were overall greater for complex objects (as confirmed by Experiments 2 and 3), there was no evidence of an interaction between complexity and load. Indeed, evidence from the BMS revealed that for most participants, guess rates increased as a function of the number of items to be remembered, even at low loads (e.g., 1 -2 items). The lack of interaction suggests that although complex objects are more difficult to encode and remember, this is not due to a reduced capacity for complex objects. The quality (i.e., precision of responses) of memory representations was also affected by stimulus complexity largely independent of set size, although complex objects were remembered with overall lower precision. Thus, while Experiment 1 was not designed to fully adjudicate between different models of VSTM, the results do point to categorically similar memory for simple and complex objects.

Although a number of studies have previously investigated VSTM for complex stimuli using continuous recall, these studies have typically relied on more naturalistic or familiar object, such as faces (Curby & Gauthier, 2007, Zhou, Mondloch, & Emrich, 2018, and Jiang, Shim, & Makovski, 2008). In contrast, I manipulated complexity through the dimensionality of the stimulus. While there are some evidence suggesting a “real world” memory advantage where more realistic objects are better remembered (Snow et al., 2014), which parallels the effect of the stimulus familiarity, the underlying mechanism mitigating this phenomenon is scarcely investigated. Results from Experiment 2 suggest that the effect of dimensionality on memory performance is not necessarily a straightforward effect; indeed, although by altering dimensionality from complex 2D to 2.5D objects, one also increases the features/information present in each stimulus, participants’ performance suggests that the 2.5D objects were relatively easier to report than the complex 2D objects. Thus, despite being more complex, the 2.5D objects may create a more feature-rich representation that helps supports better memory.

Experiment 2 also provides evidence for separate memory capacity for object color and orientation, as previously suggested by Wheeler and Treisman (2002). That is, although memory performance was strongly correlated across all measures and stimuli, capacity (K) for 2D lines explained more of the relationship between recall measures of 2D and 3D shapes than did color capacity. It is important to note that, while there is more variance accounted for by 2D line capacity, comparisons between correlations did not significant yield significant differences. The finding that there may exist some feature-specific memory stores may be consistent with sensory-recruitment models of VSTM. That is, there is a growing literature indicating that VSTM is not a unitary phenomenon affected only by the number of stimuli that have to be stored. Indeed, according to the sensory recruitment hypothesis, VSTM maintenance is sustained through the continuous recruitment of the same cortical regions that were initially required for the encoding of the information (Ester et al., 2009; Postle, 2006; Serences, 2016; Serences et al., 2009). For instance, Druzgal & D'Esposito (2001) have demonstrated the involvement of the fusiform face area (FFA) during the maintenance of face stimuli in a working memory task. Similarly, studies have demonstrated that decoding fMRI BOLD activity recorded in V1 can reliably identify the orientation of grating patterns (Albers et al., 2013; Harrison & Tong, 2009). Moreover, activity in these sensory regions (e.g., V1 and MT) is predictive of the precision of memory recall (Emrich et al., 2013; Sahan et al., 2019). Consequently, it is possible that whereas both simple lines and complex 2.5D shapes may rely on activity in the orientation-tuned neurons in primary visual cortex, committing a complex 2.5D-shape to memory may recruit more anterior regions of the ventral visual stream (Christophel et al., 2017). For example, familiar unscrambled objects recruit distinct patterns of activation in ventral temporal cortex during the maintenance period than do highly scrambled versions of the same objects (Stojanoski et al., 2020). Moreover, studies have shown that 3D shapes are processed in areas of the dorsal visual stream (Króliczak et al., 2008; Theys et al., 2015). Thus, although VSTM for complex objects may share more variance with memory for simple lines than with color, they may differ in that memory for 3D shapes requires the recruitment of large networks of visual regions not required for simple 2D line orientations. This may have important implications for models of VSTM, as manipulating stimulus complexity may necessitate the recruitment of higher visual areas, despite sharing a common, relevant feature (i.e., orientation).

Alternatively, a recent study has demonstrated that colors benefit from the configural processing afforded by the simultaneous presentation used in change detection, whereas real-world complex objects benefit from the deeper processing afforded by sequential presentation as performed in the continuous recall task use here (Brady & Störmer, 2020b). Thus, the "capacity" measures obtained in the change detection tasks and the measures obtained in the continuous recall tasks may reflect distinct processing of the stimuli, rather than some equivalent measure of performance. Thus, the fact that different measures affect the processing of simple and complex stimuli differently further emphasizes the importance of developing stimuli and procedures that allow for fair comparisons in order to properly assess memory mechanisms (Brady & Störmer, 2020a). Finally, Experiment 3 investigated participant's ability to report complex objects rotated over different axes of rotation (i.e., x-, y-, and x-axis), as most VSTM studies have investigated orientation of objects rotated around the z-axis only. I observed that memory for 2.5D shapes is not consistent across all axes of rotation, with all measures of accuracy indicating better performance around the z-axis than over the x- and y-axes. To our knowledge, these results provide the first demonstration of the psychophysical properties of VSTM for 2.5D shapes in a recall task, as well as demonstrating the additional challenges faced for remembering 2.5D shapes in space.

Overall, the results reveal that while storing complex objects is more difficult than storing simple ones, adding monocular depth cues to an object helps memory, so long as memory for orientation is around the z-axis. A possible reason for this decrease in recall performance may pertain to how participants encode the target feature. Storing the orientation of an object as a Euclidean vector is perhaps the most effective way to complete this VSTM task. As such, converting a 2D line into a vector is fairly straightforward process; the additional features on the more complex objects may make it more difficult to extract a meaningful orientation vector. As shown in Experiment 2, presenting a complex object with meaningful 2.5-dimensional cues may help to better represent the orientation information in a 2D space. The same cues, however, won't provide the same benefits when the depth of the features is affected by the orientation, as demonstrated in Experiment 3. Interference can still occur when an object changes it orientation over the x- and y-axes, as not only are the forward features partially obstructing parts of the object, but also the orientation vectors would vary along only one dimension (i.e., left to right around the y-axis; up and down around the x-axis). Moreover, some evidence suggests that, while making stimuli more tangible increases recall performance (Snow et al., 2014), working in a 3D space, however, can create confusion about an object's spatial reference to one another (Cockburn & McKenzie, 2002; Wickens et al., 1996). Although the complex objects are always depicting depth across all experiments, where only in Experiment 3 the features are moving about the depth space, the absence of any spatial references or context creates greater confusion about the object's true orientation.

In conclusion, while memory performance for complex stimuli is generally worse than for simple 2D stimuli, as previously demonstrated using dichotomous tasks, by adopting a continuous recall paradigm, I was able to demonstrate that the orientation of complex objects is primarily remembered with less precision than simple stimuli, but with some issue of memory accessibility (i.e., greater guess rate), which is more prominent in Experiment 2 and 3. These performance errors are shown to work independently from one's memory capacity (as measured by change detection) but do rely heavily on a shared processes. While there is evidence for a general VSTM mechanism that overlaps simple and complex stimuli, as demonstrated by the repeated-measures correlation in Experiment 1 and the inter-correlation between variables in Experiment 2 and 3, there is also evidence for specific, and perhaps independent, mechanisms at play. Indeed, our results from Experiment 2 suggest that the feature dimension (i.e., color or orientation) and the axis of rotation from Experiment 3 exhibit unique processes for to VSTM. The nature of these shared mechanisms, however, is still up for investigation. Furthermore, despite increasing the featural complexity of a stimulus to be remembered, providing meaningful features such as monocular depth cues can facilitate one's recall ability, given that the object is not rotated freely in a 3D space (i.e., not limited to rotation around the z-axis). The present study demonstrates that modeling the responses obtained from a continuous recall paradigm provide greater insights in the psychometrical properties of VSTM and how it is affected by complexity, dimensionality, or axes of rotation, which can only be superficially addressed using a dichotomous paradigm.

References

Adam, K. C. S., Doss, M. K., Pabon, E., Vogel, E. K., & Wit, H. de. (2019). Δ9-Tetrahydrocannabinol (THC) impairs visual working memory performance. BioRxiv, 778068. https://doi.org/10.1101/778068

Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & de Lange, F. P. (2013). Shared Representations for Working Memory and Mental Imagery in Early Visual Cortex. Current Biology, 23(15), 1427–1431. https://doi.org/10.1016/j.cub.2013.05.065

Alvarez, G. A., & Cavanagh, P. (2004). The Capacity of Visual Short-Term Memory is Set Both by Visual Information Load and by Number of Objects. Psychological Science, 15(2), 106–111. https://doi.org/10.1111/j.0963-7214.2004.01502006.x

Awh, E., Barton, B., & Vogel, E. (2007). Visual working memory represents a fixed number of items regardless of complexity. Journal of Vision, 7(9), 352–352. https://doi.org/10.1167/7.9.352

Barton, B., Ester, E. F., & Awh, E. (2009). Discrete resource allocation in visual working memory. Journal of Experimental Psychology. Human Perception and Performance, 35(5), 1359–1367. https://doi.org/10.1037/a0015792

Blajenkova, O., Kozhevnikov, M., & Motes, M. A. (2006). Object-spatial imagery: A new self-report imagery questionnaire. Applied Cognitive Psychology, 20(2), 239–263. https://doi.org/10.1002/acp.1182

Brady, T., & Störmer, V. S. (2020a). Comparing memory capacity across stimuli requires maximally dissimilar foils: Using deep convolutional neural networks to understand visual working memory capacity for real-world objects. PsyArXiv. https://doi.org/10.31234/osf.io/25t76

Brady, T., & Störmer, V. S. (2020b). The role of meaning in visual working memory: Real-world objects, but not simple features, benefit from deeper processing. PsyArXiv. https://doi.org/10.31234/osf.io/kzvdg

Brady, T., Störmer, V. S., & Alvarez, G. A. (2016). Working memory is not fixed-capacity: More active storage capacity for real-world objects than for simple stimuli. Proceedings of the National Academy of Sciences of the United States of America, 113(27), 7459–7464. https://doi.org/10.1073/pnas.1520027113

Brainard, D. H. (1997). The Psychophysics Toolbox. Spatial Vision, 10(4), 433–436.

Cockburn, A., & McKenzie, B. (2002). Evaluating the effectiveness of spatial memory in 2D and 3D physical and virtual environments. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 203–210. https://doi.org/10.1145/503376.503413

Delvenne, J.-F., & Bruyer, R. (2004). Does visual short‐term memory store bound features? Visual Cognition, 11(1), 1–27. https://doi.org/10.1080/13506280344000167

Dowd, E. W., Kiyonaga, A., Beck, J. M., & Egner, T. (2015). Quality and accessibility of visual working memory during cognitive control of attentional guidance: A Bayesian model comparison approach. Visual Cognition, 23(3), 337–356. https://doi.org/10.1080/13506285.2014.1003631

Druzgal, T. J., & D'Esposito, M. (2001). Activity in fusiform face area modulated as a function of working memory load11Published on the World Wide Web on 5 October 2000. Cognitive Brain Research, 10(3), 355–364. https://doi.org/10.1016/S0926-6410(00)00056-2

Eid, M., Gollwitzer, M., & Schmitt, M. (2011). Statistik und Forschungsmethoden: Lehrbuch ; mit Online-Materialien. Beltz.

Emrich, S. M., Lockhart, H. A., & Al-Aidroos, N. (2017). Attention mediates the flexible allocation of visual working memory resources. Journal of Experimental Psychology. Human Perception and Performance, 43(7), 1454–1465. https://doi.org/10.1037/xhp0000398

Emrich, S. M., Riggall, A. C., Larocque, J. J., & Postle, B. R. (2013). Distributed patterns of activity in sensory cortex reflect the precision of multiple items maintained in visual short-term memory. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 33(15), 6516–6523. https://doi.org/10.1523/JNEUROSCI.5732-12.2013

Ester, E. F., Serences, J., & Awh, E. (2009). Global mechanisms of sensory recruitment during working memory maintenance. Journal of Vision, 9(8), 597–597. https://doi.org/10.1167/9.8.597

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

Gorgoraptis, N., Catalao, R. F. G., Bays, P. M., & Husain, M. (2011). Dynamic Updating of Working Memory Resources for Visual Objects. The Journal of Neuroscience, 31(23), 8502–8511. https://doi.org/10.1523/JNEUROSCI.0208-11.2011

Hardman, K. O., Vergauwe, E., & Ricker, T. J. (2017). Categorical working memory representations are used in delayed estimation of continuous colors. Journal of Experimental Psychology: Human Perception and Performance, 43(1), 30–54. https://doi.org/10.1037/xhp0000290

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. Nature, 458(7238), 632–635. https://doi.org/10.1038/nature07832

JASP Team. (2018). JASP (0.9) [Computer software]. https://jasp-stats.org/

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. Perception, 36(14), 1–16.

Króliczak, G., McAdam, T. D., Quinlan, D. J., & Culham, J. C. (2008). The human dorsal stream adapts to real actions and 3D shape processing: A functional magnetic resonance imaging study. Journal of Neurophysiology, 100(5), 2627–2639. https://doi.org/10.1152/jn.01376.2007

Lenhard, W., & Lenhard, A. (2014). Hypothesis tests for comparing correlations. Bibergau: Psychometrica. https://www.psychometrica.de/correlation.html

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. Nature, 390(6657), 279–281. https://doi.org/10.1038/36846

Ma, W. J. (2018). Problematic usage of the Zhang and Luck mixture model. BioRxiv, 268961. https://doi.org/10.1101/268961

Marks, D. F. (1973). Visual imagery differences in the recall of pictures. British Journal of Psychology (London, England: 1953), 64(1), 17–24.

McKelvie, S. J. (1995). The VVIQ as a psychometric test of individual differences in visual imagery vividness: A critical quantitative review and plea for direction. Journal of Mental Imagery, 19(3–4), 1–106.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. Spatial Vision, 10(4), 437–442.

Postle, B. R. (2006). Working Memory as an Emergent Property of the Mind and Brain. Neuroscience, 139(1), 23–38. https://doi.org/10.1016/j.neuroscience.2005.06.005

Rademaker, R. L., Park, Y. E., Sack, A. T., & Tong, F. (2018). Evidence of gradual loss of precision for simple features and complex objects in visual working memory. Journal of Experimental Psychology. Human Perception and Performance, 44(6), 925–940. https://doi.org/10.1037/xhp0000491

Sahan, M. I., Sheldon, A. D., & Postle, B. R. (2019). The Neural Consequences of Attentional Prioritization of Internal Representations in Visual Working Memory. Journal of Cognitive Neuroscience, 1–28. https://doi.org/10.1162/jocn_a_01517

Schurgin, M. W., Wixted, J. T., & Brady, T. (2018). Psychophysical Scaling Reveals a Unified Theory of Visual Memory Strength. BioRxiv, 325472. https://doi.org/10.1101/325472

Serences, J. T. (2016). Neural mechanisms of information storage in visual short-term memory. Vision Research, 128, 53–67. https://doi.org/10.1016/j.visres.2016.09.010

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

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. Science (New York, N.Y.), 171(3972), 701–703.

Snow, J. C., Skiba, R. M., Coleman, T. L., & Berryhill, M. E. (2014). Real-world objects are more memorable than photographs of objects. Frontiers in Human Neuroscience, 8. https://doi.org/10.3389/fnhum.2014.00837

Stengård, E., & Berg, R. van den. (2019). Imperfect Bayesian inference in visual perception. BioRxiv, 402776. https://doi.org/10.1101/402776

Suchow, J. W., Brady, T., Fougnie, D., & Alvarez, G. A. (2013). Modeling visual working memory with the MemToolbox. Journal of Vision, 13(10). https://doi.org/10.1167/13.10.9

Theys, T., Romero, M. C., van Loon, J., & Janssen, P. (2015). Shape representations in the primate dorsal visual stream. Frontiers in Computational Neuroscience, 9. https://doi.org/10.3389/fncom.2015.00043

van den Berg, R., Awh, E., & Ma, W. J. (2014). Factorial Comparison of Working Memory Models. Psychological Review, 121(1), 124–149.

van den Berg, R., & Ma, W. J. (2018). A resource-rational theory of set size effects in human visual working memory. ELife, 7. https://doi.org/10.7554/eLife.34963

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

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

Wickens, C. D., Liang, C. C., Prevett, T., & Olmos, O. (1996). Electronic maps for terminal area navigation: Effects of frame of reference and dimensionality. The International Journal of Aviation Psychology, 6(3), 241–271.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. Nature, 453(7192), 233–235. https://doi.org/10.1038/nature06860

Zokaei, N., McNeill, A., Proukakis, C., Beavan, M., Jarman, P., Korlipara, P., Hughes, D., Mehta, A., Hu, M. T. M., Schapira, A. H. V., & Husain, M. (2014). Visual short-term memory deficits associated with GBA mutation and Parkinson's disease. Brain, 137(8), 2303–2311. https://doi.org/10.1093/brain/awu143

Chapter 3

Visual working memory and mental imagery: A behavioral comparison of stimuli and task design

Abstract

Although the literatures investigating visual imagery and working memory have largely remained independent, several recent findings suggest a significant overlap in the neural mechanisms that mediate the formation and manipulation of visual information. There is a paucity of evidence, however, to compare the behavioural properties of these processes, as research has focused on different stimuli, paradigms, and metrics to evaluate these mental representations. Hence, I attempt to compare the behavioral performance in visual working memory and mental rotation through similar measures. In Experiment 1, I evaluate the psychophysical properties of memory retention and manipulation for simple and complex objects using a delayed-recall continuous-response paradigm. I demonstrate that mental transformation results in lower accessibility, with the magnitude of the transformation affecting the quality of the reported information. In Experiment 2, I compare the metrics derived from a seminal dichotomous response mental rotation task to a delayed recall continuous response task to determine whether the tasks assess the same underlying processes. I show that performance is almost identical between tasks and across manipulations, and that participants perform the mental transformation at different speeds to accommodate for the magnitude of the rotation. I also applied Principal Component Analysis to all performance variables from both experiments revealing a shared underlying process between retention and transformation. Our results provide support for a common process to visual working memory and mental rotation.

Keywords: Working memory, imagery, mental rotation

Introduction

The ability to represent information that is no longer in view and the ability to perform operations on those representations are processes are critical to human cognition. Indeed, these processes are thought to be highly related, as perhaps most famously outlined in Baddeley and Hitch's (1974) model of "working memory", which established a framework in which the ability to manipulate information was directly tied to the process of holding information in short-term memory stores. Despite the influence of this theoretical framework, the processes of storing and manipulating information have often been studied independently (Tong, 2013). Particularly within the visual domain, researchers studying visual working memory (a.k.a. visual short-term memory) have extensively documented the psychophysical properties of memory storage of simple visual features, while much less focus has been placed on the "working" aspect of working memory. Interestingly, the manipulation of visual information has been extensively studied within the mental imagery field, largely independent of the psychophysical considerations common in the visual working memory literature. Therefore, the aim of this study is to address the lack of behavioral convergence between these fields of research and psychophysically evaluate the manipulation of mental representation of visual information.

Despite the sparse behavioral evidence for an overlap between mental imagery and working memory processes, there is substantial neural evidence suggesting that the mechanisms involved in imagery and working memory representation rely on overlapping brain regions. Indeed, studies have demonstrated that the primary visual cortex is substantially recruited during imagery (Cichy et al., 2012; Ishai & Sagi, 1995;

Kosslyn et al., 1997, 2001; Kosslyn & Thompson, 2003; Pearson et al., 2015), which has sometimes been equated to a weak form of perception (Pearson et al., 2015). Correspondingly, researchers studying the neural mechanisms of visual working memory have established that the primary visual cortex is considerably recruited during the storage of memory (Awh & Jonides, 2001; Curtis & D'Esposito, 2003; Emrich et al., 2013; Harrison & Tong, 2009; Silvanto & Cattaneo, 2010). These results have been interpreted as support for the sensory recruitment hypothesis developed in the working memory field, which suggests that memories are maintained in the same sensory areas responsible for encoding them (Postle, 2006; Serences, 2016; Serences et al., 2009). Thus, similar to the proposal that imagery can be characterized as a weak perception (Pearson et al., 2015), working memory maintenance also relies significantly on perceptual areas. Hence, these reports point towards a shared visual representation for both memory and imagery.

Compelling evidence demonstrating that imagery and working memory not only rely on the same regions but also share the same underlying representations was provided by Albers et al. (2013). Examining fMRI BOLD signals recorded from early visual areas, Albers et al. (2013) used multi-voxel pattern analysis to train a pattern classifier on the perceptual encoding of three orientation gradients, all separated by 60°. In a working memory task, participants were required to remember the orientation of two of the three possible orientation gradients, and, after a brief delay, they were cued to report the orientation of one of the gratings. Similarly, participants completed an imagery task, which was identical to the working memory task except that they were required to mentally rotate by 60° the cued memory sample. The pattern classifier, trained on the passive perceptual encoding of the orientation gradients, was able to recognize the orientation from BOLD activity during working memory maintenance and track the manipulation of the memoranda from one gradient to another during the mental rotation. Similar findings were observed by Christophel et al. (2015), who demonstrated that both remembered and transformed images were represented by activity in early visual areas, as well as posterior parietal cortex. These results suggest not only that both working memory and imagery manipulation recruit overlapping perceptual areas but also that the functional activity found during these tasks are also comparable.

Although there is substantial evidence suggesting that imagery and working memory rely on the same neural substrates, there have been fewer attempts to compare their psychophysical properties. One reason for the lack of direction behavioral comparison between the two fields may be that they have traditionally employed divergent methodological and theoretical approaches. On one hand, the imagery field has focused on understanding the format these visual representations take, and have employed dichotomous paradigms with complex stimuli (e.g., 3D cube structures). Specifically, there has been significant debate about whether imagery representations are best described as pictorial in nature (i.e., in a depictive or analogue format) or as a set of linguistic descriptors (i.e., propositional format) (Pearson & Kosslyn, 2015; Pylyshyn, 1973; Suchow et al., 2014). In a classic imagery study, Shepard and Metzler (1971) presented a pair of 3D figures made of cubes and participants were required to determine whether the two images depict the same object (with one of them rotated in space) or if one of the figures was mirror reversed. They demonstrated that the time required to correctly evaluate the rotated pairs was linearly related to the angular difference between the two objects. This suggested that performing a mental rotation takes proportionally the same time as though one were physically rotating an object, suggesting that the mental representation was pictorial in nature. While these results have been replicated abundantly, particularly within the sex differences literature (see Maeda & Yoon, 2013, for meta-analysis), there have been few attempts to characterize the nature of these transformed representations beyond examining accuracy or RT.

On the other hand, working memory researchers have extensively examined the qualitative and quantitative nature of representations in visual working memory by using both dichotomous and continuous response tasks. Using a forced choice, change-detection paradigm, Luck and Vogel (1997) demonstrated a sharp decline in working memory performance when more than 3-4 simple visual items were to be remembered. This paradigm requires participants to make a judgment on whether a display, which varied in the number of items (e.g., colored squares, oriented lines) presented, changed after a brief delay. While Wilken and Ma (2004) demonstrated the same sharp decline in memory performance using a continuous paradigm that allows participants to report the exact color of a probed item from a display using a continuous color wheel, Zhang and Luck (2008) used this paradigm to estimate quality and accessibility of memory representation by modeling participants' responses. Indeed, they designed a mixture model based on the idea that, on any given trial, participants can either make an informed choice about a probed item, because the colored square was successfully committed to memory, or make a random guess. These two possible responses, over the course of several trials, produces qualitatively different distributions; guesses produce a uniform pattern whereas target-related responses will cluster around the target value forming a pattern analogous to a normal distribution. The width of the target distribution also provides a metric for the precision of target responses (i.e., the fidelity of the maintained representation). Using this mixture model, Zhang and Luck (2008) demonstrated that when more than 3-4 items were displayed the level of response precision remains unchanged, while the proportion of guesses increased as a function of the number of items. These results supported a discrete-representation view of working memory, which suggests that, although strictly limited in number, items committed to memory can be recalled with high fidelity. While this conceptualization of working memory representations is currently being debated and several alternative models have been proposed (see Ma, Husain, & Bays, 2014, and van den Berg, Awh, & Ma, 2014), employing a continuous response paradigm has substantially furthered our understanding of the nature of working memory by investigating different properties of theses representations.

The type of stimulus used in the working memory and imagery fields also constitute a considerable difference that makes comparisons between the two literatures difficult. While the imagery literature has historically designed studies using complex stimuli (i.e., 3D structures, maps, common household objects) to evaluate mental manipulation of the information, the visual working memory field has typically assessed the psychophysical properties of simple items (i.e., lines, colored squares, simple geometric shapes). There is some evidence, however, suggesting that the complexity (i.e., based on the number of features) of an object affects working memory performance. For instance, Alvarez and Cavanagh (2004) demonstrated that working memory capacity decreases linearly as stimuli increased in complexity (e.g., squares, letters, random polygons, shaded cubes), where stimuli were previously operationally ranked through a visual search task. Similar effects of complexity reducing memory capacity have been observed in several behavioral (Delvenne & Bruyer, 2004; Eng et al., 2005; Wheeler & Treisman, 2002) and neuroimaging studies (Luria et al., 2010; Song & Jiang, 2006; Xu & Chun, 2006). Moreover, the nature of certain stimuli used in working memory experiments, which includes how familiar (Xie & Zhang, 2017), relatable (Zhou et al., 2018) or real (Snow et al., 2014) these stimuli appear, can have a beneficial influence on participants' performance. In an attempt to evaluate the effect of complexity absent of these benefiting factors, Robitaille and Emrich (2019) assessed the psychophysical properties of working memory for complex 2.5D block stimuli, similar to those used in the imagery literature, using a continuous report paradigm. However, while that study begins to address the effect of complexity on the psychophysical properties of memory representations, given the substantial differences in the paradigms used in the mental imagery and working memory fields, it remains difficult to make direct comparisons between these tasks.

To summarize, despite the theoretical link between working memory and the mental transformations performed in mental imagery, as well as the evidence that these representations share similar neural mechanisms, few attempts have been made to compare the nature of these representations within similar paradigms. To date, behavioral comparisons between the two fields have been made difficult due mainly to their choice of different stimuli and tasks. Therefore, the main objective of the current study is to compare the behavioral properties of imagery and working memory by using shared stimuli and tasks. More specifically, I aim to evaluate the psychophysical properties associated with the type of stimulus employed (Experiment 1) as well as comparing task specific metrics (Experiment 2) commonly used in the working memory and imagery literatures. In Experiment 1, I evaluated the psychophysical properties associated with the maintenance and manipulation of simple stimuli commonly used in the working memory literature and compared it to complex objects classically employed in the imagery field. To answer this question, I used an analytical approach typically used in the working memory literature by applying a mixture model (Zhang & Luck, 2008) to participants' responses in a continuous response paradigm. In Experiment 2, I compared the dichotomous Shepard and Metzler (1971) imagery paradigm with a continuous report version of the task to evaluate the extent to which these metrics were related to one another.

Experiment 1

Methods

Participants. Forty-five (9 females; normal or corrected-to normal vision, 17-34 years of age, M = 20.22) Brock university students were recruited for this first experiment, where one participant withdrew from the study after completing a single working memory task. Participants were screened for color-blindness using Ishihara's 38 plates test upon arrival and they were compensated with course credits or $10/hr. at the end of the experiment. All procedures were approved by the Brock University Research Ethics Board.

Apparatus. All tasks and questionnaires were displayed on a 20-in. flat LCD with a resolution of 1600-1200 ppi and a refresh rate of 60 Hz. Stimuli for the working memory tasks were generated and presented using MATLAB (mathworks.com) software along with the Psychophysics Toolbox extension (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997). Questionnaires were presented using PsychoPy2 (v1.90.3) software (Peirce, 2007).

Stimuli. For the simple stimulus, I used a 2D line 4.5° of visual angle long and 0.45° wide was used, in which the orientation is manipulated form the center of the stimulus and can span the full 180° of possible orientations. Conversely, I also used more complex objects, similar to those commonly used in the imagery literature, which consist of 25 possible 3D arrangement of 9 cubes, with 4-5 90° bends in the arrangement of those cubes. These shapes were taken from a previous study (see Robitaille & Emrich, 2019) and can be found on an open source repository (https://osf.io/9cdqz/). The size of these objects can vary between 1.32 – 2.6 visual degrees in any directions and they were given an initial random orientation at the beginning of each trial. These stimuli were presented on a grey background (RGB: .5 .5 .5) and they were given an arbitrary color by picking 2 RGB values above .2 and leaving the third value at 0 to ensure that the color is distinctive enough and can be distinguished from the background color.

Procedure. Participants completed two continuous-recall tasks with a mental rotation manipulation, which were identical with the only exception of the complexity of the stimulus presented (see Figure 3-1A). These tasks required participants to remember the orientation of a stimulus, mentally rotate the object, and report its final orientation on a continuous probe. At the beginning of each trial, a cue was presented to indicate participants whether the participant was required to report the orientation of the stimuli exactly as it was presented (i.e., no rotation) or whether they would be required to perform a mental rotation of either 60° or 120°. The direction of the rotation was conveyed by an arrow pointing to the right (>) or left (<), where participants were required to mentally transform the stimulus clockwise or counter-clockwise, respectively, while the magnitude of the transformation is indicated by a single or double arrow (e.g., < : 60° counter-clockwise, >> : 120° clockwise). For the working memory condition, a circle cue (i.e., **O**) was presented to indicate that no transformation was required on this trial.

A trial would begin by presenting the mental rotation cue for 500ms, followed by a 1000ms delay and then a colored stimulus would be displayed for 500ms. Following a 500ms delay, the colored stimulus would re-appear, and participants were tasked to report the orientation of the colored probe after being mentally transformed (or not) according to the cue. When participants make their response, their reported figure lingers for 500ms while presenting the correct orientation in the form of a white overlapping shape to provide feedback on their performance. In either the simple or complex shape task, participants were required to complete a total of 240 trials, 80 trials for each of the three possible mental rotation conditions (i.e., 0°, 60°, 120°) where half of the trials for the 60° and 120° rotations required a clockwise transformation. Note that, for the purpose of our analyses, the directions of the mental rotation were collapsed. To reduce task order effect, 24 participants began the experiment with the simple line stimulus and 21 with the complex shapes. Before the beginning of each recall task, participants were trained on a series of practice trials (at least 20) until they reported being comfortable with the task, the imagery cues and the magnitude of the rotations.

After the two continuous-recall tasks were completed, participants were then required to complete a computerized version of the Vividness of Visual Imagery

Questionnaire (VVIQ-rv) (Marks, 1973; McKelvie, 1995). This questionnaire is designed to obtain a self-report measure of one's ability to produce mental images. Participants are instructed to visualize 8 different scenarios while keeping their eyes open, and they are required to assess the vividness of the details of their imagery ability on a scale from 0 "No images at all, you only 'know' that you are thinking of the object" to 7 "Perfectly clear and as vivid as real seeing." Each scenario assesses 4 different contextual details for a total of 32 questions. Participants were then required to complete the Object-Spatial Imagery and Verbal Questionnaire (OSVIQ) (Blajenkova et al., 2006) designed to determine the type of mental visualization strategy that participants engage in on a day-to-day basis. The questionnaire requires participants to rate how they relate to a particular statement using a scale ranging from 1 "Totally disagree" to 5 "Totally agree". All three subscales (i.e., object, spatial, and verbal) contain 15 statements for a total of 45 statements. These questionnaires were included in order to evaluate the relationship between participant's task performance and their subjective report of their ability to visualize objects and their preferred visualization strategy.

Analysis. For each participant, performance was evaluated using raw response error, where the smaller values indicate that responses are close to the target. This measure has been shown to produce a reliable measure of participants' performance (Ma, 2018; van den Berg & Ma, 2018). For each condition, raw error was computed by taking the standard deviation of the absolute angular difference between the target orientation of the object and the response.

While the raw error is statistically analyzed without any further transformation, error (i.e., angular difference) was also modeled according to a two-component mixture model suggested by Zhang and Luck (see Zhang & Luck (2008), for a full description of the model). This model assumes that the distribution of responses is a mixture of two distributions that each represents a different type of response. In the case where participants fail to commit the information to memory, they will guess their answer on the continuous probe, which will produce a uniform – or flat – distribution over multiple trials. On the other hand, when participants have some accurate memory about the stimulus, their response will cluster around the target orientation and create a Von-Mises distribution (i.e. circular analog to the Gaussian distribution). The proportion of guesses (P_g) and target responses (P_T) are approximated using maximum likelihood estimation (MLE) and, since these proportions are complementary in this model (i.e., $P_g + P_T = 1$), only the proportion of guesses are reported here. The standard deviation (SD) of the Von-Mises distribution also provides a measure inversely related to response precision (or precision^{-1}). The parameters of the mixture model were computed using the Memtoolbox Matlab extension (see Suchow, Brady, Fougnie, & Alvarez, 2013; Suchow et al., 2013; memtoolbox.org).

All statistical analyses were conducted using JASP (JASP Team, 2018) statistical software, where 2 (simple vs. complex stimulus) x 3 (rotations: 0°, 60°, and 120°) within-subject repeated measures frequentist and Bayesian ANOVAs were computed from participants' raw error, proportion of guesses and response precision. For the frequentist analyses, degrees of freedom were corrected using Greenhouse-Geisser estimate when the assumption of sphericity was violated and all post-hoc analyses were Bonferroni corrected to account for multiple comparisons. Bayesian ANOVAs creates a series of models that accounts for each possible main effect separately, all possible combinations of the main effects, and the main effects with an interaction, which assesses the strength of each model at explaining the data. Therefore, each Bayesian model was evaluated as a function of their likelihood in contrast to a null model (i.e., $\log BF_{10}$) as well as the evidence against the other models tested ($\log BF_M$). The Bayes Factor (BF) reported here represents the strength of the evidence in favour of the alternative hypothesis and they are interpreted according to Lee & Wagenmakers (2014) revised version of Jeffreys's guideline (1961), where $BF_{10} = 1$ or $\log BF_{10} = 0$ represents no evidence, $BF_{10} = 10$ or $\log BF_{10} = 1$ is moderate evidence, $BF_{10} = 30$ or $\log BF_{10} = 1.5$ is strong evidence, $BF_{10} > 100$ or $\log BF_{10} > 2$ is extreme evidence in favour of the alternative hypothesis. In contrast, $BF_{10} < 1$ or $\log BF_{10} < 0$ represents evidence in favour of the null hypothesis. Note that Bayesian ANOVAs were conducted with JASP's default Cauchy prior term.

To investigate the underlying mechanisms driving task performance, I used Principal Component Analysis (PCA) as a variable reduction method. Since I have a total of six performance variables that reflect all combinations of working memory/imagery transformation, stimulus complexity and transformation magnitude, this method enables the exploration of the nature of the mechanism(s) mitigating overall performance. I determined the number of components using the default JASP parallel analysis, Kaiser rule (i.e., eigenvalue > 1) and the Scree plot test, and a Promax oblique rotation between components when appropriate. To investigate the association between performance and self-report measures of visual imagery (i.e., VVIQ-rv) and cognitive strategies (i.e., OSIVQ), I computed correlations between task performance and the questionnaire scores. Task performance was represented by a weighted composite scores (CS) derived from the component loadings from the PCA for each participant. Composite scores were calculated by averaging the product of all six ($N = 6$) performance variables (x_i) with their corresponding component loadings (l_i) (i.e., $CS = \frac{\sum_{i=1}^{N} l_i x_i}{N}$) (Chao & Wu, 2017). These composites scores were then correlated with the VVIQ and OSIVQ scores.

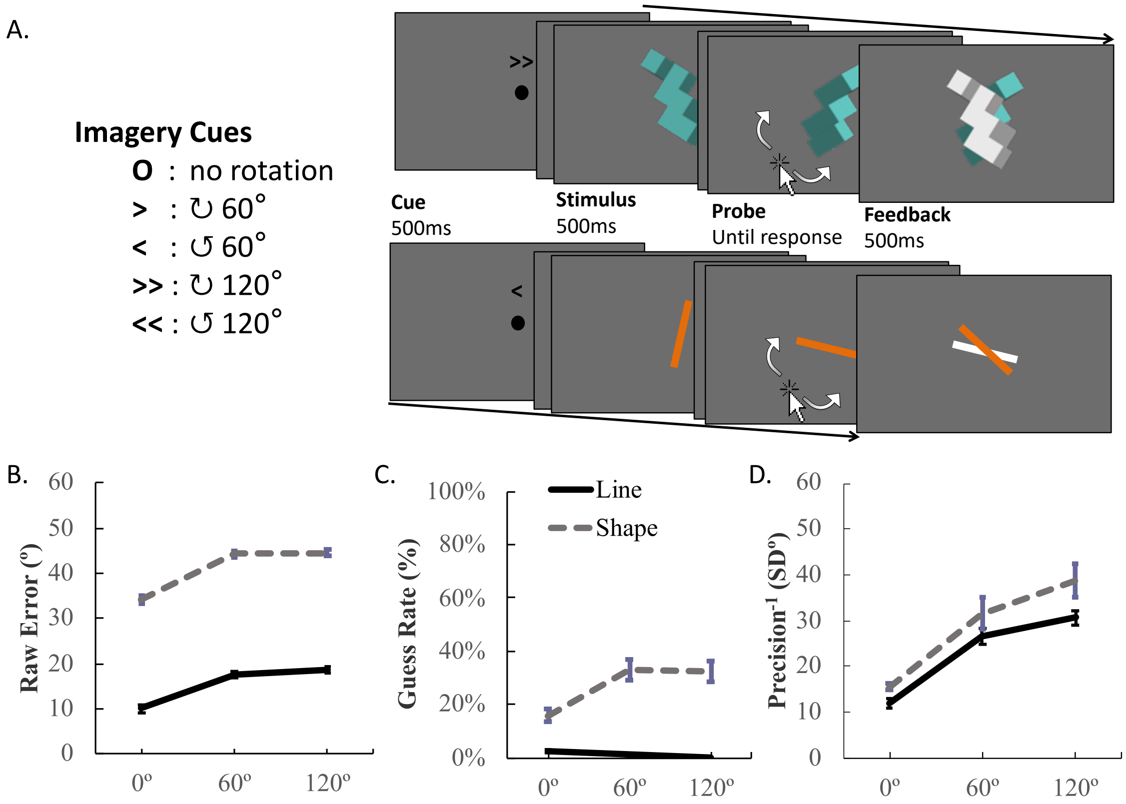

Figure 3-1. Task schematics and behavioural results of Experiment 1.(A) Schematic of the two continuous recall tasks. At the beginning of each trial, participants are presented with a mental rotation cue for 500ms. After a 1000ms delay, a stimulus is displayed for 500ms, where participants were required to mentally rotate the object according to the imagery cue. After a 500ms delay, the probe appears where participants use the mouse to rotate the object to the proper orientation and click to report the angle. When the response is made, the response stimulus lingers while a second white stimulus depicting the correct orientation appears on the screen to provide feedback on their performance. Performance results across stimulus type and rotation manipulation (x-axis) for (B) raw error, (C) guess rate, and (D) precision^{-1}, where the solid black represent the line stimulus, the dashed grey line depicts the complex stimulus, and the error bars are showing standard error.

Results

Raw Error. I investigated participants' raw error as an initial measure of performance, which consist of the circular standard deviation (SD) of the absolute angular difference between the target and reported orientations (see Figure 3-1B). A frequentist and Bayesian repeated measures ANOVA were conducted to compare stimulus type (simple vs. complex) and imagery manipulation condition (0°, 60°, and 120°). The Bayesian analysis provides greater evidence in favor of a model that includes object complexity and rotation manipulations without an interaction in comparison to the null model, $BF_M = 2.918, \log(BF_{10}) = 227.8$, closely followed by the model that also includes the interaction, $BF_M = -0.145, \log(BF_{10}) = 226.2$. Likewise, frequentist repeated measures ANOVA revealed a significant main effect of object complexity, $F(1, 43) = 640.8, p < .001, \eta^2 = 0.69$, and rotation, $F(1.60, 68.92) = 96.02, p < .001, \eta^2 = 0.08$, but no interaction, $F(1.57, 67.6) = 2.41, p = .1091, \eta^2 = 0.001$. Therefore, I further investigated the main effects without interaction through Bayesian and frequentist post-hoc comparisons. The main effects of complexity revealed that simple objects produce lower raw error than complex objects, $M_{diff} = 25.46°, SE = 1.006°, \log(BF_{10}) = 151.76$, whereas the main effect of transformation showed significant differences between the working memory (i.e., 0°) condition and both the 60°, $M_{diff} = 8.945°, SE = 0.860°, \log BF_{10} = 37.20$, and the 120°, $M_{diff} = 9.523°, SE = 0.862°, \log BF_{10} = 39.77$, mental transformations, but not between the two mental rotation magnitudes, $M_{diff} = 0.578°, SE = 0.546°, \log BF_{10} = -1.47$.

Mixture Model.

Guess rate. Using Zhang & Luck's (2008) mixture model, I derived the proportion of guesses from participants' distribution of error for each condition (see Figure 3-1C). A Bayesian repeated measures ANOVA reveals that the best model includes an interaction between the object complexity and rotation, $\log(BF_M) = 9.184, \log(BF_{10}) = 76.50$. Similarly, a frequentist repeated measure ANOVA reveals significant main effects of object complexity, $F(1,43) = 80.82, p < .001, \eta^2 = 0.37$, and rotation, $F(1,86) = 13.40, p < .001, \eta^2 = 0.028$, as well as an interaction, $F(1,86) = 21.38, p < .001, \eta^2 = 0.043$. Investigating the interaction reveals that changes in rotation magnitude did not impact the proportion of guesses for simple stimulus ($p > .999$), unlike with the complex stimuli ($p < .001$). Post-hoc frequentist t-test comparisons for complex objects demonstrate significantly fewer guesses when no transformation is performed to the memory representation than when transformed by 60°, $M_{diff} = 17.3\%,\ SE = 2.4\%, p < .001, \log BF_{10} = 5.88$, or 120°, 16.6%, $SE = 2.4\%, p < .001, \log BF_{10} = 4.16$. Interestingly, no difference in guess rates were observed between 60° or 120° rotations, $M_{diff} = 0.7\%,\ SE = 2.4\%, p < .001, \log BF_{10} = -2.03$. These results indicate that participants tend to produce guesses only when exposed to complex stimuli, and while rotating complex stimuli increases overall guessing, the proportion of guesses do not scale as a function the magnitude of rotation.

Precision^{-1} (SD). Next, I looked at participants' response precision^{-1} defined as the width (SD) of the participant's response distribution (see Figure 3-1D). Frequentist repeated measure ANOVA revealed a main effect of stimulus type, $F(1,43) = 7.55, p =$

$.009, \eta^2 = 0.024$, and mental rotation, $F(1, 86) = 71.73, p < .001, \eta^2 = 0.24$, but no interaction, $F(1, 86) = 0.83, p = .439, \eta^2 = 0.003$. Similarly, Bayesian repeated measure ANOVA demonstrates that the best model includes the main effect of stimulus type and the mental rotation without any interaction, $\log(BF_M) = 3.23, \log(BF_{10}) = 45.58$. Post-hoc frequentist and Bayesian t-test comparisons revealed that, for the main effect of stimulus type, the simple stimulus was reported with significantly greater precision (lower error) than complex objects, $M_{diff} = 5.63°, SE = 2.053°, p = .009, \log(BF_{10}) = 2.998$. Post-hoc comparisons for main effect of rotation reveals that response for items that were not mentally transformed are significantly more precise than when a 60°, $M_{diff} = 15.32°, SE = 1.74°, p < .001, \log BF_{10} = 24.58$, or 120°, $M_{diff} = 21.00°, SE = 1.85°, p < .001, \log(BF_{10}) = 37.48$, rotations is applied. Precision was also better for 60° than 120° rotations, $M_{diff} = 5.68°, SE = 1.86°, p = .012, \log(BF_{10}) = 2.33$.

Principal component analysis (PCA) and correlations. To evaluate the shared mechanisms involved in the retention and manipulation of simple and complex objects, I conducted a principal component analysis (PCA) using raw error for simple and complex objects across all conditions (see Table 3-1). A parallel analysis suggested that a single component model best represents the relationship between all variables, where this component explains 59.4% of the variance. This model provides evidence for a common mechanism responsible for working memory and mental rotation of simple and complex objects. I also evaluated a two-component model, which is supported by the Keiser criterion (i.e., eigenvalues above 1), using a Promax oblique rotation, which demonstrates that a different component loads on each stimulus type (i.e., simple features vs complex shapes). By opting for rotated components rather than forcing an orthogonal solution, I allow these components to not be strictly independent from each other. However, while these two components reflected the two different stimulus types, they were indeed highly correlated, $r_{RC1-RC2} = .513$, which suggest that performance across the two stimulus types are related. It also demonstrates that there is more shared variance between the maintenance and manipulation of the visual representation than between stimulus types, as a single component can account for the maintenance (i.e., 0-degree rotation) and transformation of each stimulus.

In order to examine the relationship between memory and mental rotation abilities and self-report measures of mental imagery, I computed correlation between the VVIQ-rv and the OSIVQ subscales with participants performance measures. Since I obtained six behavioural measures for each combination of stimulus type and mental transformations, I calculated weighted composite score (CS) based on the PCA component loadings for the single component model (PC1) ($M_{CS.PC1} = 21.76°, SE_{CS.PC1} = 0.628°$) as well as two composite scores from the two-components solution ($M_{CS.RC1} = 13.65°, SE_{CS.RC1} = 0.600°; M_{CS.RC2} = 36.15°, SE_{CS.RC2} = 1.019°$). Based on the loading pattern of the two-component model, the latter composite scores reflect a performance measure associated with the simple and complex stimulus respectively. Pearson and Bayesian correlation between composite scores and the self-report questionnaire scores are found in Table 3-2. Interestingly, not only I did not find any significant relationship between the three performance composite scores and self-report measures of vividness of imagery or imagery strategies, but also I generally found moderate to strong evidence in favour of a null relationship across all correlations.

Table 3-1. Principal component loadings for a single and two component model of performance for Experiment 1 and 2

Experiment 1	Single Component*		Two-Components**		
	PC 1	Uniqueness	RC1	RC2	Uniqueness
Simple					
0°	**.725**	.475	**.866**	-.033	.279
60°	**.783**	.386	**.930**	-.031	.163
120°	**.808**	.347	**.849**	.079	.204
Complex					
0°	**.752**	.435	.057	**.808**	.297
60°	**.788**	.379	.001	**.906**	.179
120°	**.766**	.414	-.040	**.921**	.188
Variance accounted	.594		.779	.774	

Experiment 2	Single Component*		Two-Components**		
	PC 1	Uniqueness	RC1	RC2	Uniqueness
Dichotomous					
0°	**.790**	.376	**. 935**	**-.**057	.171
50°	**.783**	.386	**. 982**	**-.**019	.053
100°	**.889**	.210	**.876**	**.**132	.112
150°	**.862**	.257	**.982**	**-.**019	.053
Continuous					
0°	**.635**	.597	**-.**001	**.793**	.371
50°	**.756**	.428	.022	**.921**	.134
100°	**.608**	.630	**-.**131	**.906**	.268
150°	**.777**	.397	.133	**.822**	.210
Variance accounted	.606		.893	.743	

Note: PCA is performed on raw error across stimulus type and rotation for Experiment 1, whereas raw error and percent inaccuracy is used for the continuous and dichotomous task respectively across rotation for Experiment 2. Bolded numbers represent the primary loadings on each of the components. The variance accounted for is computed from the average of the squared bolded loadings. *No rotation method applied to the single Principal Component (PC1). **Promax oblique rotation method applied to the two Rotated Components (RC1 and RC2), where the components are correlated with each other, Exp1: $r_{RC1-RC2} = .513$, Exp2: $r_{RC1-RC2} = .447$

Table 3-2. Correlation between weighted composite scores and VVIQ and OSIVQ scales for Experiment 1 and 2.

Experiment 1

	Composite Score (PC1)		Simple object Composite Score (RC1)		Complex object Composite Score (RC2)	
	r	log BF$_{10}$	r	log BF$_{10}$	r	log BF$_{10}$
VVIQ	.099	-1.475^{-}	.097	-1.484^{-}	.121	-1.376^{-}
OSIVQ						
Object	-.082	-1.529^{--}	-.081	-1.532^{--}	-.068	-1.570^{--}
Spatial	-.052	-1.608^{--}	.058	-1.593^{--}	-.015	-1.656^{--}
Verbal	-.017	-1.655^{--}	.186	-0.973	.050	-1.611^{--}

Experiment 2

	Composite Score (PC1)		Continuous task Composite Score (RC1)		Dichotomous task Composite Score (RC2)	
	r	log BF$_{10}$	r	log BF$_{10}$	r	log BF$_{10}$
VVIQ	- .403**	1.894^{++}	-.317*	0.451	-.391**	1.662^{++}
OSIVQ						
Object	.026	-1.680^{--}	-.017	-1.688^{--}	.078	0.185
Spatial	-.525***	5.110^{+++}	-.470***	3.564^{+++}	-.427**_	34.445^{+++}
Verbal	.261	-0.203	.212	-0.724	-.246	0.493

Note: Correlations of weighted composite scores from the component loadings of the PCA and VVIQ and OSIVQ scores.
* p < .05, ** p < .01, *** p < .001
Evidence in favour for H$_1$: $^{+}$ $\log BF_{10} \geq 1$ (moderate), $^{++}$ $\log BF_{10} \geq 1.5$ (strong), $^{+++}$ $\log BF_{10} \geq 2$ (extreme)
Evidence in favour for H$_0$: $^{-}$ $\log BF_{10} \leq -1$ (moderate), $^{--}$ $\log BF_{10} \leq -1.5$ (strong), $^{---}$ $\log BF_{10} \leq -2$ (extreme)

Discussion

In this first experiment, I evaluated the effect of mental rotation on raw error, proportion of guesses, and memory precision of simple and complex stimuli using a continuous recall paradigm. Our results successfully replicated those typically observed in working memory experiments; namely, that performance is reduced when reporting the orientation of complex objects relative to simple stimuli. Indeed, several studies that used a dichotomous response tasks have demonstrated that working memory accuracy is worse for complex compared to simple stimuli (Alvarez & Cavanagh, 2004; Delvenne & Bruyer, 2004; Wheeler & Treisman, 2002). The results from Alvarez and Cavanagh (2004) in particular suggest that the performance costs are likely the result of increasing cognitive load associated with greater featural information, as they demonstrated that participants working memory capacity decreases as item complexity increases. Here, I replicate our previous findings (Robitaille & Emrich, 2019), and demonstrate that this cognitive cost is not limited to a dichotomous recall paradigm. Importantly, I also demonstrate that response error (i.e., raw error) is impaired following any of the tested mental rotations. Interestingly, however, there is no interaction between stimulus type and mental rotation, and raw error did not scale as a function of the magnitude of rotation. These results suggest that, not only is the additional error acquired from the manipulation of these objects not cumulative, but it also appears to affect both stimuli to the same degree.

However, when I further investigated the source of these errors using the mixture model, I observed that the nature of the stimuli produces different types of errors. While the simple stimuli are easily remembered regardless of how they are manipulated, the accessibility of the mental representations for complex objects is more compromised. Indeed, not only do guess responses increase significantly when a complex object is to be remembered compared to a simple line stimulus, guess rates almost double once participants are required to mentally rotate the object. Hence, the act of transforming the memoranda for complex stimulus puts memory retention at risk. Once manipulated, however, the maintenance of the memory is not affected by the degree in which it is rotated. It is unclear from these results whether this reduced memory retention is a fixed cost due to mental rotation that plateaus across any rotation magnitude. It is likely that these retention errors are largely responsible for the difference in raw error observed between simple and complex stimuli is mainly driven by guessed responses. While stimulus complexity also significantly affects recall precision^{-1}, it appears that complexity has a greater impact memory encoding or retention than the precision of the memory. The quality of the mental representation, however, appears to be more directly affected by the magnitude of the rotation.

I also used principal component analysis to examine the shared and unique contributions to memory and mental rotation, which revealed a strong reliance on common mechanisms driving participant's performance regardless of stimulus type or the presence or extent of mental transformation. These results further support the idea that working memory and imagery manipulations rely on the same underlying cognitive processes. As discussed previously, there is substantial neural evidence suggesting that the visual cortex is recruited during both imagery (Cichy et al., 2012; Ishai & Sagi, 1995; Kosslyn et al., 1997, 2001; Kosslyn & Thompson, 2003; Pearson et al., 2015) and working memory (Awh & Jonides, 2001; Curtis & D'Esposito, 2003; Emrich et al., 2013;

Harrison & Tong, 2009; Silvanto & Cattaneo, 2010), as well as functional activity relating these processes together (Albers et al., 2013; Christophel et al., 2015). It appears that our behavioural measures also support the representational overlap between these processes as I observed significant shared variance among performance variables. The complete absence of a correlation between our composite scores and the subjective reports of imagery abilities (i.e., VVIQ-rv) and strategy (i.e., OSIVQ) raises some concern about the involvement of imagery processes. Typically, the imagery literature reports strong correlation between task performance and the VVIQ score and the object subscale of the OSIVQ (Blazhenkova & Kozhevnikov, 2008). Our results suggest that participants are completing in a way that is distinct from their self-reported imagery abilities. Note that reports about the relationship between subjective rating of imagery and spatial abilities are not consistent (Dean & Morris, 2003; McAvinue & Robertson, 2007) and it is proposed that correlation between self-report measures and imagery ability is task-specific (Dean & Morris, 2003). I further evaluate this hypothesis in Experiment 2.

Experiment 2

In Experiment 1, I investigated the psychophysical properties of stimuli common in the working memory and imagery literature (i.e., lines and 3D block objects, respectively) using a continuous recall paradigm. This second experiment specifically evaluates how analogous a dichotomous and continuous task design are at evaluating mental rotation.

Methods

Participants. A total of 47 undergraduate Brock students (35 females; normal or corrected-to normal vision, 17-29 years of age, M = 19.49) were recruited for this second experiment. Participants were screened for color blindness using Ishihara's 38 plates test and they were compensated with course credits or $10/hr. All procedures were approved by the Brock University Research Ethics Board.

Apparatus and Stimuli. The same computers, monitors, questionnaires, and programs were used as in Experiment 1. The stimuli for this experiment are based on Ganis & Kievit (2015) three dimensional shapes available in image format. Because I needed a dynamic version of these objects for the continuous response paradigm, I recreated their stimuli using MATLAB (mathworks.com) software and they were rendered using the OpenGL functions of the Psychophysics Toolbox extension (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997).

Procedure. Participants completed an imagery task based on Sheppard & Metzler's (1971) original paradigm (see Figure 3-2A), where two static 3D objects were presented side-by-side. The 3D objects could either depict the same objects or a mirror image of the one object. In each trial the object on the right was rotated by 0°, 50°, 100°, or 150°, as reported in Ganis & Kievit (2015). Participants were instructed to mentally rotate the right object to an orientation that matches the other object to determine whether the objects are the same or different. They were also instructed to go as fast as they can, while remaining accurate, and to report their response by pressing the "Z" key for mirrored objects or the "M" key when the objects were the same. After their response, a "Correct" or "Incorrect" message was displayed on the screen for 500ms to provide participants with feedback on their performance. Following with an inter-trial interval blank screen for 500ms, a new pair of objects would appear. Participants were presented with all 48 shapes created by Ganis & Kievit (2015) in a randomized 2 (mirror/same) x 4 (rotations) for a total of 384 trials.

The working memory task consisted of a delayed continuous-recall paradigm using the same stimuli from the previous task. A trial would begin with a fixation point followed by the display of a single 3D object for 500ms, where participants were instructed to remember its shape and orientation. After a 500ms delay, a number cue appeared at the bottom of the screen instructing the participants to either report the orientation of the object exactly as it was presented (i.e. 0°) or preform a mental rotation of 50°, 100°, or 150° over the y-axis, where the direction of the rotation was denoted by an arrow (i.e., < or >). The imagery manipulation cue stayed on the screen for as long as the participant needed to mentally rotate the object in their mind. Moving the mouse triggered the probe window to appear and participants were instructed to move the mouse only when they have a clear idea of the orientation they wish to report. When the mouse was moved, a grey version of the 3D object reappeared, initially in a random orientation over the y-axis, and rotated freely accordingly to the mouse movement. Participants were then able to reproduce the orientation of the object that matched their mental representation, and they would submit their response with a click of the mouse. To provide some feedback, the grey probe would linger for 500ms and a white 3D shape in the correct orientation would overlap the probe, which highlighted the discrepancy with the actual orientation. Participants completed 80 trials per rotation angle, where half were in the clockwise direction, for a total of 320 trials. Like in the previous experiment, participants completed the VVIQ-rv and the OSIVQ questionnaires at the end the session.

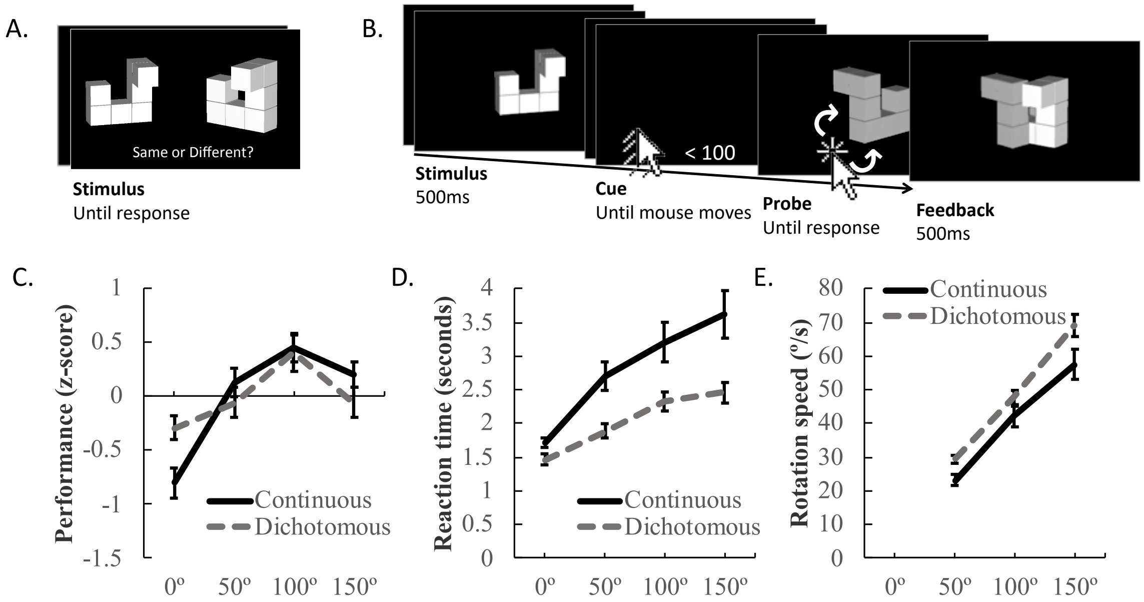

Figure 3-2. Task schematic and behavioural results of Experiment 2. (A) Schematic representation of the dichotomous imagery task, where participants are instructed to repot whether the two images are depicting the same object or a mirror image of one another. (B) Schematic representation of the continuous recall task, where participants are presented with an 3D object for 500ms (Stimulus window). After a 500ms delay, an imagery cue appears to instruct participants to mentally rotate the object. Since the probe window does not appear until the mouse is moved, participants were instructed to complete their mental rotation before moving their mouse. Participants were required to report the result of their mental rotation by rotating a grey probe object using the mouse and clicking to report their answer. A second object, in white, depicting the correct angle would appear on top of their answer to provide feedback for 500ms. (C) Standardized inaccuracy scores comparison between the continuous (black line) and the dichotomous task (grey line), (D) reaction time for the mental rotation (in seconds) and (E) speed of rotation (in °/seconds). Error bars represent standard error.

Analysis. Performance on the dichotomous task was assessed in terms of percent inaccuracy (i.e., %inacc. = 1 - %acc.) so that, analogously to raw error from the continuous task, greater %inacc corresponds to worse performance. In order to evaluate whether both tasks assess mental rotation performance in a similar manner, I standardized the raw scores from the continuous task and the %inacc. from the dichotomous task by standardizing (i.e., z-scores) performance measures for every participant using the mean and standard deviation for each mental rotation condition.

I also evaluated RT in both tasks. In the dichotomous task, reaction time was measured from the moment the stimulus window appeared until the participant made their response. To approximate the time required for participants to make their mental rotation in the continuous task, participants were instructed to complete the mental transformation before they triggered the appearance of the probe by moving the mouse. Therefore, the reaction time period for the continuous task began when the rotation cue appears and ends when participants moved the mouse to make their response. I also evaluated the speed of rotation (°/s) for each rotation and each task, which is computed by dividing the number of degrees rotated by the reaction time. Note that, as opposed to our previous measures, I only have three levels of rotation as there is no speed associated with the 0° rotation.

As described in Experiment 1, here too I conducted a PCA as a variable reduction method as well as a way to investigate the shared variance among performance measures. Furthermore, I used the component loadings to compute weighted composite scores for each relevant component and calculated correlations between task performance and the VVIQ-rv and OSIVQ scores.

Results

Task performance. I used a Bayesian and frequentist repeated measures ANOVA to evaluate the effect of task and mental rotation magnitude on overall inaccuracy (See Figure 3-2C). The Bayesian analysis revealed that a model including the interaction between task and rotation best describes the data, $\log BF_M = 4.61, \log BF_{10} = 42.96$, followed by a main effect of rotation alone, $\log BF_M = -1.946,\ \log BF_{10} = 39.73$. Similarly, the frequentist repeated measures ANOVA reveals a significant main effect of rotation, $F(2.28,104.84) = 65.474, p < .001, \eta^2 = 0.123$, and interaction between task and rotation, $F(2.33,107.23) = 16.009, p < .001\ , \eta^2 = 0.023$, but no main effect of task, $F(1,46) = 9.881\text{e}^{-21}, p > .999, \eta^2 < 0.001$. I further explored the interaction by investigating the simple effect of task on each level of rotation. A significant difference was found only when no rotation is performed, where performance is better on the dichotomous than the continuous task, $M_{diff} = 0.512°,\ SE = 0.149°, p = .027$, whereas there was no significant difference between task when the objects were rotated by 50°, $M_{diff} = 0.190, SE = 0.149, p > .999$, 100°, $M_{diff} = 0.056, SE = 0.149, p > .999$, or 150°, $M_{diff} = 0.266,\ SE = 0.149, p > .999$.

Investigating the rotation simple effect for the dichotomous task only shows significant differences with the 100° rotation, where performance is significantly worse (i.e., greater z-score) in comparison to no rotation, $M_{diff} = 0.692,\ SE = 0.094, p < .001$, to 50°, $M_{diff} = 0.461,\ SE = 0.094, p < .001$, and to 150°, $M_{diff} = 0.461,\ SE = 0.094, p < .001, d = 0.714$. No difference was found between 0° and 50°, $M_{diff} = 0.231,\ SE = 0.094, p = .410$, 0° and 150°, $M_{diff} = 0.231, SE = 0.094, p = .410$, or between 50° and 150°, $M_{diff} = 2.165\text{e}^{-15},\ SE = 0.094, p > .999$. I also evaluated the significant simple main effect of rotation on the continuous task only, $F = 53.096\ , p < .001$, which reveals that applying no rotation to the stimulus results in significantly less errors than 50°, $M_{diff} = 0.933,\ SE = 0.094, p < .001$, 100°, $M_{diff} = 1.259,\ SE = 0.094, p < .001$, or 150°, $M_{diff} = 1.009, SE = 0.094, p < .001$, and that performance on the 50° rotation is better than 100°, $M_{diff} = 0.327,\ SE = 0.094, p = .017$. No differences were found between 150° and 50°, $M_{diff} = 0.076,\ SE = 0.094, p > .999$, and 100°, $M_{diff} = 0.251,\ SE = 0.094, p = .228$.

Reaction time. I also examined participants reaction times (see Figure 3-3D) when performing both tasks. Bayesian two-way repeated measures ANOVA revealed that the main effects of both task and rotation best describe the data, $\log BF_M = 1.739, \log BF_{10} = 46.73$, closely followed by a model that includes an interaction, $\log BF_M = 1.034, \log BF_{10} = 46.47$. The frequentist ANOVA reveals a significant effect of task, $F(1{,}46) = 10.251, p = .002, \eta^2 = 0.065$, rotation, $F(1.176{,}54.112) = 70.573, p < .001, \eta^2 = 0.130$, as well as an interaction, $F(1.302{,}59.907) = 6.322, p = .009\ , \eta^2 = 0.012$. I therefore investigated the interaction further by evaluating the simple main effect of task on each rotation manipulation, which showed a significant difference only for the 150° rotation, $M_{diff} = 1.159\text{s},\ SE = 0.276\text{s}, p = .002$, where it takes more time performing the mental rotation for the continuous task than the dichotomous. All other level of rotation were not significantly different between task; that is for 0°, $M_{diff} = 0.251\text{s},\ SE = 0.276\text{s}, p > .999, d = 0.133$, 50°, $M_{diff} = 0.819\text{s}, SE = 0.276\text{s}, p = .112\ , d = 0.433$, and 100° rotations, $M_{diff} = 0.878\text{s}, SE = 0.276\text{s}, p = .059, d = 0.464$, which indicates that time necessary for performing the mental rotation does not differ significantly between the continuous and dichotomous response tasks in relatively small rotation magnitude.

Next, I found a significant simple main effect of rotation for the dichotomous task, $F = 110.65, p < .001$, where the reaction time is significantly ($p < .001$) faster when no rotation is applied compared to both 100° and 150°, and faster at 50° than 150°, at $p = .005$. There was no difference between no rotation and 50°, $M_{diff} = 0.420\text{s}$, $SE = 0.151\text{s}, p = 0.161, d = 0.406$, between 50° and 100°, $M_{diff} = 0.442s,\ SE = 0.151\text{s}, p = 0.105, d = 0.426$, as well as between 100° and 150°, $M_{diff} = 0.131\text{s}$, $SE = 0.151\text{s}, p > .999, d = 0.126$. Despite the fact that each level of rotation is not significantly different from their neighboring levels, there is still evidence for a slow but linear increase in reaction time happening between rotation magnitude in the dichotomous task, which is often reported in mental rotation studies (Cooper & Shepard, 1973; Ganis & Kievit, 2015; Kung & Hamm, 2010; Searle & Hamm, 2012; Shepard & Metzler, 1971).

I also evaluated the simple main effect of rotation on the continuous task, which revealed a significant difference, $F = 31.994, p < .001$, where no rotation showed significantly faster reaction times than all other rotations ($p < .001$) as well as an observed between 50° and 150° ($p < .001$). Reaction time was not significantly different between 50°-100° rotations, $M_{diff} = 0.501s,\ SE = 0.151\text{s}, p = 0.029, d = 0.483$, and between 100°-150° rotations, $M_{diff} = 0.411\text{s},\ SE = 0.151\text{s}, p = .193, d = 0.397$. Comparable to the results from the dichotomous task, here too we observe a linear increase in reaction time across rotation magnitude. Overall, the reaction time between rotations indicate that the mental rotation performed on either task are following a similar pattern, despite being executed at a different rate.

Rotation speed. Bayesian repeated-measures ANOVA demonstrated that the best model that describes rotation speed (See Figure 3-2E) includes the main effect of task and rotation, $\log BF_{10} = 84.18, \log BF_M = 3.12$, closely followed by the model that also includes an interaction, $\log BF_{10} = 82.43, \log BF_M = -0.365$. Similarly, frequentist repeated measures ANOVA revealed a significant main effect of task, $F(1{,}46) = 5.014\,, p = .030\,, \eta^2 = 0.025$, and rotation, $F(1.20{,}55.38) = 354.974, p < .001, \eta^2 = 0.370$, but no interaction, $F(1.320{,}60.740) = 2.913\,, p = .082\,, \eta^2 = 0.003$. Post-hoc comparisons evaluating the main effect of task reveals that mental rotations are generally completed faster on the dichotomous than the continuous, $M_{diff} = 7.885°/\mathrm{s}, SE = 3.521°/\mathrm{s}$. Looking at the main effect of rotation, post-hoc comparisons reveal that participants tend to rotate objects faster as a function of rotation magnitude; that is, rotations of 150° are completed faster than 100°, $M_{diff} = 18.70°/\mathrm{s}$, $SE = 1.005°/\mathrm{s}, p < .001$, or 50°, $M_{diff} = 37.06°/\mathrm{s}$, $SE = 1.870°/\mathrm{s}, p < .001$, and that 100° rotation are completed faster than 50°, $M_{diff} = 18.36°/\mathrm{s}$, $SE = 1.138°/\mathrm{s}, p < .001$. These results suggest that participants tend to linearly increase their mental rotation abilities as a function of the magnitude of rotation to be performed.

Principal component analysis (PCA) and correlations. As in Experiment 1, I conducted a principal component analysis between the performance measures of each task (i.e., raw error for the continuous response task and %inacc for the dichotomous task) for all level of mental rotation. The parallel analysis, Kaiser rule (i.e., eigenvalue > 1) and the scree plot suggest that two components best represent the data (see component loadings in Table 3-1). These two components were computed with a an oblique Promax rotation, which allows the components to correlate with each other ($\mathrm{r}_{\mathrm{RC1-RC2}} = .447$), and they indicated that variables associated with each task are loading on separate components. Although there is a stronger case for two components, I manually computed a single component model to parallel the previous experiment. Interestingly, this single component appears to load very well on each variable, where the lowest loading of this component still accounts for 60.8% of the variance. Using these components, I computed weighted composite z-scores for each participant and calculated correlations with the VVIQ and OSIVQ scores (see Table 3-2). Results identified significant correlations ($p < .05$), with strong to extreme evidence in favour of these correlations, between all composite scores and the VVIQ and OSIVQ-spatial scores only, but not with the object and verbal subscales of the OSIVQ. Thus, as opposed to Experiment 1, I find evidence that participants are using imagery to perform these tasks and that participants who report preferring a spatial imagery strategy show better performance as well.

Discussion

This second experiment evaluated the measures of two different paradigms designed to assess working memory and mental rotation, in order to identify whether they assess the same underlying processes. I compared accuracy on a dichotomous simultaneous comparison task as well as raw error from a continuous recall adaptation of the paradigm similar to that from Experiment 1. The results revealed that participants performed equally well on either task once a mental transformation was performed, suggesting that tasks assess imagery abilities in a similar fashion. There was, however, a significant performance difference between tasks when no mental rotation was applied. This finding is likely a reflection of psychophysical differences associated with perceptual (or iconic) memory compared to working memory. Indeed, the continuous task requires participants to recall the orientation of the stimulus from memory, which is not the case for the dichotomous task. Hence, it is unsurprising that participants showed better performance when the unrotated images are presented simultaneously in the dichotomous task. Once a transformation must be performed, however, performance was similarly accurate for both tasks.

In terms of rotation performance in the dichotomous task, Ganis & Kievit (2015) showed a steadying decrease in performance as a function of rotation magnitude, whereas I observed that participants error peaks at 100°. Depending on the type of stimuli employed, a peak in performance inaccuracy at around 90° is sometimes observed in mirror/normal discrimination tasks (Cooper & Shepard, 1973; Kung & Hamm, 2010; Searle & Hamm, 2012), as an object presented at 180° can as easily be evaluated as to whether it is mirrored as when it is not rotated (i.e., 0°). This means that the comparison between items is most difficult when the object is rotated by 90°, which closely corresponds to our 100° rotation condition. Interestingly, the same pattern of results is present in the continuous task, in which error also peaks at 100°. Since in this version of the task participants cannot rely on a static referential stimulus, it is surprising that participants would perform equally well on the 50° and 150° rotation. These results suggest that perhaps participants adopt a particular strategy to aid their performance. While it is a reasonable strategy in the case of the dichotomous task to mentally rotate an object in whichever direction is closest to 0° or 180° to perform the discrimination, this strategy is not as sound for the continuous task since there is no object available for comparison. It is possible, however, that participants apply a rapid 180° rotation and backtrack to approximately by 30° to match the 150°. Therefore, the cognitive cost of completing a 180° turn and a 30° correction may closely resemble that of a direct 30° turn. The possibility of a rotation-jump strategy being employed in both tasks is supported by participants' rotation speed (see Figure 3-2E) as our analysis demonstrates that a constant rotation speed is not maintained across conditions. While reaction times (See Figure 3-2D) show that it generally takes more time to mentally rotate an object by a larger angle, looking at rotation speed shows that participants perform larger rotations at greater speed. Therefore, it appears that participants are not mentally rotating the object steadily around its axis, but instead either increase the speed as a function of the magnitude of the transformation or they make a coarse jump to the estimated target orientation and then fine tune their response.

As in Experiment 1, PCA results show strong overlap between all performance measures. Although there is greater support for a two-components model, I cannot ignore that the single component accounts for a significant amount of the shared variance among all performance measures, demonstrating that performance on each task and across manipulation is driven by a global process rather than independent mechanisms. These results support the idea that performance on both tasks is capturing similar processes. On the other hand, the two components solution shows that there is some task-specific variance that distinguishes these components, since either component is loading variables primarily associated to the separate task designs. However, these components appear to be strongly related to each other as the rotated solution shows that both components correlate significantly with one another $r_{RC1-RC2} = .447$. The two components solution shows similar results as in Experiment 1, specifically that, for each task, memory retention and imagery manipulation load together on the same component. As described in Experiment 1, here too the PCA results indicates that retention and manipulation are likely driven by a single process.

By contrast, the correlation analysis with the composite scores calculated from the PCA analysis and the VVIQ-rv and OSIVQ scales shows very contrasting results from our first experiment. Indeed, here I show strong evidence of a positive relationship between the composite scores in performance and the VVIQ-rv and the spatial subscale of the OSIVQ. These results suggest that the higher a participant rates the vividness of their imagery ability shows better performance and that participants tend to rely on spatial imagery strategies to complete these tasks. Given that the task requires participants to remember and transform the orientation of an a 2.5D object, it makes sense that individuals who engage in spatial strategies are more likely to perform well on the tasks. Although I cannot confirm which specific strategy participants are adopting, as I am using self-report measure only, it seems clear that participants who do well on these tasks rely on a more vivid imagery abilities that is mainly spatially driven. Interestingly, I can observe that neither task engages participants in a different type of memory or imagery vividness or strategy. Indeed, given that the two components solution are loading distinctly on each task and mainly reflects task-specific variance, our correlations shows that both composite scores based on the rotated components demonstrate the same relation to self-reported vividness of imagery and imagery strategy scores. These results further indicate that the dichotomous and continuous tasks ultimately assess same cognitive process. Finally, the proposal that the link between self-report measure and cognitive processes are task specific (Dean & Morris, 2003) appears to be unsupported by our results, as there are no distinction between the pattern of correlation for the composite scores based on the rotated components, which are distinctly loading on each task.

General Discussion

The primary objective of this paper was to evaluate the behavioural properties of imagery as assessed by mental rotation and compare them to those of visual working memory. Although a significant body of neural evidence suggests that imagery and working memory representations rely on the same brain areas and functional activity within these regions, very little work has been done to demonstrate that these representations are behaviourally related. Importantly, those studies that have linked memory and imagery have tended to use very different stimuli and behavioural measures of each ability (Keogh & Pearson, 2011, 2014). Here, I directly compared the type of stimulus (Experiment 1) and performance metrics obtained from different task paradigms (Experiment 2) commonly reported in these literatures.

Our results demonstrate a behavioural cost associated with the manipulation of visual representation that exists for both simple and complex stimuli, and which was observed using both a continuous and dichotomous paradigm. Thus, performing internal manipulations of mental representations reduces the accuracy or fidelity of those representations compared to those generated externally. The extent to which this cognitive cost scales as a function of the magnitude of the manipulation remains unclear; however, Experiment 1 revealed that raw error plateaued between 60° and 120°, whereas Experiment 2 showed that errors increase with the degree of rotation, peaking at 100°. One possible explanation for this difference is that participants perform 120° rotations by simply flipping the object around the axis and then performing a similar 60° rotation leading to similar performance. Examining performance by breaking down errors into guesses and correct responses using the method of Zhang and Luck (2008), however, reveals important differences. That is, for both simple and complex stimuli, precision was lower at 120° compared to 60°. Thus, our study reveals that using a continuous-report task along with a mixture-model analysis can provide important insights into the accuracy of mental rotation.

Experiment 1 also revealed overall worse performance for more complex stimuli in both memory and mental rotation tasks. Behavioural costs due to complexity have been previously reported in the working memory literature (Alvarez & Cavanagh, 2004; Robitaille & Emrich, 2019; Snow et al., 2014; Xie & Zhang, 2017; Zhou et al., 2018); to our knowledge, our results are the first to demonstrate that a similar effect of complexity exists for rotated stimuli. Indeed, with the help of the mixture model, I demonstrate that the source of the errors comes primarily from one's ability to either encode or retain the representation of a complex object, while the precision (SD^{-1}) of the mental representation was unaffected by complexity. Importantly, an interaction was observed, such that guess rates also increased for complex objects once a mental rotation was performed. Thus, not only are complex objects more difficult to encode, they may also be more susceptible to forgetting once transformations are performed on those objects. However, once those transformations are performed, they are accomplished with a similar level of precision, suggesting a common cognitive mechanism for mentally rotating simple and complex objects.

In Experiment 2, I compared the metrics associated with two experimental paradigms used to behaviourally evaluate the mental representation of visual information. While the imagery literature often relies on the percent accuracy and reaction time derived from dichotomous paradigms, in the past couple of decades, the working memory literature investigated raw error as a measure of accuracy on continuous response paradigms. By standardizing the performance scores derived from each task, I demonstrate that participants perform almost identically between mentally rotations and that both tasks produce comparable linear increase in reaction time as a function of rotation magnitude. Therefore, these results suggest that again that, regardless of the nature of the stimulus and the assessment method, these tasks assess a similar underlying mechanism.

More compelling evidence supporting the idea that the mental representations described in the imagery and the working memory literatures are behaviourally related is provided by the PCAs. Indeed, a significant proportion of variability found among the performance metrics across conditions can be effectively accounted for using a single component in each experiment (respectively 59.4% and 60.6%). These results demonstrate strong evidence in favour of overlapping cognitive processes mediating the formation and transformation of mental representations for simple or complex stimuli as well as evaluated using dichotomous or continuous paradigms. These results therefore corroborate the neuro-imaging literature, which demonstrate that the brain areas associated with the visual cortex is recruited for both working memory (Awh & Jonides, 2001; Curtis & D’Esposito, 2003; Emrich et al., 2013; Harrison & Tong, 2009; Silvanto & Cattaneo, 2010) and imagery (Cichy et al., 2012; Ishai & Sagi, 1995; Kosslyn et al., 1997; Kosslyn & Thompson, 2003; Pearson et al., 2015) and that the functional activity found in this region is also comparable (Albers et al., 2013; Christophel et al., 2015) . It is also possible that both processes involve representation within the parietal cortex (Christophel et al., 2015; Kosslyn et al., 2001; Pearson et al., 2015), although the relative contributions of sensory vs parietal and frontal contributions to working memory remain debated (Christophel et al., 2015).

Correlations between performance composite scores and the imagery questionnaires provide conflicting results. Indeed, in Experiment 1, I did not find any significant associations between the vividness or type of imagery and performance, suggesting that one's performance is not associated with the self-reported vividness of their imagery or with any of the three possible imagery strategies (i.e., verbal, object, and spatial). This is true for both a general performance measure computed from the one component PCA solution and the stimuli specific composite scores derived from the rotated solution. On the other hand, strong correlations were observed in Experiment 2 between performance and the self-reported vividness of imagery as well as spatial imagery strategy; these relationships are also found for both rotated components that represent each paradigm. It is quite puzzling to see a different pattern of correlations for the composite score based on the second rotated component (RC2) in Experiment 1, which represents the Complex stimulus, and the second rotated component (RC2) in Experiment 2, which represents the continuous composite score (RC2), as they both derived from the same task. These results indicate that perhaps self-reported imagery may not be reliable enough measure to assess the internal processes driving memory retention and manipulation. The link between self-reports and cognitive abilities has been unclear in the literature (Dean & Morris, 2003; McAvinue & Robertson, 2007). Indeed, Experiment 2 shows that performance correlations with the questionnaire scores are not specific to the dichotomous or continuous task. Given the potentially unreliable nature of self-report data, perhaps behavioral measures (such as those used here) may prove a more sensitive measure of mental imagery abilities.

To conclude, our results provide some of the first behavioral evidence that the mental representations created from working memory and imagery are rely on the same cognitive processes. Indeed, we observed that this was true independent of stimulus type, or whether imagery was assessed using a traditional mirror-rotation task, or a continuous-response paradigm more common in the visual working memory field. Thus, these results add to the findings observed in the neuroimaging literature that suggest that imagery and working memory are mediated by similar mechanisms. Although further research may be required for fully isolate the unique cognitive and neural contributions, our results provide evidence that similar paradigms can be used to assess both visual working memory and mental imagery, providing a more controlled and rigorous way to make these comparisons.

References

Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & de Lange, F. P. (2013). Shared Representations for Working Memory and Mental Imagery in Early Visual Cortex. Current Biology, 23(15), 1427–1431. https://doi.org/10.1016/j.cub.2013.05.065

Alvarez, G. A., & Cavanagh, P. (2004). The Capacity of Visual Short-Term Memory is Set Both by Visual Information Load and by Number of Objects. Psychological Science, 15(2), 106–111. https://doi.org/10.1111/j.0963-7214.2004.01502006.x

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. Trends in Cognitive Sciences, 5(3), 119–126. https://doi.org/10.1016/S1364-6613(00)01593-X

Baddeley, A. D., & Hitch, G. (1974). Working Memory. In G. H. Bower (Ed.), Psychology of Learning and Motivation (Vol. 8, pp. 47–89). Academic Press. https://doi.org/10.1016/S0079-7421(08)60452-1

Blajenkova, O., Kozhevnikov, M., & Motes, M. A. (2006). Object-spatial imagery: A new self-report imagery questionnaire. Applied Cognitive Psychology, 20(2), 239–263. https://doi.org/10.1002/acp.1182

Blazhenkova, O., & Kozhevnikov, M. (2008). The new object-spatial-verbal cognitive style model: Theory and measurement. Applied Cognitive Psychology, 23(5), 638–663. https://doi.org/10.1002/acp.1473

Brainard, D. H. (1997). The Psychophysics Toolbox. Spatial Vision, 10(4), 433–436.

Chao, Y.-S., & Wu, C.-J. (2017). Principal component-based weighted indices and a framework to evaluate indices: Results from the Medical Expenditure Panel Survey 1996 to 2011. PLoS ONE, 12(9). https://doi.org/10.1371/journal.pone.0183997

Christophel, T. B., Cichy, R. M., Hebart, M. N., & Haynes, J.-D. (2015). Parietal and early visual cortices encode working memory content across mental transformations. NeuroImage, 106, 198–206. https://doi.org/10.1016/j.neuroimage.2014.11.018

Cichy, R. M., Heinzle, J., & Haynes, J.-D. (2012). Imagery and perception share cortical representations of content and location. Cerebral Cortex (New York, N.Y.: 1991), 22(2), 372–380. https://doi.org/10.1093/cercor/bhr106

Cooper, L. A., & Shepard, R. N. (1973). Chronometric studies of the rotation of mental images. In W. G. Chase (Ed.), Visual Information Processing (pp. 75–176). Academic Press. https://doi.org/10.1016/B978-0-12-170150-5.50009-3

Curtis, C. E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. Trends in Cognitive Sciences, 7(9), 415–423. https://doi.org/10.1016/S1364-6613(03)00197-9

Dean, G. M., & Morris, P. E. (2003). The relationship between self-reports of imagery and spatial ability. British Journal of Psychology, 94(2), 245–273. https://doi.org/10.1348/000712603321661912

Delvenne, J.-F., & Bruyer, R. (2004). Does visual short - term memory store bound features? Visual Cognition, 11(1), 1 - 27. https://doi.org/10.1080/13506280344000167

Emrich, S. M., Riggall, A. C., Larocque, J. J., & Postle, B. R. (2013). Distributed patterns of activity in sensory cortex reflect the precision of multiple items maintained in visual short-term memory. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 33(15), 6516–6523. https://doi.org/10.1523/JNEUROSCI.5732-12.2013

Eng, H. Y., Chen, D., & Jiang, Y. (2005). Visual working memory for simple and complex visual stimuli. Journal of Vision, 5(8), 611–611. https://doi.org/10.1167/5.8.611

Ganis, G., & Kievit, R. (2015). A New Set of Three-Dimensional Shapes for Investigating Mental Rotation Processes: Validation Data and Stimulus Set. Journal of Open Psychology Data, 3(1), e3. https://doi.org/10.5334/jopd.ai

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. Nature, 458(7238), 632–635. https://doi.org/10.1038/nature07832

Ishai, A., & Sagi, D. (1995). Common mechanisms of visual imagery and perception. Science (New York, N.Y.), 268(5218), 1772–1774.

JASP Team. (2018). JASP (0.9) [Computer software]. https://jasp-stats.org/

Jeffreys, H. (1961). The Theory of Probability. OUP Oxford.

Keogh, R., & Pearson, J. (2011). Mental Imagery and Visual Working Memory. PLoS ONE, 6(12). https://doi.org/10.1371/journal.pone.0029221

Keogh, R., & Pearson, J. (2014). The sensory strength of voluntary visual imagery predicts visual working memory capacity. Journal of Vision, 14(12), 7–7. https://doi.org/10.1167/14.12.7

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. Perception, 36(14), 1–16.

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural Foundations of Imagery. Nature Reviews Neuroscience, 2(9), 635–642. https://doi.org/10.1038/35090055

Kosslyn, S. M., & Thompson, W. L. (2003). When is early visual cortex activated during visual mental imagery? Psychological Bulletin, 129(5), 723–746. https://doi.org/10.1037/0033-2909.129.5.723

Kosslyn, S. M., Thompson, W. L., & Alpert, N. M. (1997). Neural Systems Shared by Visual Imagery and Visual Perception: A Positron Emission Tomography Study. NeuroImage, 6(4), 320–334. https://doi.org/10.1006/nimg.1997.0295

Kung, E., & Hamm, J. P. (2010). A model of rotated mirror/normal letter discriminations. Memory & Cognition, 38(2), 206–220. https://doi.org/10.3758/MC.38.2.206

Lee, M. D., & Wagenmakers, E.-J. (2014, April). Bayesian Cognitive Modeling: A Practical Course. Cambridge Core. https://doi.org/10.1017/CBO9781139087759

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. Nature, 390(6657), 279–281. https://doi.org/10.1038/36846

Luria, R., Sessa, P., Gotler, A., Jolicoeur, P., & Dell'Acqua, R. (2010). Visual short-term memory capacity for simple and complex objects. Journal of Cognitive Neuroscience, 22(3), 496–512. https://doi.org/10.1162/jocn.2009.21214

Ma, W. J. (2018). Problematic usage of the Zhang and Luck mixture model. BioRxiv, 268961. https://doi.org/10.1101/268961

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

Maeda, Y., & Yoon, S. Y. (2013). A Meta-Analysis on Gender Differences in Mental Rotation Ability Measured by the Purdue Spatial Visualization Tests: Visualization of Rotations (PSVT:R). Educational Psychology Review, 25(1), 69–94. https://doi.org/10.1007/s10648-012-9215-x

Marks, D. F. (1973). Visual imagery differences in the recall of pictures. British Journal of Psychology (London, England: 1953), 64(1), 17–24.

McAvinue, L. P., & Robertson, I. H. (2007). Measuring Visual Imagery Ability: A Review. Imagination, Cognition and Personality, 26(3), 191–211. https://doi.org/10.2190/3515-8169-24J8-7157

McKelvie, S. J. (1995). The VVIQ as a psychometric test of individual differences in visual imagery vividness: A critical quantitative review and plea for direction. Journal of Mental Imagery, 19(3–4), 1–106.

Pearson, J., & Kosslyn, S. M. (2015). The heterogeneity of mental representation: Ending the imagery debate. Proceedings of the National Academy of Sciences, 112(33), 10089–10092. https://doi.org/10.1073/pnas.1504933112

Pearson, J., Naselaris, T., Holmes, E. A., & Kosslyn, S. M. (2015). Mental Imagery: Functional Mechanisms and Clinical Applications. Trends in Cognitive Sciences, 19(10), 590–602. https://doi.org/10.1016/j.tics.2015.08.003

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. Journal of Neuroscience Methods, 162(1), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. Spatial Vision, 10(4), 437–442.

Postle, B. R. (2006). Working Memory as an Emergent Property of the Mind and Brain. Neuroscience, 139(1), 23–38. https://doi.org/10.1016/j.neuroscience.2005.06.005

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. Psychological Bulletin, 80(1), 1–24. https://doi.org/10.1037/h0034650

Robitaille, J., & Emrich, S. (2019). Visual short-term memory for 3D shapes reveals similar properties to simple 2D features. PsyArXiv. https://doi.org/10.31234/osf.io/hp8fy

Searle, J. A., & Hamm, J. P. (2012). Individual differences in the mixture ratio of rotation and nonrotation trials during rotated mirror/normal letter discriminations. Memory & Cognition, 40(4), 594–613. https://doi.org/10.3758/s13421-011-0172-2

Serences, J. T. (2016). Neural mechanisms of information storage in visual short-term memory. Vision Research, 128, 53–67. https://doi.org/10.1016/j.visres.2016.09.010

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

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. Science (New York, N.Y.), 171(3972), 701–703.

Silvanto, J., & Cattaneo, Z. (2010). Transcranial magnetic stimulation reveals the content of visual short-term memory in the visual cortex. NeuroImage, 50(4), 1683–1689. https://doi.org/10.1016/j.neuroimage.2010.01.021

Snow, J. C., Skiba, R. M., Coleman, T. L., & Berryhill, M. E. (2014). Real-world objects are more memorable than photographs of objects. Frontiers in Human Neuroscience, 8. https://doi.org/10.3389/fnhum.2014.00837

Song, J.-H., & Jiang, Y. (2006). Visual working memory for simple and complex features: An fMRI study. NeuroImage, 30(3), 963–972. https://doi.org/10.1016/j.neuroimage.2005.10.006

Suchow, J. W., Brady, T., Fougnie, D., & Alvarez, G. A. (2013). Modeling visual working memory with the MemToolbox. Journal of Vision, 13(10). https://doi.org/10.1167/13.10.9

Suchow, J. W., Fougnie, D., Brady, T., & Alvarez, G. A. (2014). Terms of the debate on the format and structure of visual memory. Attention, Perception & Psychophysics, 76(7), 2071–2079. https://doi.org/10.3758/s13414-014-0690-7

Tong, F. (2013). Imagery and visual working memory: One and the same? Trends in Cognitive Sciences, 17(10), 489–490. https://doi.org/10.1016/j.tics.2013.08.005

van den Berg, R., Awh, E., & Ma, W. J. (2014). Factorial Comparison of Working Memory Models. Psychological Review, 121(1), 124–149. https://doi.org/10.1037/a0035234

van den Berg, R., & Ma, W. J. (2018). A resource-rational theory of set size effects in human visual working memory. ELife, 7. https://doi.org/10.7554/eLife.34963

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

Xie, W., & Zhang, W. (2017). Familiarity increases the number of remembered Pokémon in visual short-term memory. Memory & Cognition, 45(4), 677–689. https://doi.org/10.3758/s13421-016-0679-7

Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. Nature, 440(7080), 91–95. https://doi.org/10.1038/nature04262

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. Nature, 453(7192), 233–235. https://doi.org/10.1038/nature06860

Zhou, X., Mondloch, C. J., & Emrich, S. M. (2018). Encoding differences affect the number and precision of own-race versus other-race faces stored in visual working memory. Attention, Perception, & Psychophysics, 80(3), 702–712. https://doi.org/10.3758/s13414-017-1467-6

Chapter 4

Using forward-encoded raw EEG activity to track the mental manipulation of visual information

Abstract

While the working memory and imagery fields both investigate the mental representation of visual information, the neural mechanism allowing for the manipulation of these representation are still unknown. Several recent reports using brain imaging and EEG have demonstrated that these mental representations rely on the visual cortex during working memory and imagery tasks. In a previous project (Robitaille, 2016), a forward encoding model was applied to alpha band (8-15Hz) EEG activity recorded over visual cortex, and showed that the orientation of a stimulus is reliably detected during retention, but not after a mental rotation was applied. In an attempt to track the manipulation of the mental representation, the current study extends the work of Robitaille (2016) by applying the same encoding model to raw EEG activity (.01 – 40 Hz), which has been shown to track orientation information better. Participants were instructed to remember the orientation of Gabor patches, and were cued to mentally rotate the Gabor but 60° clockwise or counter-clockwise, and report the orientation after rotation. I also introduced a task-irrelevant stimulus (i.e., impulse) to probe "activity-silent" activity either before or after subjects were cued to perform the mental rotation. The results show that it is possible to identify the orientation of the Gabor from raw EEG activity during perception, but the activity brought back by the impulse is not reliable enough to clearly track the orientation during retention and manipulation.

Keywords: Working memory, imagery, mental representation, encoding model, EEG

Introduction

Working memory is a core cognitive function that has the property to flexibly manipulate information (A. Baddeley, 2003). While the earlier conceptualizations of short-term memory was almost exclusively described as a placeholder for information, the working memory model, as described by Baddeley and Hitch (1974), emphasizes the ability to manipulate information. Their model proposes that a mechanism referred to as the central executive would mediate between the active processes of maintaining and manipulating information. While significant progress has been made in describing the nature of the processes that support the retention of information in working memory, the neural mechanisms surrounding its manipulation is still under-examined.

As the working memory literature focused on the psychophysical properties of memory retention, the field of mental imagery has studied extensively the transformation of internally generated visual representations (Tong, 2013). To describe the format of these visual representations, two theories have emerged. According to the depictive (or pictorial) theory of mental imagery, our mental representation of an object shares properties with physical objects. As such, these representations will exhibit properties that are consistent with the manipulation of physical objects. For instance, Shepard and Metzler (1971) demonstrated a direct linear relationship between the time participant required to perform a mental rotation and the magnitude of the rotation. While it takes longer to physically rotate an object by 180° than 90°, the same is true to mental rotations. Kosslyn (1973) showed a similar relationship when participants were asked to scan through elements of a memorized map, revealing that scanning time increased as the distance between items was greater. Therefore, these results illustrate that our mental representations retain features of the physical objects, particularly the spatial relationship between its features.

On the other hand, the propositional theory of mental imagery suggests that representations are stored in a language-like format and that the visual experience associated with imagery is epiphenomenal and serves no functional purpose. For instance, the same relationship found by Shepard and Metzler (1971) is true for congenitally blind individuals (Marmor & Zaback, 1976). Similarly, Fleming et al. (2006) showed no performance difference between sighted, blind, and blindfolded individuals in a task that required participants to spatially organize items according to a set of verbal instructions. While there is more evidence provided in favour of the depictive theory (see Kosslyn, 1996), the debate around what format imagery representations take is still unsettled, in part because it has depended primarily on behavioural evidence. However, a significant body of evidence in favour of the depictive format has also stemmed from neural imaging research. Using univariate analyses, several studies have reported that cortical areas involved during visual perception, mainly in the occipital lobe, are also recruited during imagery (Cichy et al., 2012; Ishai & Sagi, 1995; Kosslyn et al., 2001; Pearson et al., 2015), which has led to the idea that imagery can be thought of as a form of a "weak perception" (Pearson et al., 2015).

This association between perception and imagery provides an interesting parallel with the sensory recruitment hypothesis developed within the working memory field. This hypothesis describes how working memory maintenance reactivates the same cortical areas that were needed to encode the sensory information in the first place (Postle, 2006; Serences, 2016; Serences et al., 2009). Indeed, while the imagery and working memory literatures have struggled to link their subject of study with one another due to a divergence in methodology (Tong, 2013) there is a growing body of evidence suggesting that working memory and imagery representations rely on similar brain areas (e.g., Albers et al., 2013).

In particular, the application of multivariate pattern analysis (MVPA) in brain imaging studies has provided substantial evidence in support of the sensory recruitment hypothesis in the working memory field, by identifying how working memory content is represented within neural activity. Specifically, studies have demonstrated that visual working memory representations can be tracked from activity recorded from the occipital lobe, whereas the frontal cortex does not appear to reliably encode working memory contents, as univariate analyses had previously suggested (Serences & Saproo, 2012). Because it was formerly believed that working memory maintenance should produce sustained spiking activity, the increase in fMRI BOLD signals in the prefrontal cortex seemed to reflect this phenomenon very well (Chaudhuri & Fiete, 2016; Costa et al., 2015; Eriksson et al., 2015; also see Fuster & Alexander, 1971). Multivariate analysis of brain imaging data, however, revealed that the dynamic pattern of activity found within visual cortex is more reflective of the feature-selective neural response associated with the memory representation (Albers et al., 2013; Emrich et al., 2013; Harrison & Tong, 2009; Haynes & Rees, 2005; Kamitani & Tong, 2005). These methods are believed to capture slight changes in the modulation of columnar activity in the early visual cortex (Cichy et al., 2015; Serences & Saproo, 2012), thus allowing to track specific features (e.g., orientation, colour, etc.) associated with the visual representation.

Based on the results of these findings in the working memory field, it is possible that these multivariate analyses could provide a means to investigate the manipulation of mental representation of visual information – specifically, by providing a direct test of the depictive format theory of imagery. For instance, Harrison and Tong (2009) were able to decode working memory representations of grating patterns using MVPA on BOLD signals recorded from the visual cortex well beyond the presentation of the visual stimulus. They trained a pattern classifier to recognize the BOLD activity recorded at in the visual cortex during the perception of grating patterns. In a delayed response task, participants were required to remember two grating patterns presented sequentially and, after a short delay, they were cued to report one of the stimuli. The pattern classifier was able to identify the activity of both grating patterns during the retention period and identify only the cued stimulus when participants dropped one from memory. Extending these findings, Albers et al. (2013) trained a classifier to recognize three grating patterns, and they were able to track the orientation of the representation as participants performed a mental rotation of the original stimulus. These results demonstrate that not only the maintenance of working memory representations can be detected in early visual cortex, but also that the manipulation of working memory representation seems to similarly recruit sensory cortex.

The adaptation of multivariate analyses to EEG data (e.g., decoding models, forward or inverse encoding models) has also provided a means to track feature-selection patterns of activity with a better temporal resolution (Brouwer & Heeger, 2009). These multivariate models are sensitive to the small changes in the pattern of activation, which can reveal the representational content of a cortical activity based on the pattern of activity recorded (Brouwer & Heeger, 2009; Harrison & Tong, 2009; Haynes & Rees, 2005; Serences & Saproo, 2012). Encoding models in particular make *a priori* assumptions about the pattern of activity associated with the variation of a feature that varies continuously, such as orientation and colour. These assumptions are grounded in the well-establish feature-selective pattern of neural activation (Carandini et al., 2005). Previous studies have used these forward encoding models applied to EEG activity to track working memory features such as spatial location (Foster et al., 2015; Samaha et al., 2016; Sprague et al., 2016), colour (Hajonides et al., 2021), and orientation (Garcia et al., 2013). Therefore, it is possible that these models could be used to evaluate how the neural representations are affected by manipulations (i.e., mental rotations) and track these changes.

In a previous experiment (Robitaille, 2016), I applied a forward encoding model, which uses the continuous nature of a feature (i.e., orientation) and neural population code to model brain activity, on the induced and evoked alpha (8-12Hz) activity in a working memory and imagery manipulation paradigm. The forward encoding model (described in detail in the method section) applies an *a priori* assumption, which is grounded in the appropriate neural population code, about the pattern of brain activity generated by the presentation of a feature. Then this assumption is evaluated using brain activity and the encoding model assesses how consistent this pattern is present in the data. The encoding model was applied to data collected from a task that consisted of presenting a line stimulus for 250ms, followed by a retention period lasting 500ms where the fixation point changed colour to indicate to the participant to mentally rotate the orientation of the stimulus by 60° clockwise or counter-clockwise (Figure 4-1). The primary objective of this study was to replicate previous reports (in particular Foster et al., 2015 and Garcia et al., 2013) and demonstrate that the orientation of a stimulus can be identified during the presentation and retention periods by applying an encoding model to the EEG activity recorded at the posterior sites (i.e., P1/2, P3/4, P5/6, PO3/4, PO7/8, O1/2, PZ, POZ, OZ, and IZ based on the extended 10-20 electrode placement system). Consistent with previous reports, encoding induced alpha activity (see Figure 4-1B.), but not evoked alpha activity (see Figure 4-1A.), revealed sustained stimulus encoding between the onset of the stimulus and slightly beyond the imagery cue onset. These results demonstrate that there is reliable information present in induced alpha activity to detect the orientation of the stimulus presented. The second objective was to determine whether the encoding model could track the mental representation as participants performed a mental rotation. The results of the encoding model revealed that the current design did not allow the identification of the rotated orientation, but it demonstrated that that the original orientation is significantly “de-selected” in the induced alpha activity; in other words, the encoding model showed reliable activity against the original orientation. These results suggest that, after the memory manipulation, participants no longer held the original orientation in memory, but it was not possible to identify the mentally transformed orientation.

While most reports that successfully reconstructed the feature of a stimulus in memory with alpha activity tracked spatial position (Foster et al., 2015; Samaha et al., 2016; Sprague et al., 2016) and not orientation, some authors suggest that alpha band activity may not track orientation information as well was initially believed (Garcia et al., 2013; Stokes, 2015; Wolff et al., 2015, 2017). However, some reports that successfully tracked orientation held in working memory examined raw EEG instead of limiting the frequency range (Wolff et al., 2015, 2017). In an exploratory analysis using the data from this data (Robitaille, 2016), I applied the encoding model to raw EEG activity (0.1 – 40 Hz band pass filter) instead of restricting the activity to the alpha band only (see Figure 4-1C.). The results of this analysis showed comparable results to the induced alpha activity, where we can see a clear selectivity of the orientation held in memory shortly after the onset of the stimulus. Unlike the encoding of induced alpha activity, the selectivity is not sustained but instead decreases steadily until the onset of the imagery cue. In their reports, Wolff et al. (2015, 2017) showed a similar pattern when encoding raw EEG activity, where there appeared to be insufficient information to detect the orientation of the stimulus held in memory.

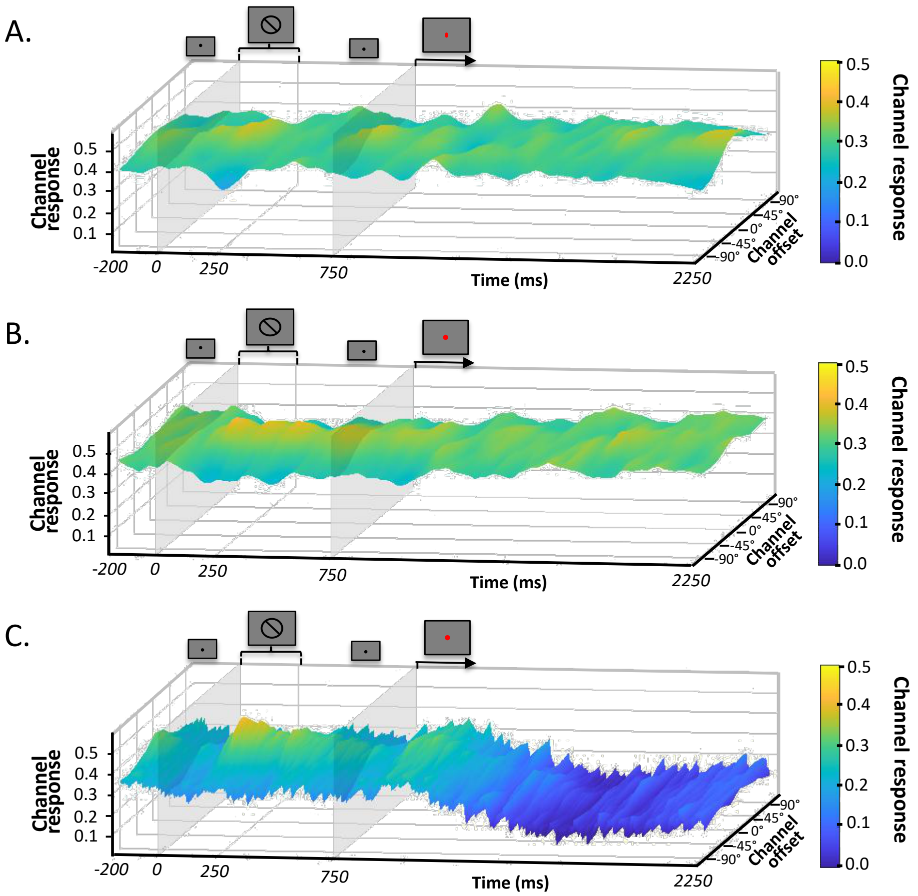

Figure 4-1. Channel Tuning Functions (CTFs) from Robitaille (2016). Channel tuning functions (CTFs) across time demonstrating orientation selectivity for (A.) evoked and (B.) induced alpha activity from Robitaille (2016) and an exploratory analysis using (C.) raw EEG activity. Greater channel response at the 0° channel offset represents successful detection and recreation of the orientation feature. The time course of this experiment (showed above the CTFs) consisted of line stimulus appearing for 250ms, followed by a fixation point for 500ms, and a coloured cue indicating the direction of the mental rotation to be performed (i.e., red for a counter-clockwise rotation and blue for a clockwise rotation). The grey windows represent the onset of the stimulus at 0ms and the imagery cue at 750ms.

While at first glance it seems that there is detectible activity associated with the mental representation of orientation in the exploratory encoding of raw EEG activity from Robitaille (2016), Stokes (2015) proposed that the assumption that working memory maintenance relies on the sustained neural activation may not be well founded. Indeed, neural population-level analyses suggest that activity states are more dynamic than previously thought. The relevant information associated with the representation might not be constantly active during memory maintenance, but rather resides in the functional connectivity or the structural connectivity of synapses (see Stokes, 2015 for more details). Therefore, memory representations would be maintained in an “activity-silent” state waiting to be re-activated. Based on Stokes (2015) framework, Wolff et al. (2015, 2017) demonstrate that these “activity-silent” states can be reactivated and detected by presenting a task-irrelevant visual impulse (i.e., ‘ping’). A glimmer of this effect is present in Robitaille’s (2016) data depicted in Figure 4-1C, as there is an increase in selectivity resurging briefly after the onset of the imagery cue. Although mental representations associated with memory retention and imagery manipulation might go into a “silent-state” that is difficult to detect from raw EEG activity, pinging the visual cortex during the memory and manipulation delays might bring the activity online.

To summarize, given that previous reports using MVPA applied to fMRI BOLD signals show that WM representations during retention and manipulation relies on the primary visual cortex and appear to rely on the same functional activity, and that we can track features of a memoranda from EEG activity using similar techniques, it should also be possible to track these features as participants are mental manipulating them. In this current study, I attempt to probe the mental representation of grating patterns while participants perform a mental rotation of the object’s orientation. As was performed in previous work (Garcia et al., 2013; Stokes, 2015; Wolff et al., 2015, 2017), I applied a forward encoding model on EEG activity recorded over the occipital lobe with the intent to track the working memory maintenance and transformation. I incorporated a high-contrast stimulus designed to bring forward the “activity-silent” state that is thought to contain relevant information about the orientation feature of the mental representation, as demonstrated by Wolff et al. (2015, 2017). The paradigm remained essentially the same as described in Robitaille (2016) except with greater delays during retention and transformation, which allows for the introduction of a task-irrelevant ‘impulse’ stimulus during either delay to “ping” the orientation feature present in a silent-activity state.

Methods

Participants. A total of 23 Brock University students (10 males: 13 females) were recruited through posters and Psychology Department Research Pool (i.e., SONA) system. Participants were first screened via a short phone interview to assess their eligibility in participating in an EEG experiment. This phone screening included a short description of the experiment as well as questions about participant’s handedness as well as whether they had a history that includes psychiatric/mental health, head injury/concussion/loss of consciousness, and any condition that may affect their nervous system. Participants that qualified for the study and completed the study were compensated with either course credit or $15/hr.

Apparatus and EEG Acquisition. The experimental paradigm was displayed on a 20-in LCD monitor with a resolution of 1600 x 1200 and a refresh rate of 60 Hz. Participants were seated about 60 cm from the screen and completed the task using a mouse and a keyboard. Electroencephalography was recorded using a BioSEMI Active-II system with built-in amplified Ag-AgCl electrodes. Participants were fitted with an EEG cap following the extended 10-20 configuration for 64 electrodes with online CMS/DRL references. Two additional electrodes were placed on the bone behind the ear (mastoids) to serve as references. Two additional electrodes were placed ~1cm lateral to the external canthi, and two were placed below each orbit, which served as the horizontal and vertical extraoculogram (VEOG and HEOG) electrodes to detect lateral eye movements and eye blinks.

Stimulus and Paradigm. To engage the visual cortex maximally, I opted to use a Gabor patch as previous work has been demonstrated that this type of stimulus maximally activates the primary visual cortex response (Thompson & Kosslyn, 2000). Moreover, this stimulus replicates that of a previous study using the same type of stimulus, which applied the impulse stimulus to examine activity-silent representation of WM representations (Wolff et al., 2015, 2017). A Gabor patch of 14 visual degrees and 0.65 cycles/deg (sine wave) was generated as using the *GratingStim()* function of the PsychoPy software (Peirce, 2007). To avoid participants relying on sharp edges around the stimulus and to maximize the orientation feature, a Gaussian alpha mask was applied using the default *GratingStim()* mask parameters.

Each trial presented a single Gabor patch at a random orientation (see Figure 4-2). To ensure that the full spectrum of orientations was represented equally across the task and between participants, the orientation of the Gabor patch was sampled from one of eight orientation bins equidistant from one another spanning the full 180° range. These bins were centered around 11.5°, 33.75°, 56.25°, 78.25°, 101.25°, 123.75, 146.25°, and 168.75°. In a given trial block, each of these bins are sampled 8 times (for a total of 64 trials per block) and, in a given trial, a random orientation was selected in a ±11.5° window around the center of the orientation bin.

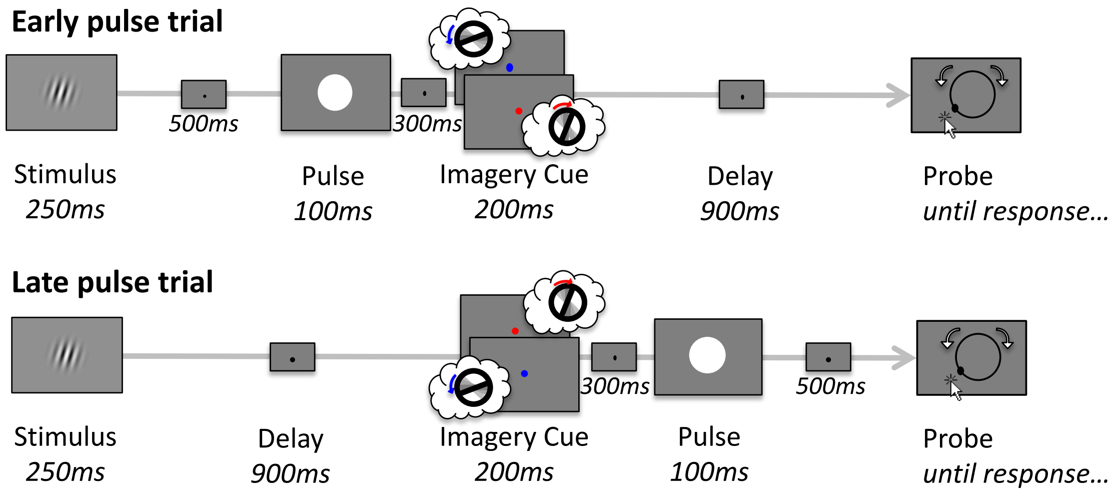

Figure 4-2. A schematic of early and late impulse trials of the behavioural paradigm. Both types of trials begin with a fixation point presented for 1,000 ± 500ms. The sample stimulus, a Gabor patch, is then presented for 250ms, followed by a fixation point for the delay window of 900ms. After the initial working memory delay period, participants were cued, via a change in the colour of the fixation, on the direction of the mental manipulation to be performed. Participants were instructed that a red fixation points requires them to perform a clockwise mental rotation of 60°, whereas a blue fixation points indicate a clockwise mental rotation of 60°. The rotation cue was present for 200ms followed by a second 900ms delay. Finally, a probe window appears that required the participants to report the orientation of the transformed stimulus using a mouse click. The distinction between early and late impulse trials is the apparition of a task irrelevant stimulus (i.e., white circle) for 100ms during either the first or second delay respectively. For the early impulse trial, the delay following the stimulus is separated into an initial 500ms fixation, then the 100ms impulse stimulus, followed by 300ms fixation, while the second delay span 900ms without interruption. Similarly, in the late impulse trials, the first delay is uninterrupted and spans 900ms whereas the second delay is interrupted by the 100ms impulse stimulus between an initial 300ms fixation window and followed by a 500ms fixation before the probe window.

The task began with the presentation of a Gabor patch for 250ms, followed by a 900 ms fixation point, and an imagery cue for 200 ms. Participants were instructed to remember the orientation of the Gabor patch, and to mentally rotate the Gabor patch by 60° after the cue appears. The imagery cue was a change in the fixation point colour, where red indicated a clockwise (CW) rotation and blue indicated a counter-clockwise (CC) rotation. After the imagery cue, there was a 900ms delay, where the fixation point returned to black, followed by the probe window, where participants could report the resulting orientation after the mental rotation using the mouse. Participants submitted their response with a click of the mouse.

Procedure. Upon arrival, participants were required to complete a consent form as well as demographic information. After the EEG cap and electrodes were applied to the participants, they were trained on a series of practice trials (1-2 blocks of 32 trials) until they were able to understand the task instructions. The task was broken into 15 blocks of 64 trials for a total of 960 trials, and participants were allowed to take a short break between blocks. A longer break was offered halfway through the task the participant was offered a drink and light snack. Participants were given course credit and/or monetary compensation for their time at the end of the experiment. The entire experiment took ~2hrs to complete.

Behavioural Data. For every trial, participants were required to report the orientation of the stimulus presented after performing a mental rotation of 60°. It is important to note that the orientation of a Gabor patch can only span 180°. For each condition, I used raw error as the primary measure of behavioural performance, which was defined as the standard deviation of the absolute angular distance between the participant's response and the target orientation. This measure has shown to be a reliable measure of working memory performance (see van den Berg & Ma, 2018, about summary statistics; Ma, 2018). To ensure that the EEG recordings reflected the proper direction of the mental rotation, trials that were beyond ±30° from the target rotation were rejected. Three participants were rejected from the study as more than 25% of the trials were rejected (i.e., 28.02%, 30.10%, 56.15%). Therefore, a mean of 13.17% (SD = 3.378%) of trials were rejected for the remaining participants.

EEG Data Preprocessing. Independent component analysis (ICA) was performed on every participant's raw EEG recordings to identify and remove sources of noise before applying the encoding model. ICA was completed through the *pop_runica()* function of EEGlab and, with the assistance of the ICLabels EEGlab extension, and each component extracted was visually inspected to reject muscle, eye, line, and channel noise. An average of 5.31 ICs were removed, where the average global offset from each component is about .4277.

Using the EEG toolbox artefact rejection functions, I also performed eye-blink and movement artefact rejection using EOG electrodes. Trials that were flagged when a peak-to-peak difference moving window identified activity on the VEOG that exceeds 80μV as well as the identification of square waves that exceeds 32μV on the HEOG channel. These flagged trials were visual inspected to confirm that they exhibited eye movements. A mean of 1.53% of trials (SD = 1.03%) per participant were rejected due to eye movement contamination. Along with the behavioural rejection, a total mean of 14.45% of trials (SD = 3.65) were rejected per participants.

Following artefact rejection, the subsequent analyses were performed on the 16 posterior electrodes (i.e., P1/2, P3/4, P5/6, PO3/4, PO7/8, O1/2, PZ, POZ, OZ, and IZ based on the extended 10-20 electrode placement system), which were re-referenced off-line to the left and right mastoids electrodes and filtered using a high pass of .01Hz and low pass of 60Hz.

Encoding Model

To attempt to track the mental representation during memory retention and its manipulation, I applied an encoding model using raw EEG activity. This encoding model (see Figure 4-3) is based on the principle that the activity recorded by a given electrode is going to detect activity that is more predictive of one variation of a feature as opposed to deviation from this variation. The encoding model makes *a priori* assumptions about the pattern of activity recorded at each site as a function of the variation of the feature – in our case, orientation. Each electrode site is given multiple feature (tuning) channels, one for each variation of the feature (i.e., 8 orientations), that track the neural activity associated with each variation. The *a priori* assumption applied to these channels is grounded in the neural population coding distribution, where a channel will respond differently to a specific variation of the orientation feature and will decreasingly respond as the feature moves away from its "preferred" orientation (see Figure 4-3a.). Therefore, at each electrode site, tuning channels will respond differently to each of the 8 orientation bins to track every possible variation of the orientation feature. Our hypothetical tuning function (R), which is our prediction of how the tuning channels will respond to the orientation feature (θ), is depicted by a basis set that follows a half-sinusoid function:

$$R = \sin\left(\frac{\theta}{2}\right)^7$$

Because I am using 8 tuning functions to track each orientation bin – each of which peaking maximally at their preferred orientation – the equation above is raised to the 7th power. As a result, it narrows the curve and allows the accommodation of 8 tuning functions while providing meaningful overlap between neighboring tuning curves (see Figure 4-3b.).

Using recorded EEG activity, we can apply this prediction to a training set and extract weights (see Figure 4-3g.). These weights serve as the link between the predicted response and the recorded activity. This relationship can be depicted by the following formula:

$$B_1 = WC_1$$

Where the data from a training set (B_1: m electrodes X n observations) is the result of a weight matrix (W: m electrodes $*$ k channels) and our predicted channel response (C_1: k channels $*$ n observations). We can obtain and estimate of the weight matrix by applying a regression analysis and finding the least-square solution through matrix division:

$$\widehat{W} = B_1 C_1^T (C_1 C_1^T)^{-1}$$

If the activity recorded for the training set has enough reliable information associated with the orientation feature of the stimulus, then applying the weighted matrix to a test set (B_2) that is similar to the training set will estimate channel responses (C_2) and should be able to recreate the predicted pattern of activity:

$$\hat{C}_2 = \left(\widehat{W}^T \widehat{W}\right)^{-1} \widehat{W}^T B_2$$

The training and test sets were created by randomly assigning trials into three blocks, where two of these blocks serve as the training data and the last block is used as the testing set. To avoid sampling bias, all possible combinations of these blocks assigned to the training or test sets were evaluated, and the random assignment of trials to blocks was repeated for 10 iterations as well as at every timepoints. The results of these three training/testing combination blocks and ten block resampling iterations were then averaged.

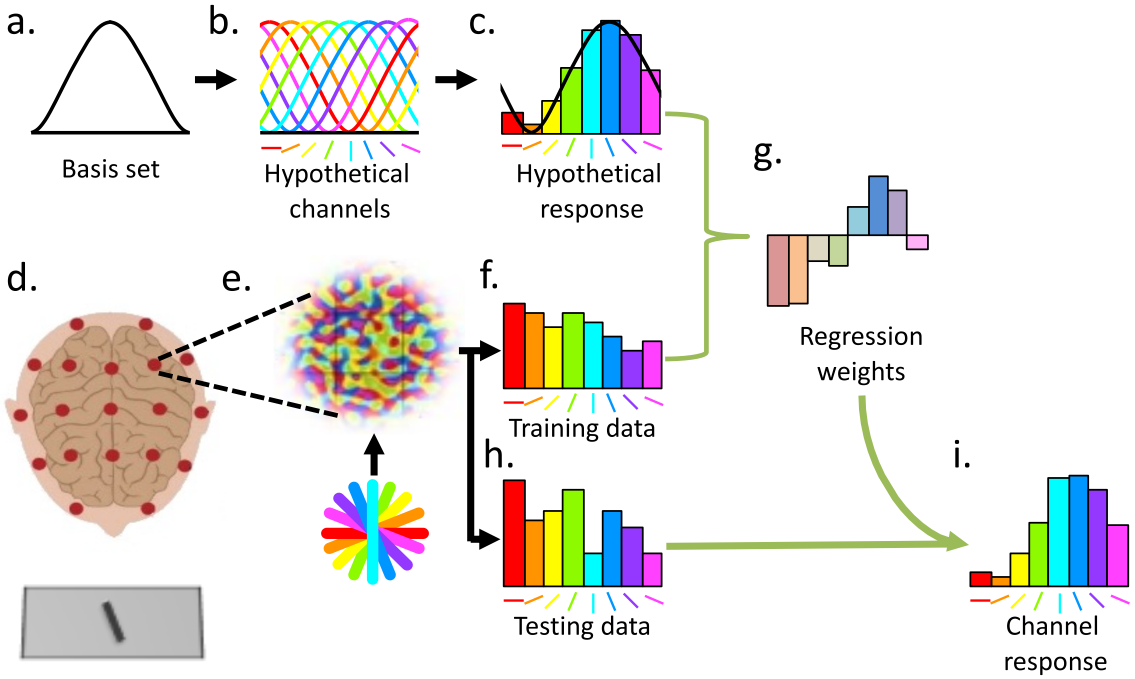

Figure 4-3. A schematic representation of the forward encoding model. (a.) The expected channel response is predicted to follow a half-sinusoid function as a basis set, which is characterized by the greatest response to a specific orientation and a progressively reduced response when moving away from the selected orientation. (b.) A series of 8 hypothetical channels are created to track each of the 8 orientation bins (each assigned a different colour) and follows the basis set. (c.) Given the presentation of a particular stimulus orientation, the hypothetical (or expected) response to the presentation would depict the basis set to be centered around the presented orientation. (d.) During the presentation of stimulus, every electrode site will record cortical activity, which (e.) contains overlapping areas that produces feature-sensitive activity. To train the encoding model, regression is applied (i.e., matrix division) to the hypothetical channel response and (f.) training data, to obtain (g.) weights. These weights are then applied to the (h.) testing set to produce the (i.) channel response.

To simplify the results and increase the strength of the effect, all orientation bins presented were recentered to a common channel offset, where 0° corresponds to the orientation bins of the stimulus presented at the beginning of the trial and any deviation from this orientation is depicted as an orientation offset that ranges between -90° to +90° - which covers the full 180° span of line orientations. This leaves us with four types of trials that combines the two types of impulse stimulus (i.e., early and late) and clockwise (CW) or counter-clockwise (CC) imagery rotation. To further combine trials, the channel tuning functions (CTFs) of the CC 60° mental rotations trials were inversed to correspond to the pattern of CW 60° so both conditions can be displayed as one. Therefore, I have combined all orientation and mental rotation conditions in a way that retains their effect on the CTFs, while presenting the late and early impulse trials on separate plots.

Previous reports have evaluated the performance of the encoding model by measuring the slope of the estimated channel responses (see Foster et al., 2016). Here, I opted for receiver operating characteristics (ROC) as a method to classify and evaluate the reliability of the channel tuning function to detect the target orientation (Fawcett, 2004). I assigned the responses of each tunning channel to different classes, where one class consisted of the value of the channel response for the target orientation while the rest were assigned to a separate class. By obtaining the true-positive and false-positive rates of this binary classification, I computed the area under the ROC curve (AUC). I performed two separate ROC and AUC procedures to track the orientation of the stimulus that was initially presented as well as the target orientation (after the mental rotation is completed). Using this approach, the aim is to attempt to identify which orientation is present in memory at any point during the trial.

To statistically evaluate the performance of the encoding model over the course of the trial, I used a Monte Carlo procedure to create a null distribution. This null distribution was generated by running the same encoding model procedure as described above 1,000 times with randomized orientation labels. Then, I computed t-statistics that represent the comparison between the real AUC performance to the null AUC.

Results

Behavioural Performance

Since participants were required to report the exact orientation of the stimulus presented at the beginning of the trial followed by a mental rotation of 60°, I computed the angular difference between the reported angle and the target orientation for each trial. Then, for each condition (2 levels of mental rotation direction and 2 levels of impulse timing), raw error was calculated as a measure of performance by taking the standard deviation of the absolute angular distance (see Figure 4-4.). A frequentist and Bayesian ANOVA were conducted to determine whether there was any difference between conditions. A significant main effect of rotation direction was present, $F(1,18) = 1.125, p = .025, \eta^2 = 0.028, BF_M = 8.075, BF_{10} = 7.945$, where post-hoc comparisons revealed better performance for CW rotations, $M_{diff} = 0.523°, SE = 0.213°, p = .025, d = 0.562, BF_{10,U} = 3.032$. Note that this is a very small, but significant, difference in favour of CW over CC rotations and only anecdotal evidence is found in support of this difference. There was no significant or evidence for the effect of impulse timing (i.e., early vs late), $F(1,18) = 0.011, p = .373, \eta^2 = 9.467e^{-5}, BF_M = 0.093, BF_{10} = 0.369$, or interaction between rotation direction and impulse timing, $F(1,18) = 0.104, p = .994, \eta^2 = 0.001, BF_M = 0.224, BF_{10} = 0.630$.

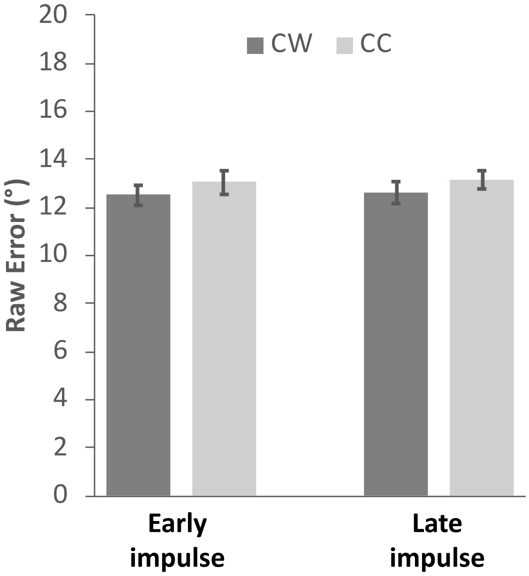

Figure 4-4. Behavioural performance across conditions. Performance results for each rotation (CW: clockwise in dark grey, CC: counterclockwise in light grey) and impulse (early on the left and late on the right) for mean (A.) raw error, (B.) guess rate, and (C.) precision. Error bars represent standard error of the mean.

Encoding Model

The channel tuning functions (CTFs) of the resulting encoding analysis, applied to the raw EEG activity for the early and late impulse trials (see Figure 4-5) reveals an increase in channel response following the presentation of the Gabor patch at 0 ms, and is sustained past the display of the stimulus for a couple of hundred milliseconds. The channel response shows a maximal response at the 0° channel offset corresponding to the orientation that was presented. At first glance, these results suggest that the encoding model is successfully reconstructing the activity associated with the orientation from the EEG activity until the model cannot find reliable information, that is before 750ms in Figure 4-5A and until 1150ms in Figure 4-5B. In Figure 4-5A, the appearance of the impulse stimulus generates a comparable channel response to the presentation of Gabor patch that peaks maximally at 0°. This is particularly interesting given that the impulse does not carry any orientation features, and that the absence of this impulse at the same time latency in Figure 4-5B shows reduced channel response with no sustained maximal peak. These results support the idea that silent-state can be reactivated using the impulse stimulus as a “ping” (Stokes, 2015; Wolff et al., 2015, 2017).

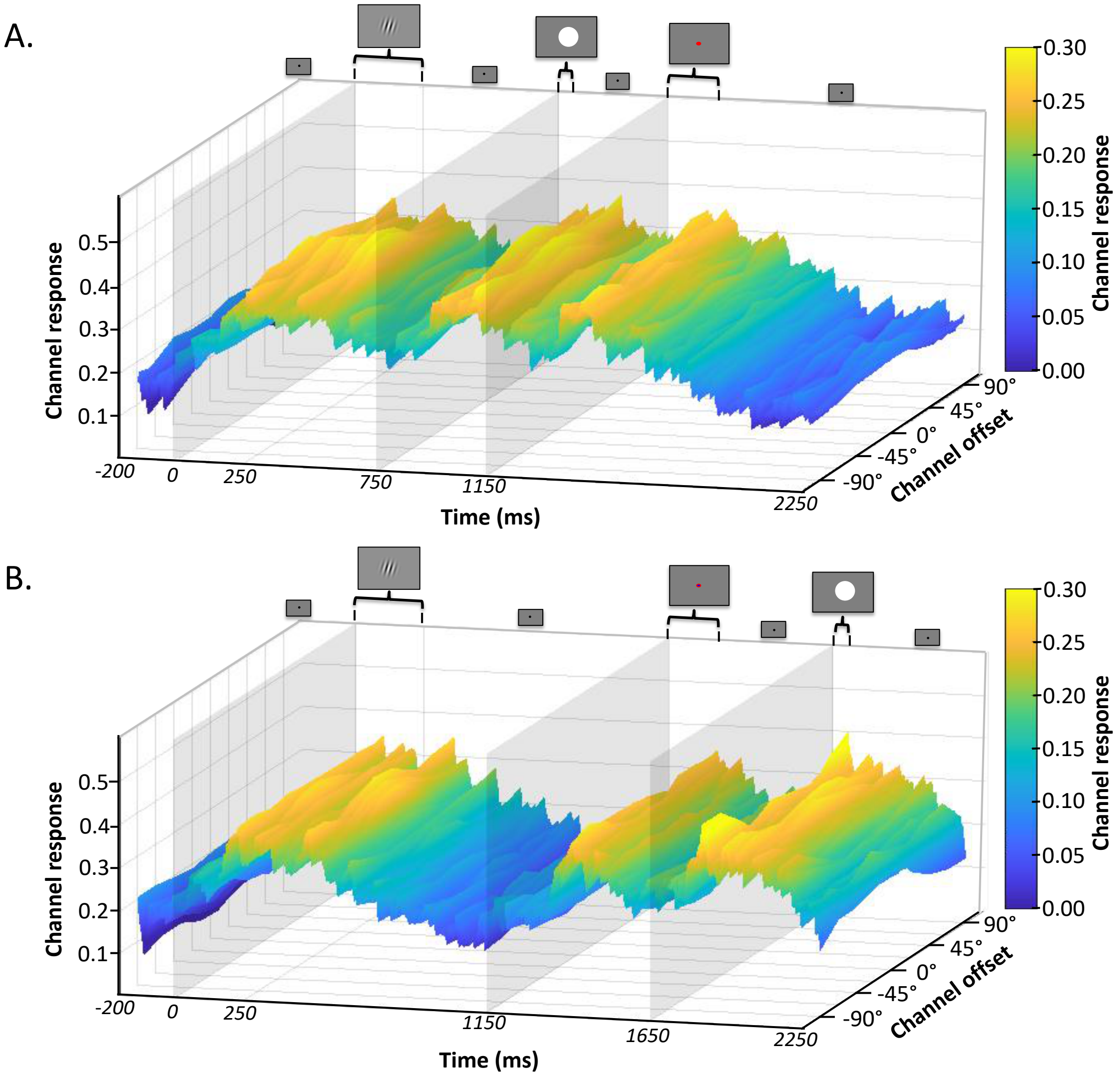

Figure 4-5. Orientation selectivity of raw EEG activity over time (x-axis) averaged across participants for (A.) early or (B.) late impulse trials, where the impulse appears at the 750ms or 1650ms mark respectively. The y-axis and colour map depicts the strength of the channel response, where greater response signifies greater selection with respect to the corresponding angle offset on the z-axis (i.e., depth). On the z-axis, 0° represents the orientation of the target as it was presented at the beginning of the trial where selectivity for other orientations is represented by the channel offset. The shaded windows highlight the significant event in the trials, namely the onset of the stimulus, rotation cue and impulse stimulus, where the latter occurs either between the stimulus and the rotation cue in the early impulse condition or after the rotation cue in the late impulse condition.

Then again, the presentation of the mental rotation cue at 1150 ms in both impulse conditions is followed by an increase in channel responses. This time, however, in neither early nor late impulse conditions can we observe a discernible peak that indicate that the information is associated with a particular orientation. During this time of the trial, participants are instructed to mentally rotate the Gabor patch by 60°, which may explain why the activity found in the raw EEG activity may not be reliable enough for the encoding model to reconstruct the CTFs that exhibit a specific orientation. While the channel response decreases as the trial comes to an end in the early impulse condition (Figure 5A), the late impulse condition (Figure 5B) shows an increase in channel response following the impulse stimulus. In this case, however, we can observe a clear maximal peak between -90° and -45° which may correspond to an orientation offset of 60°. This maximal peak is short-lived, however, as we can discern maximal peaks close the 0° orientation shortly after that sustained for a longer period. Therefore, it appears that the encoding model can detect the changes in orientation from raw EEG activity if they are reactivated using an impulse stimulus.

To determine whether these effects are performing above chance, ROC classification was applied to the CTFs at every time point for both early and late impulse trials. As a measure of classification performance, I obtained the AUC where the results are depicted in Figure 4-6. The shaded coloured regions correspond to the 95% confidence of the AUC. Since the classification accuracy shows quick changes over time, I applied a Savitzky-Golay filter, which provides a more defined trend without changing the signal tendency (Savitzky & Golay, 1964). This filtered data is represented as thick coloured lines, where blue represents the detection of the orientation of the initial stimulus whereas the red line tacks the orientation of the transformed orientation. In both early and late impulse trials, the CTFs appears to predictably detect the orientation of the Gabor patch as the stimulus is presented on the screen between 0-250ms. This effect is confirmed through statistically significant deviation for both the blue and red line, where the orientation of the initial stimulus is more reliably detected above chance levels whereas the detection of the transformed orientation is performing worse than chance. These results indicate that the encoding model can accurately detect the orientation of the stimulus from the activity in detected on posterior electrodes.

As is seen in Figure 4-6A, classification generally detects the presence of the orientation of the initial Gabor patch over the transformed one throughout the entire trial. It does not appear that the transformed orientation is significantly present in the signal past the mental rotation (i.e., past 1150 ms), however. Instead, there are greater detection of the initial rotation during the imagery phase of the trial. Conversely, in Figure 4-6B, it seems that the encoding model is not reliably detecting either rotation past the presentation of the Gabor patch (i.e., past ~300 ms). There is some significant detection of the transformed orientation at the onset of the mental rotation, indicating that the encoding model is detecting the transformation of the stimulus. Despite showing greater CTFs response, the presentation of the impulse stimulus did not bring information that allows for the detection of the orientation held in memory in either the early or late impulse trials.

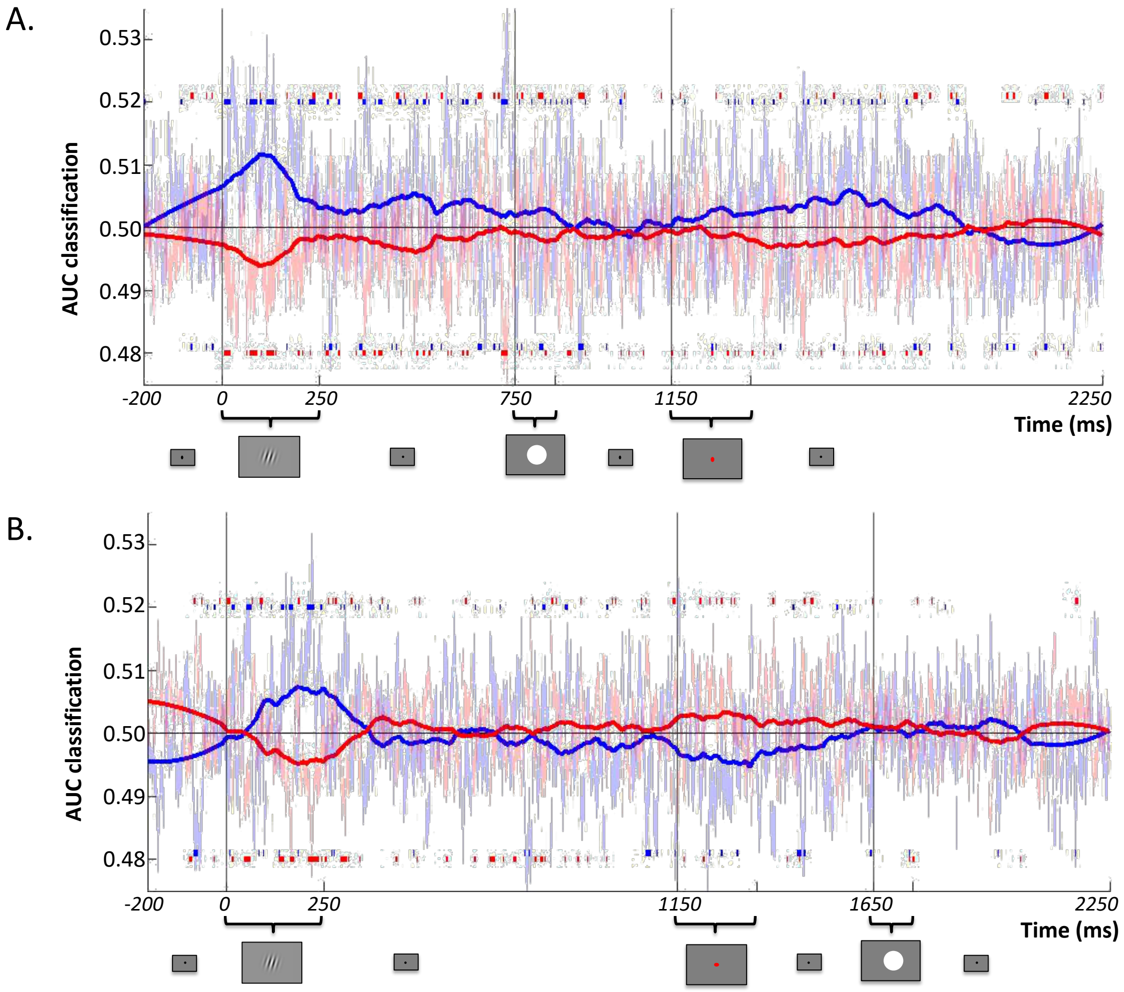

Figure 4-6. AUCs of the CTFs ROC classification over time tracking the orientation of the original stimulus (blue) and the rotated orientation (red) for the (A.) early and (B.) late trials. Shaded regions represent the real classification value with a 95% confidence interval. Thicker coloured lines represent the smooth trend of the AUCs overtime, using a Savitzky-Golay filter. Statistically significant deviation from the chance line, where an AUC of .50, is depicted by the solid straight lines at the top, when the classification detects the orientation at rate better than chance, or at the bottom of the plot, when the classification is worse than chance.

Discussion

The purpose of this experiment was to further extend the results from Robitaille (2016) and attempt to track the mental representations associated with the transformation of the orientation of a Gabor patch. While Robitaille (2016) applied a forward encoding model to evoked and induced alpha EEG activity as participants mentally rotated a line, here I applied the forward encoding model on raw EEG activity and probed activity silent state while participants either retained or transformed a mental representation of a Gabor patch. While we observe clear evidence that the orientation of the Gabor patch can be detected when it is presented to the screen (i.e., during perception), the detection does not reliably go further than perception into memory retention and imagery manipulation.

I evaluated behavioural performance not only to determine whether participants' performance is affected by the impulse stimulus and the direction of the mental rotation, but also to ensure that they were performing the task accurately so that the EEG activity recorded reflected the corresponding mental representation during the transformation. Here, averaging across all conditions, participants exhibited a mean raw error (i.e., standard deviation) of 12.5° for reporting the Gabor patch after a rotation of 60°. Although it appears that participants are doing better on this task than what is reported in previous chapters (i.e., raw error of ~18° with a mental rotation of 60° in Chapter 3, in contrast with ~9° without mental rotation in Chapter 2), it is important to note that I rejected trials that are too far away the target orientation. Therefore, this measure of performance of the participants is biased to retain only accurate trials. Looking at differences between conditions, I found no substantial differences in performance between the later and early impulse confirming that the onset of the task-irrelevant stimulus did not disrupt the processing, retention, and manipulation of the Gabor's orientation. On the other hand, I found a small (i.e., ~0.5°), yet significant difference in performance in favour of the clockwise rotation. There have been reports about the asymmetry in mental rotation with a clockwise rotation advantage (Koriat & Norman, 1985a; Robertson & Palmer, 1983), but it has not been reliably replicated (Koriat & Norman, 1985b). Liesefeld & Zimmer (2011) demonstrated that this asymmetry is not universal and depends on participants' rotation strategy. Thus, while the behavioural results indicate a statistically significant difference between rotation direction, it is not of practical significance given the magnitude of the difference. As such, it remains justified to average across both rotational directions to retain a higher statistical power for the analysis of the encoding model.

As for the EEG encoding, the channel tuning functions (CTFs) replicated some of the effects depicted in Robitaille (2016). Indeed, the encoding model for both the early and late impulse trials correctly identified the orientation of the Gabor patch during the presentation of the stimulus, but the channel tuning peak was not sustained for as long as observed in Robitaille (2016), either in the induced alpha or the exploratory raw EEG analyses (see Figure 4-1A and C respectively). Furthermore, our results do not successfully replicate that of Wolff et al. (2015, 2017) as they were able to use task irrelevant visual stimulus ("ping") to bring back the detection of the orientation of the stimulus held in memory. While our channel tuning functions (CTFs) show that the impulse stimulus increased the underlying channel response, the early impulse trials brought back the Gabor's original orientation, whereas it is not clear from the ROC classification what orientations were associated with the late response. Perhaps our ROC classification method is not suited to reliably detect the underlying patterns. While it may be the case that the activity present in raw EEG is too variable and contains too much noise irrelevant to the orientation of the stimulus to be clearly detected, it seems more likely that it is the result of a methodological decision; particularly since the decoding of feature from EEG activity is shown to be possible (Foster et al., 2016; Garcia et al., 2013;

Hajonides et al., 2021; Samaha et al., 2016; Sprague et al., 2016; Wolff et al., 2015, 2017).

Although the encoding model has the potential to add to the imagery format debate by tracking the mental manipulation of a depictive representation, our results offer limited contribution in favour of the depictive format. It seems that the initial detection of the orientation feature could be viewed as a depictive representation of the information. During the retention and manipulation of this representation, however, there was unreliable and barely detectable data about the orientation of the rotated stimulus. This is particularly interesting that the impulse stimulus could not bring back the relevant information, which suggests that the representation is perhaps not held in an activity-silent state either.

Although the purpose of the task irrelevant stimulus was to probe the activity-silent state of the mental representation present in the visual cortex, it may be the case that the imagery representations are not held in such a state. Based on recent reports using high-field (7T) fMRI studying feedforward and feedback processes in the visual cortex, Koenig-Robert & Pearson (2021) propose that visual imagery is driven by feedback connections that are found deeper within the visual cortex than that of perception-driven forward connections. Although it has been shown that there are close to double the number of feedback signals to forward signals in the primate brain (Markov et al., 2014; Perkel et al., 1986; Salin & Bullier, 1995), the functional purpose of these feedback connections are still unclear (Koenig-Robert & Pearson, 2021). While the feedforward connections are located in the granular layer of the primary visual cortex, a large portion of feedback signals are found deeper in the cortex, in the infra-granular layer (Callaway, 2004; Felleman & Van Essen, 1991). The activity in these deep layer of the visual cortex using high-field fMRI is found to be predictive of internally generated imagery (Bergmann et al., 2019) or illusory contours (Kok et al., 2016).

If the activity associated with imagery is found deeper within the cortex, this might explain why I could not reliably detect the imagery representation in the raw EEG activity. Moreover, in Robitaille (2016), the encoding of induced alpha activity showed significant evidence against the original orientation of the stimulus when participants rotated the object. If the signals for the imagery manipulation is found deeper within the visual cortex, and that the perceptually driven activity is found above, it is likely that a lot of conflicting activity is being detected at scalp. This functional organization of V1 can account for the inconsistent fluctuation between the detection of the original and transformed orientations during the imagery phase of the trial (see Figure 4-6). Therefore, it might explain why there are little evidence of tracking imagery manipulation from EEG, while features of working memory maintenance is more commonly reported (Foster et al., 2016; Garcia et al., 2013; Hajonides et al., 2021; Samaha et al., 2016; Sprague et al., 2016; Wolff et al., 2015, 2017).

The purpose of this chapter was to extend the findings of Robitaille (2016) and track the orientation of mental representations as they are being rotated. Despite focusing on raw EEG, which has shown to contain orientation-specific activity (Wolff et al., 2015, 2017), and using a task-irrelevant stimulus to “ping” activity-silent states, I could not reliably detect orientation of the mental rotation past perception. While individual difference can play a part in the lack of reliable detection, from different mental representation onset or speed, individual plots does not provide much insight into possible causes (see supplementary figures in Appendix A). However, given that new reports (Bergmann et al., 2019; Koenig-Robert & Pearson, 2021) suggest that the relevant neural activity associated with imagery representation are found deeper in the cortex, this presents a challenge to its detection using EEG. To track imagery manipulation, further studies are likely going to require to tease apart the activity associated with the perceptual feedforward and imagery backward signals.

References

Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & de Lange, F. P. (2013). Shared Representations for Working Memory and Mental Imagery in Early Visual Cortex. Current Biology, 23(15), 1427–1431. https://doi.org/10.1016/j.cub.2013.05.065

Baddeley, A. (2003). Working memory: Looking back and looking forward. Nature Reviews Neuroscience, 4(10), 829–839. https://doi.org/10.1038/nrn1201

Baddeley, A. D., & Hitch, G. (1974). Working Memory. In G. H. Bower (Ed.), Psychology of Learning and Motivation (Vol. 8, pp. 47–89). Academic Press. https://doi.org/10.1016/S0079-7421(08)60452-1

Bergmann, J., Morgan, A. T., & Muckli, L. (2019). Two distinct feedback codes in V1 for 'real' and 'imaginary' internal experiences. BioRxiv, 664870. https://doi.org/10.1101/664870

Brouwer, G. J., & Heeger, D. J. (2009). Decoding and Reconstructing Color from Responses in Human Visual Cortex. The Journal of Neuroscience, 29(44), 13992–14003. https://doi.org/10.1523/JNEUROSCI.3577-09.2009

Callaway, E. M. (2004). Feedforward, feedback and inhibitory connections in primate visual cortex. Neural Networks: The Official Journal of the International Neural Network Society, 17(5–6), 625–632. https://doi.org/10.1016/j.neunet.2004.04.004

Carandini, M., Demb, J. B., Mante, V., Tolhurst, D. J., Dan, Y., Olshausen, B. A., Gallant, J. L., & Rust, N. C. (2005). Do we know what the early visual system does? The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 25(46), 10577–10597. https://doi.org/10.1523/JNEUROSCI.3726-05.2005

Chaudhuri, R., & Fiete, I. (2016). Computational principles of memory. Nature Neuroscience, 19(3), 394–403. https://doi.org/10.1038/nn.4237

Cichy, R. M., Heinzle, J., & Haynes, J.-D. (2012). Imagery and perception share cortical representations of content and location. Cerebral Cortex (New York, N.Y.: 1991), 22(2), 372–380. https://doi.org/10.1093/cercor/bhr106

Cichy, R. M., Ramirez, F. M., & Pantazis, D. (2015). Can visual information encoded in cortical columns be decoded from magnetoencephalography data in humans? NeuroImage, 121, 193–204. https://doi.org/10.1016/j.neuroimage.2015.07.011

Costa, V. D., Tran, V. L., Turchi, J., & Averbeck, B. B. (2015). Reversal learning and dopamine: A bayesian perspective. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 35(6), 2407–2416. https://doi.org/10.1523/JNEUROSCI.1989-14.2015

Emrich, S. M., Riggall, A. C., Larocque, J. J., & Postle, B. R. (2013). Distributed patterns of activity in sensory cortex reflect the precision of multiple items maintained in visual short-term memory. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 33(15), 6516–6523. https://doi.org/10.1523/JNEUROSCI.5732-12.2013

Eriksson, J., Vogel, E. K., Lansner, A., Bergström, F., & Nyberg, L. (2015). Neurocognitive Architecture of Working Memory. Neuron, 88(1), 33–46. https://doi.org/10.1016/j.neuron.2015.09.020

Fawcett, T. (2004). ROC graphs: Notes and practical considerations for researchers. Machine Learning, 31(1), 1–38.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. Cerebral Cortex (New York, N.Y.: 1991), 1(1), 1–47. https://doi.org/10.1093/cercor/1.1.1

Fleming, P., Ball, L. J., Ormerod, T. C., & Collins, A. F. (2006). Analogue versus propositional representation in congenitally blind individuals. Psychonomic Bulletin & Review, 13(6), 1049–1055. https://doi.org/10.3758/bf03213924

Foster, J. J., Sutterer, D. W., Serences, J. T., Vogel, E. K., & Awh, E. (2016). The topography of alpha-band activity tracks the content of spatial working memory. Journal of Neurophysiology, 115(1), 168–177. https://doi.org/10.1152/jn.00860.2015

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

Garcia, J. O., Srinivasan, R., & Serences, J. T. (2013). Near-Real-Time Feature-Selective Modulations in Human Cortex. Current Biology, 23(6), 515–522. https://doi.org/10.1016/j.cub.2013.02.013

Hajonides, J. E., Nobre, A. C., van Ede, F., & Stokes, M. G. (2021). Decoding visual colour from scalp electroencephalography measurements. NeuroImage, 237, 118030. https://doi.org/10.1016/j.neuroimage.2021.118030

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. Nature, 458(7238), 632–635. https://doi.org/10.1038/nature07832

Haynes, J.-D., & Rees, G. (2005). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. Nature Neuroscience, 8(5), 686–691. https://doi.org/10.1038/nn1445

Ishai, A., & Sagi, D. (1995). Common mechanisms of visual imagery and perception. Science (New York, N.Y.), 268(5218), 1772–1774.

Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. Nature Neuroscience, 8(5), 679–685. https://doi.org/10.1038/nn1444

Koenig-Robert, R., & Pearson, J. (2021). Why do imagery and perception look and feel so different? Philosophical Transactions of the Royal Society B: Biological Sciences, 376(1817), 20190703. https://doi.org/10.1098/rstb.2019.0703

Kok, P., Bains, L. J., van Mourik, T., Norris, D. G., & de Lange, F. P. (2016). Selective Activation of the Deep Layers of the Human Primary Visual Cortex by Top-Down Feedback. Current Biology: CB, 26(3), 371–376. https://doi.org/10.1016/j.cub.2015.12.038

Koriat, A., & Norman, J. (1985a). Reading rotated words. Journal of Experimental Psychology: Human Perception and Performance, 11(4), 490–508. https://doi.org/10.1037/0096-1523.11.4.490

Koriat, A., & Norman, J. (1985b). Mental rotation and visual familiarity. Perception & Psychophysics, 37(5), 429–439. https://doi.org/10.3758/BF03202874

Kosslyn, S. M. (1973). Scanning visual images:Some structural implications. Perception & Psychophysics, 14(1), 90–94. https://doi.org/10.3758/BF03198621

Kosslyn, S. M. (1996). Image And Brain: The Resolution of the Imagery Debate (1 edition). A Bradford Book.

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural Foundations of Imagery. Nature Reviews Neuroscience, 2(9), 635–642. https://doi.org/10.1038/35090055

Liesefeld, H. R., & Zimmer, H. D. (2011). The advantage of mentally rotating clockwise. Brain and Cognition, 75(2), 101–110. https://doi.org/10.1016/j.bandc.2010.10.012

Ma, W. J. (2018). Problematic usage of the Zhang and Luck mixture model. BioRxiv, 268961. https://doi.org/10.1101/268961

Markov, N. T., Vezoli, J., Chameau, P., Falchier, A., Quilodran, R., Huissoud, C., Lamy, C., Misery, P., Giroud, P., Ullman, S., Barone, P., Dehay, C., Knoblauch, K., & Kennedy, H. (2014). Anatomy of hierarchy: Feedforward and feedback pathways in macaque visual cortex. The Journal of Comparative Neurology, 522(1), 225–259. https://doi.org/10.1002/cne.23458

Marmor, G. S., & Zaback, L. A. (1976). Mental rotation by the blind: Does mental rotation depend on visual imagery? Journal of Experimental Psychology: Human Perception and Performance, 2(4), 515–521. https://doi.org/10.1037/0096-1523.2.4.515

Pearson, J., Naselaris, T., Holmes, E. A., & Kosslyn, S. M. (2015). Mental Imagery: Functional Mechanisms and Clinical Applications. Trends in Cognitive Sciences, 19(10), 590–602. https://doi.org/10.1016/j.tics.2015.08.003

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. Journal of Neuroscience Methods, 162(1), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Perkel, D. J., Bullier, J., & Kennedy, H. (1986). Topography of the afferent connectivity of area 17 in the macaque monkey: A double-labelling study. The Journal of Comparative Neurology, 253(3), 374–402. https://doi.org/10.1002/cne.902530307

Postle, B. R. (2006). Working Memory as an Emergent Property of the Mind and Brain. Neuroscience, 139(1), 23–38. https://doi.org/10.1016/j.neuroscience.2005.06.005

Robertson, L. C., & Palmer, S. E. (1983). Holistic processes in the perception and transformation of disoriented figures. Journal of Experimental Psychology. Human Perception and Performance, 9(2), 203–214. https://doi.org/10.1037//0096-1523.9.2.203

Robitaille, J. (2016). Encoding Induced Alpha EEG Activity Reveals Visual Working Memory Representations and their Manipulations. http://hdl.handle.net/10464/10688

Salin, P. A., & Bullier, J. (1995). Corticocortical connections in the visual system: Structure and function. Physiological Reviews, 75(1), 107–154. https://doi.org/10.1152/physrev.1995.75.1.107

Samaha, J., Sprague, T. C., & Postle, B. R. (2016). Decoding and Reconstructing the Focus of Spatial Attention from the Topography of Alpha-band Oscillations. Journal of Cognitive Neuroscience, 28(8), 1090–1097. https://doi.org/10.1162/jocn_a_00955

Savitzky, Abraham., & Golay, M. J. E. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Analytical Chemistry, 36(8), 1627–1639. https://doi.org/10.1021/ac60214a047

Serences, J. T. (2016). Neural mechanisms of information storage in visual short-term memory. Vision Research, 128, 53–67. https://doi.org/10.1016/j.visres.2016.09.010

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

Serences, J. T., & Saproo, S. (2012). Computational advances towards linking BOLD and behavior. Neuropsychologia, 50(4), 435–446. https://doi.org/10.1016/j.neuropsychologia.2011.07.013

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. Science (New York, N.Y.), 171(3972), 701–703.

Sprague, T. C., Ester, E. F., & Serences, J. T. (2016). Restoring Latent Visual Working Memory Representations in Human Cortex. Neuron, 91(3), 694–707. https://doi.org/10.1016/j.neuron.2016.07.006

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

Thompson, W. L., & Kosslyn, S. M. (2000). Neural systems activated during visual mental imagery. In A. W. Toga & J. C. Mazziotta (Eds.), Brain mapping: The systems (pp. 535–560). Academic Press.

Tong, F. (2013). Imagery and visual working memory: One and the same? Trends in Cognitive Sciences, 17(10), 489–490. https://doi.org/10.1016/j.tics.2013.08.005

van den Berg, R., & Ma, W. J. (2018). A resource-rational theory of set size effects in human visual working memory. ELife, 7. https://doi.org/10.7554/eLife.34963

Wolff, M. J., Ding, J., Myers, N. E., & Stokes, M. G. (2015). Revealing hidden states in visual working memory using electroencephalography. Frontiers in Systems Neuroscience, 123. https://doi.org/10.3389/fnsys.2015.00123

Wolff, M. J., Jochim, J., Akyürek, E. G., & Stokes, M. G. (2017). Dynamic hidden states underlying working-memory-guided behavior. Nature Neuroscience, advance online publication. https://doi.org/10.1038/nn.4546

Appendix A

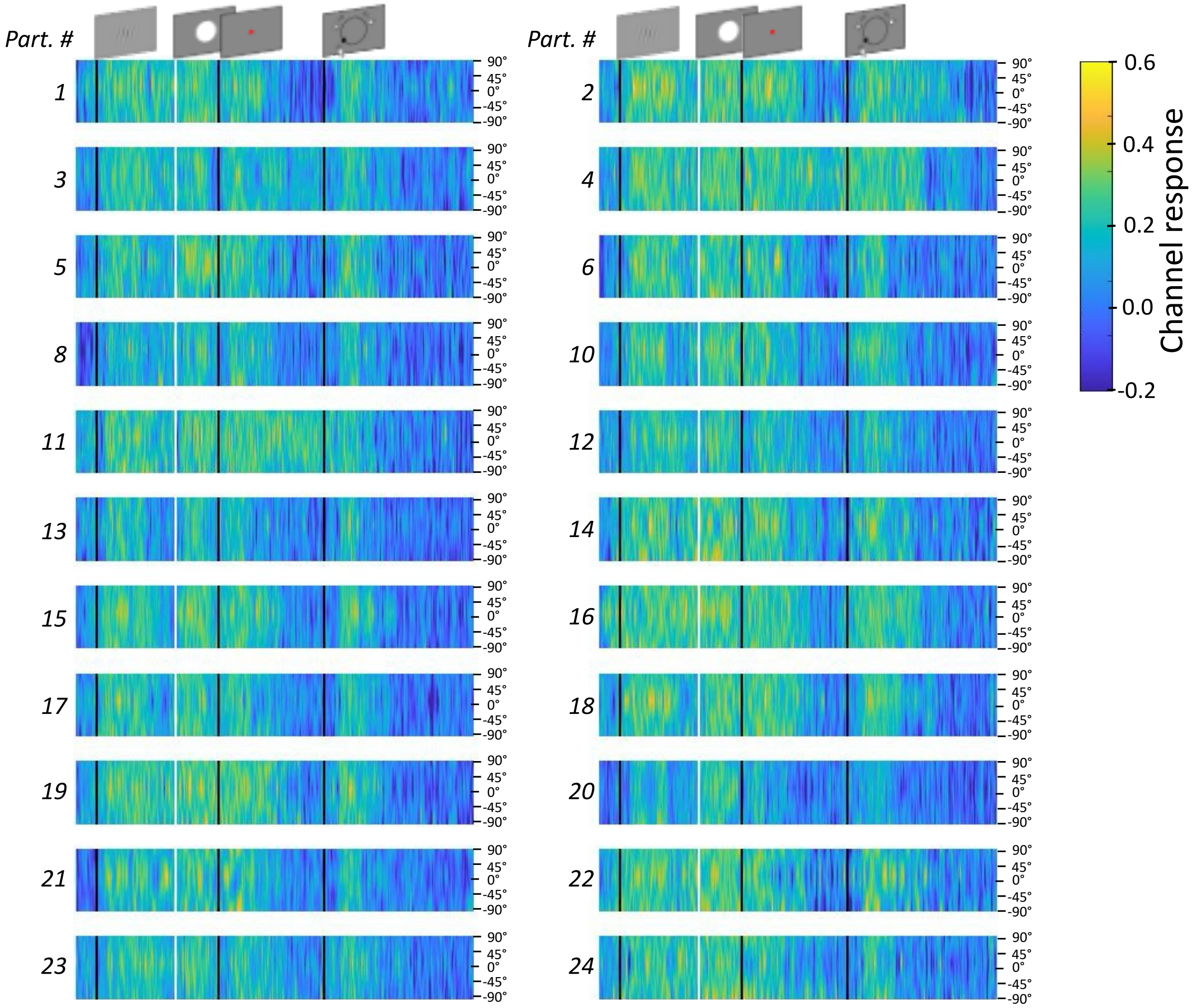

Figure 4-7. Supplementary figure of the orientation selectivity of raw EEG activity over time (x-axis) averaged for each participant for early impulse trials. The beginning of the trial, the presentation of the rotation cue, and the onset of the probe are marked by a vertical black line, and the impulse stimulus is marked (white line).

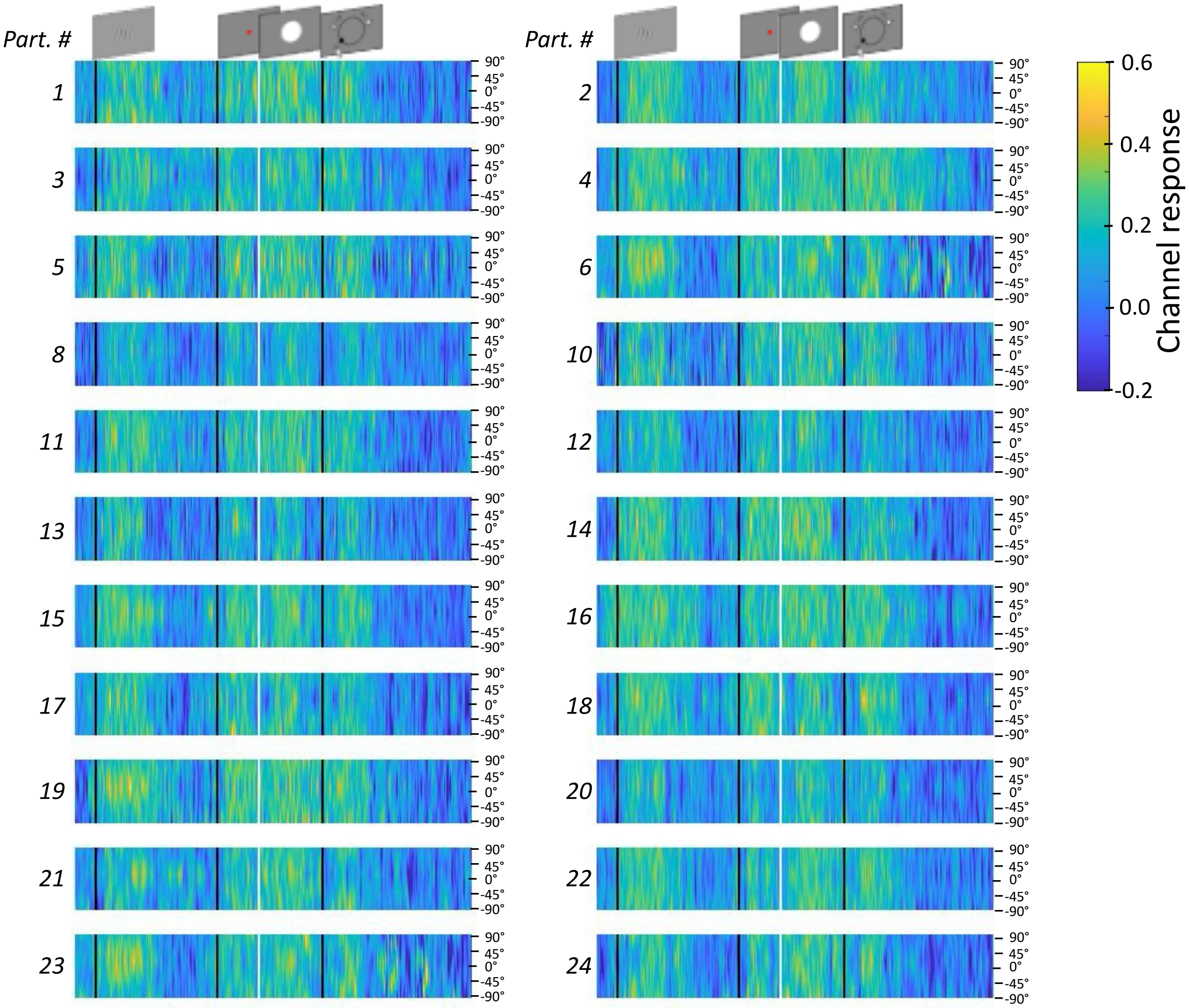

Figure 4-8. Supplementary figure of the orientation selectivity of raw EEG activity over time (x-axis) averaged for each participant for late impulse trials, where the impulse appears at the 1650ms mark. The beginning of the trial, the presentation of the rotation cue, and the onset of the probe are marked by a vertical black line, and the impulse stimulus is marked (white line).

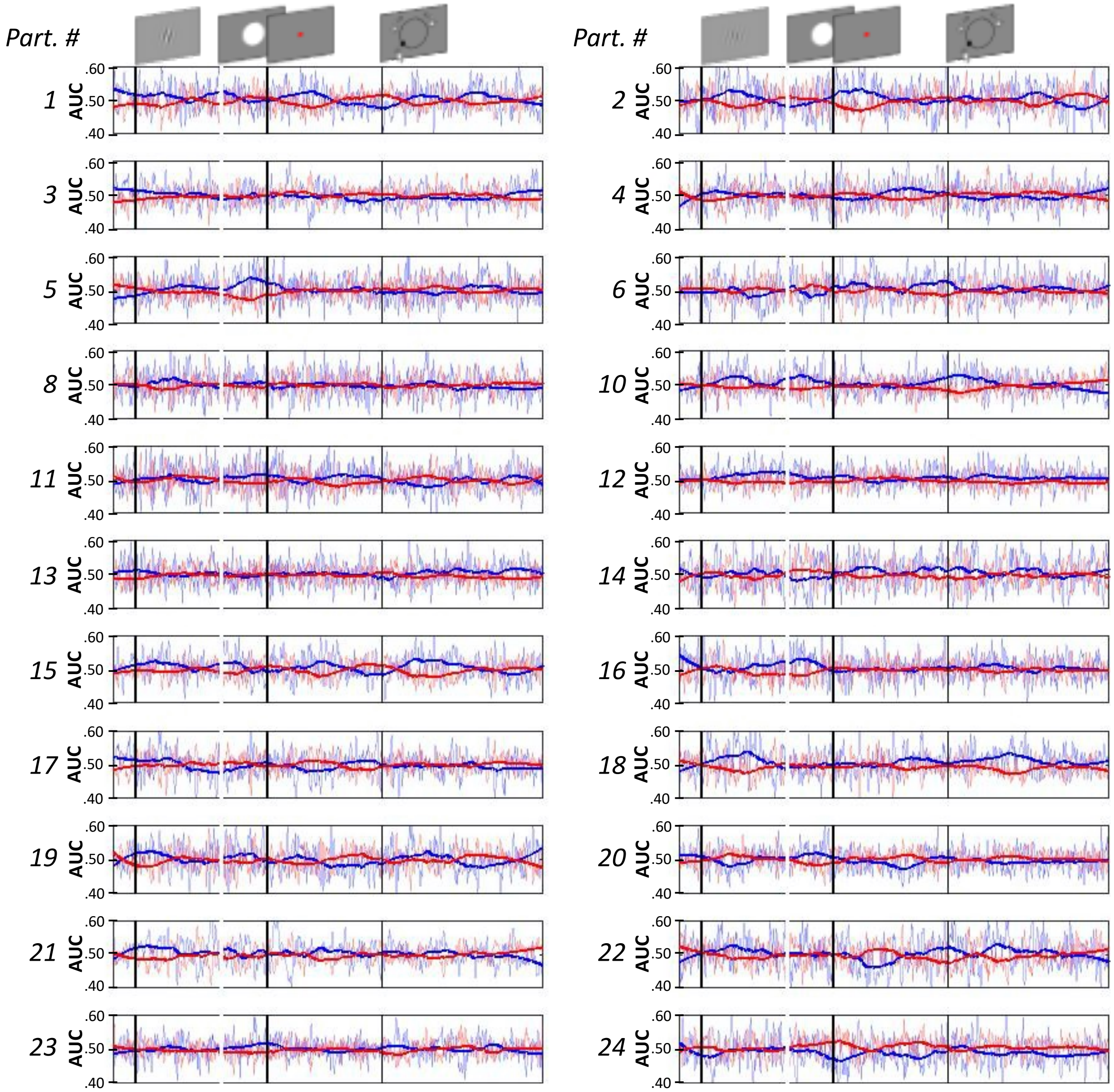

Figure 4-9. Supplementary figure of AUCs of the CTFs ROC classification over time tracking the orientation of the original stimulus (blue) and the rotated orientation (red) for each participant for the early trials. The beginning of the trial, the presentation of the rotation cue, and the onset of the probe are marked by a vertical black line, and the impulse stimulus is marked (white line).

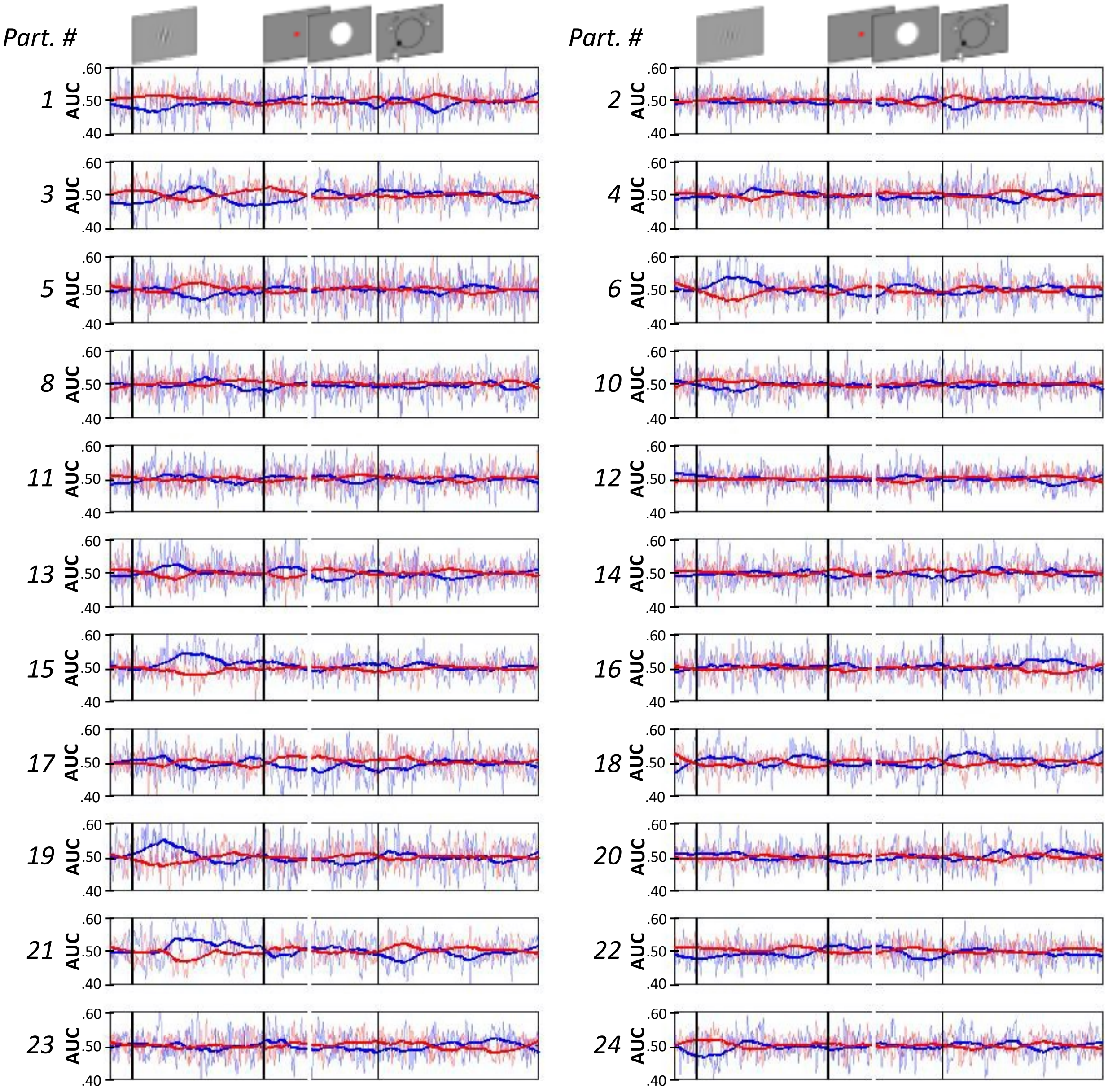

Figure 4-10. Supplementary figure of the AUCs of the CTFs ROC classification over time tracking the orientation of the original stimulus (blue) and the rotated orientation (red) for each participant for the late trials.The beginning of the trial, the presentation of the rotation cue, and the onset of the probe are marked by a vertical black line, and the impulse stimulus is marked (white line).

Chapter 5

General Discussion

The imagery and working memory fields exhibit a great deal of overlap in the conceptual depiction of the mental representation of visual information, despite working in relative independence (Tong, 2013). The investigation of the neural structures and functional activity supporting these representations has also revealed a strong overlap between these two areas of study. With the current state of the literature, however, it is rather difficult to demonstrate that memory and imagery representations are related, as research on these topics have not adopted methodologies that permit a direct comparison. Therefore, the overarching purpose of this thesis was to investigate the mental representation of visual information and to bridge between the working memory and imagery literatures.

Chapter 2 addressed the issue that imagery and working memory paradigms use markedly different stimuli, which makes it difficult to compare the properties of mental representations as described by both literatures. During three experiments, I evaluated the psychophysical properties (i.e., raw error, guess rate, and precision) of working memory for the orientation of a simple line stimulus compared to a complex 2.5D object typically used in visual imagery experiments. All experiments demonstrated a cost associated with complexity that increased participants' response errors, where the predominant impact is on the precision of the mental representation, with some increase in retention errors. Strong correlations between performance on the different stimuli were found, suggesting that performance between simple and complex stimulus are indicative of a common mechanism, while also distinguishing between specific stimulus properties (e.g., feature dimension and axis of rotation) that work affect performance independently.

Chapter 3 examined the effect of mental manipulation of these mental representations as well as evaluated whether different performance metrics derived from distinct paradigms assess the same construct. In the first experiment, reported errors declined for both the simple and complex objects when participants were required to mentally rotate their mental representation of the object, but it did not scale as a function of the magnitude of the rotation. Further investigation into these report errors using the mixture model showed that the precision of responses declines with the scope of the rotation, while retention for representation of the complex object only is affected when the object is mentally manipulated. The second experiment shows that the performance metrics obtained from different paradigms traditionally adopted by the imagery and working memory literatures capture the same decline in performance as a function of the degree of manipulation. Using Principal Component Analysis, a common mechanism is identified that is driving performance for simple and complex objects across memory retention and manipulation in Experiment 1, as well as between paradigms in Experiment 2. These results demonstrate that the representations prompted from a across stimuli and procedures produce similar behavioural results and are driven by strongly overlapping mechanisms.

Finally, Chapter 4 investigated whether it is possible to track imagery manipulation by using an encoding method applied to raw EEG activity recorded over the visual cortex. The task introduced a task-irrelevant stimulus (i.e., impulse) that is believed to probe “activity-silent” neural representations. The results demonstrated some evidence that the orientation of the stimulus can be detected during the perception of the stimulus, but that there wasn't reliable information beyond this point that would allow the tracking of the orientation held in memory or during the manipulation. While at first glance, the task-irrelevant stimulus appears to have brought back the representation of the orientation, the information probed by the impulse was not reliable enough to clearly identify the orientation.

Mental representations of object complexity through dimensionality

One of the reasons why there is little convergence between the imagery and working memory research (Tong, 2013), particularly from the behaviour investigation standpoint, is that the two literatures have historically adopted different methodological approaches. The working memory literature has focused on simplistic stimuli that carries a limited number of features to investigate the psychophysical properties of memory representations. On the other hand, the imagery literature tends to work with more everyday common objects. While there is evidence that increasing the number of features in a stimulus (i.e., increase complexity) influences the psychophysical properties of working memory representations, the nature of complexity tends to reveal mixed effect on performance. Indeed, when complex items are perceived as familiar (Xie & Zhang, 2017), relatable (Zhou et al., 2018) or real (Snow et al., 2014), participants tend to do better on the task whereas performance decreases when the items are more abstract (Delvenne & Bruyer, 2004; Eng et al., 2005; Luria et al., 2010; Song & Jiang, 2006; Wheeler & Treisman, 2002; Xu & Chun, 2006). To attempt to address the effect of item complexity and avoid the familiarity confound, Chapters 1 and 2 adopted a 2.5D block stimuli that are unfamiliar but recognizable 3D objects. Focusing on the orientation feature, the psychophysical properties of the mental representation of these objects were compared against a simple line stimulus to determine the effect of spatial dimensions, by manipulating dimensionality through monocular depth cues to an object, on working memory and mental rotation.

Although the literature has reported costs to performance when participants are required to remember more complex objects, the manifestation of this cost is not well agreed upon. The results from all three experiments in Chapter 2 and Experiment 1 in Chapter 3 show a clear decrease in performance (i.e., greater raw error) when participants are required to report the orientation of a complex object compared to that of a simple line. Instead of manipulating stimuli complexity through conjunctions (Luck & Vogel, 1997), colour patches (Christophel et al., 2015), symbols and characters (Alvarez & Cavanagh, 2004), etc., in this thesis, the increase in complexity of the object was achieved through dimensionality. While we observe a cost due to complexity, the manipulation of dimensionality shows an interesting relationship with performance: Experiment 2 of Chapter 2 shows that worse performance is found for complex 2D objects compared to 2.5D objects, even though the 2.5D object contains the greatest number of features. Snow et al. (2014) reported a similar effect when participants showed the best memory performance when presented with the physical objects rather than a digital picture of the same objects, and the worst performance when presented with drawings. These results suggest that despite the real objects and pictures depicting mostly the same features, - excluding binocular cues perhaps - the physical object provided a greater memory advantage. Thus, it appears that, to a certain extent, dimensionality can facilitate recall performance and that the effect of item complexity on memory performance can be mitigated by the features provided by the addition of this depth cue. As mentioned before, it is not the first time that the nature of the complexity has shown to facilitate memory encoding and retention. Like faces, familiar, relatable, and real objects, it seems that the presence of monocular depth cues is sufficient to facilitate performance.

The manipulation of complexity, as defined by the number of features present in a stimulus, is not as straightforward as if first appears. While there are a number of ways complexity has been manipulated in the literature or can be conceived, this thesis reports the manipulation of complexity using dimensionality to avoid previously reported advantages to performance such as familiarity (Xie & Zhang, 2017), relatability (Zhou et al., 2018) or realism (Snow et al., 2014). While we generally find that the complexity of an object negatively impacts memory or imagery performance, this relationship is not consistent across stimulus complexity given these advantages. As discussed previously, Chapter 2 Experiment 2 is a great example of this as we observed worse performance for the 2D shapes in comparison to the 2.5D shapes. Therefore, investigating "objective" complexity as being associated with the number of features is perhaps the wrong way to explore complexity and establishing "subjective" complexity is possibly a more efficient way to investigate the impact of stimuli on memory performance. For instance, Alvarez and Cavanagh (2004) operationally defined the complexity of a series of stimulus types based on participants' performance on a search task before evaluating the memory capacity for these stimuli. Thus, complexity defined and evaluated subjectively might offer a more straightforward way to explore how the visual information is treated for the maintenance and manipulation of mental representations.

Mental representations of static stimuli

Despite the cost of complexity and the facilitation of depth cues on memory performance, the question remains: are simple and complex object cognitively represented the same way? The working memory literature has been inferring the type of mental representation from the psychometric evaluation of memory, particularly by assessing memory capacity. According to one of the first models proposed, discrete memory representations are believed to be stored with great precision but limited by a number of placeholders called “slots” (Luck & Vogel, 1997; Zhang & Luck, 2008). The results of the Bayesian Model Selection (BMS) from Chapter 2 – Experiment 1 shows that most participants do exhibit characteristics of the discrete model of working memory. Indeed, most participants (~64.2% for simple lines, ~66.7% for 2.5D objects) show a fixed level of precision across set sizes, which is a characteristic of the discrete view of mental representations. Therefore, it appears that the majority of participants represent memory items that is best characterized by the discrete model of memory representations.

It would not be accurate, however, to say that there is only evidence in favour of the discrete model present in this thesis. Although there is some support in favour of the discrete view of working memory representations (Todd & Marois, 2004; Vogel & Machizawa, 2004; Xu & Chun, 2006; Zhang & Luck, 2008), most of the recent evidence suggest that the resource-based models, which are characterized by a monotonic decrease in response precision as cognitive load increases, are better supported by the data (Bays & Husain, 2008; Wilken & Ma, 2004). Resource-based models have also outperformed discrete models when compared against each other (Ma et al., 2014; van den Berg et al., 2014). Moreover, some of the early evidence in favour of the discrete models (e.g., Todd

& Marois, 2004; Xu & Chun, 2006) have been reinterpreted under the resource-based models (Bays, 2018). The results from the standard mixture model in Chapter 2 Experiment 1 shows that there is a significant decrease in precision across set size, which is more typical of the resource view of working memory. Likewise, the Bayesian Model Selection (BMS) also shows a significant number of participants (~30.1% for simple lines, ~27.8% for 2.5D objects) depicting better fits for models that has decreasing levels of precision (i.e., ordered precision) as memory load increases. These indicate that not all participants depict the characteristics associated with discrete memory representations, but rather reflects characteristics of resource-based memory representations.

More importantly, while it is still not clear which working memory model best represents the data, most participants tend to exhibit similar psychometric pattern for both types of stimuli. Indeed, the results from the BMS in Chapter 2 Experiment 1 reveals that 30 out of 44 participants show better fits for the same model applied to simple and complex objects. Despite the fact that all 30 participants do not agree on the same model – 26 prefers the ordered guesses and fixed precision (*o*G-*f*SD), and 4 prefers the ordered guesses and ordered precision (*o*G-*o*SD) – this suggest that most participants are treating or representing simple and complex objects similarly. This gives some strong evidence that the mental representation for stimuli for working memory tends to be consistent. Nevertheless, there are still 14 participants that are not treating these stimuli in the same way, suggesting that individual differences play a significant role in the way we create mental representation of visual information.

While the validity of the mixture model is being questioned, particularly about the existence of guessed responses (Schurgin et al., 2020), by comparing the psychophysical properties of the mental representation of simple and complex stimulus, this thesis shows that there is consistency in how these stimuli are being represented. Although there is some evidence that individual differences have an impact on the psychophysical properties of mental representation, there is still a generally preferred model that indicates it is consistent across participants and stimuli to be remembered.

The Manipulation of Mental Representations

Although the mental representation for objects with different complexity tends to be similar, the static representation of this visual information is only one aspect of working memory and imagery. Indeed, it is built in both imagery and working memory concepts that these representations can also be manipulated. There are few reports investigating this aspect of working memory despite the fact that Baddeley and Hitch (1974) proposed their working memory model as a way to account the flexibility of memory representations that was not addressed in the short-term memory model. The working memory field, however, has focused more on the psychophysical properties of static representations, while the manipulation of this information has been more or less ignored. On the other hand, the imagery literature has a long history of investigating the transformation of mental representations. It is only in recent years and with the neural investigation of working memory and imagery representations that the literature began considering that the mental representation described in these fields may be the same. While there is considerable evidence suggesting a conceptual and neural overlap between visual memory and imagery, it is still unclear whether these representations are the same, which should produce similar behavioural properties. Moreover, because both literatures have tackled the investigation of the mental representation of visual information using different approaches, they have adopted different methods to assess cognitive performance.

While there isn’t a consensus on what differentiates memory representations from imagery, this thesis offers some psychophysical distinctions between the two types of representations. As depicted in the results from the imagery manipulation in Chapter 3 Experiment 1, there is a performance cost to the manipulation of the mental representation. This decrease in performance, however, is comparable to the cost associated with an increase in set size (Chapter 2 – Experiment 1), or rotations along the x- and y-axes (Chapter 2 – Experiment 3). It is also interesting that simple and complex stimuli are affected in the same fashion once the object is transformed and that the manipulation cost does not scale as a function of rotation magnitude. Although this psychophysical discrepancy may suggest that the representations are qualitatively different, it does not resolve whether the format of imagery and working memory are fundamentally different. Rather, it suggests that representations that are manipulated perhaps decrease in quality or that the representation is more susceptible to internal interference.

Despite the presence of a cost to performance when mental representations are manipulated, there is some compelling evidence that the mental manipulation and static retention of visual representation is driven by similar processes as depicted by the Principal Component Analysis of Chapter 3. Looking at the two components’ solutions specifically, both experiments maintain the overlap in loadings for memory retention and mental rotation performance. Therefore, these results suggest that the representations are driven by overlapping processes, where the static and manipulated representations are strongly associated, but are still exhibiting distinct psychophysical properties.

It is important to note, however, some limitations to these correlation analyses (i.e., repeated-measure correlation, partial-correlations, PCAs). Although they do provide evidence of shared variance the source(s) of this variance is unclear. For instance, by virtue of presenting tasks on a computer, participants ability to operate the mouse accurately, etc. are all aspects of the task that can produce shared variance. Then, to make a more convincing case, one should include more tasks that assess imagery and working memory differently, and perhaps also include some perceptual and motor imagery tasks to either partial-out these processes or show their involvement in the retention and manipulation of visual information. Despite this, limitation, the two-component models of Chapter 3 still offers some convincing evidence that there is a greater separation of the variance between stimulus type and method of assessment than between working memory and imagery processes.

The Format of Mental Representations

For decades, the imagery literature has been debating whether the format of mental images take the depictive form (Kosslyn, 2005) or that these mental representations are propositional or symbolic (Pylyshyn, 2003). Most of the recent findings, mainly from brain imaging (Kosslyn et al., 2001), is in support of the depictive format, particularly because feature-specific activity associated with the imagery representation can be identified within primary visual cortex (Albers et al., 2013; Serences et al., 2009). The Vividness of Visual Imagery Questionnaire (VVIQ) (Marks, 1973) was devised to capture one’s ability to form these depictive representations.

Indeed, some studies have shown an association between the subjective experience of mental imagery and the activation of the visual cortex during an imagery task (Amedi et al., 2005; Cui et al., 2007), which has been used to demonstrate that participants are engaging in imagery, as well as to diagnose aphatasia – a condition where an individual is unable to experience voluntarily bring about mental images (Zeman et al., 2015). Thus, the VVIQ has been used as reliable tool to assess participants engagement in depictive imagery.

Although there are some validity to this subjective measure with imagery tasks (Campos, 2011; McKelvie, 1995), Chapter 3 shows some mixed results. The correlations between the composite scores, created from the Principal Component Analysis loadings in Chapter 3, and the VVIQ scores show that there was no association between the component extracted in the PCA and general performance (i.e., PC1) or for simple and complex composite scores (i.e., RC1 and RC2) in Experiment 1, whereas strong correlations were found in Experiment 2. This is puzzling given that the complex objects in Experiment 1 and the stimuli used in Experiment 2 are very similar, which should produce comparable mental representations. Therefore, based on the VVIQ alone it is not clear upon which imagery format participants are relying.

One explanation for this discrepancy may be that the paradigm employed might have an influence on the format of imagery. Although there is a significant correlation between the VVIQ scores and the composite score created that captures performance on the continuous task (i.e., RC1) in Experiment 2 of Chapter 2, there is only anecdotal evidence from the Bayesian correlation that supports this relationship. This coupled with the absence of correlation in Experiment 1 for any composite scores and vividness of imagery suggest that the continuous paradigm might not rely on vivid imagery to complete the task. The continuous recall paradigm used in experiment 2 does allow for the stimuli to be manipulated externally whereas participants that are presented with the dichotomous probe must rely on their internal abilities to manipulate the stimulus. Therefore, it may be the case that participants are relying on the continuous probe as a means to mentally rotate the object without using their own imagery representations. Although these results are not decisively supporting either format, it appears that when participants are required to rely on their own internal abilities to perform the mental rotation, the format is consistent with the depictive format.

The results of Chapter 4 had the greatest potential to provide positive evidence in favour of the depictive format. Indeed, the ability to track the mental representation of the Gabor patch as it is mentally rotated would have extended previous reports identifying features of the mental representation held in memory from EEG activity (Foster et al., 2015; Garcia et al., 2013; Hajonides et al., 2021; Samaha et al., 2016; Sprague et al., 2016) and it would corroborate participants' self-reports. While at first glance the encoding model appears to reveal the orientation of the stimulus during perception and retention phase, the statistical results demonstrate that the activity is not reliable enough to track the orientation of the stimulus past perception. Although it is difficult to interpret negative findings, this result does favour the propositional format by default because participants were able to successfully complete the task while I have no evidence that they relied on depictive representations. It is important to note, however, it does not provide positive evidence in favour of the propositional format either.

Moreover, despite not having any direct evidence in favour of the propositional format. Chapter 3 does show some evidence against the typical conceptualization of the depictive format. Indeed, the rate of mental rotation reported in Experiment 2 of Chapter 3 goes against the idea that mental rotation as done uniformly as it has been demonstrated previously (Shepard & Cooper, 1986; Shepard & Metzler, 1971). In one of his critique of the depictive format, Pylyshyn (1979) argued that it the rate of mental rotation is not always done at a fixed rate. Liesefeld & Zimmer (2011), however, showed that some participants can make quick mental flip of 180° that significantly reduces the time to mentally rotate an object. Although this strategy does not necessarily negate the possibility that the mental representation is in a depictive format, it does conflicts with some of the idea that these representations are manipulated the same way as one would with a physical object. Therefore, it seems that the results from this thesis is mostly in agreement the depictive view of mental representations.

Shared and Distinct Mechanisms for Mental Representations

Although, the initial discovery of frontal cortex activation using univariate analyses of fMRI BOLD signals during working memory tasks have suggested a possible location of memory representations (Curtis & D’Esposito, 2003; Todd & Marois, 2004; D’Esposito & Postle, 2015; Chaudhuri & Fiete, 2016; Costa et al., 2015; Eriksson et al., 2015), investigations using multivariate analyses showed that the feature sensitive information associated with the nature of the mental representation can be found in the visual cortex (Albers et al., 2013; Emrich et al., 2013; Harrison & Tong, 2009; Haynes & Rees, 2005; Kamitani & Tong, 2005). Similarly, the imagery field reported similar fMRI BOLD activation between imagery representations and visual perception (Cichy et al., 2012; Ishai & Sagi, 1995; Kosslyn et al., 2001; Pearson et al., 2015) leading to the claim that imagery may be seen as a form of “weak” perception (Pearson et al., 2015). Although it seems clear that from the neuroimaging data that the cognitive representations generated from imagery and working memory appears equivalent, the network responsible for these visual representations found in the visual cortex may not hold these representations as part of a unitary mechanism. Indeed, evidence from this thesis supports the idea that perhaps subparts of the visual cortex might offer varying contributions depending on the nature of the representation.

As mentioned previously, this thesis provides evidence demonstrating that there are clear shared mechanisms that drive performance across imagery and working memory representations. There is also evidence of independent mechanisms at play that appears to be tied to the additional features (e.g., colour, 3D cues, etc.) of the mental representation that are irrelevant to the critical orientation feature. The PCAs of Chapter 3 shows that a single component explains most of the variance found in performance on working memory retention and imagery manipulation for simple and complex stimuli (i.e., Experiment 1) and regardless of the paradigm used (i.e., Experiment 2). However, two-components solutions suggest that the variance can distinguish between stimuli or assessment method (i.e., dichotomous or continuous reports) more easily than between the mechanisms that are involved during memory retention or imagery manipulations. Similarly, across all three experiment, Chapter 2 also shows strong correlations between the representation of simple and complex objects while suggesting the involvement of distinct mechanisms in the creation of visual representations. Indeed, the partial correlations of Experiment 2 suggest controlling for memory capacity show that the type of the feature creates some differences in the mental representation created.

As proposed in the literature (Delvenne & Bruyer, 2004; Wheeler & Treisman, 2002), it appears that individuals possess distinct capacity for different features (e.g., orientation, colour, shape, etc.), where these separate capacities likely engage distinctive brain areas. The correlation between axes of rotation in Experiment 3 of Chapter 2 also shows that complex objects rotated around the y- and x-axes shows greater association to one another than they are to similar objects rotated around the z-axis, while showing even less association with a simple stimulus. These results demonstrate that the closer the features of a mental representations are to one another the closer the association, despite the fact that the critical feature (i.e., orientation) is constant across all experiments. Therefore, these results suggest that, instead of holding visual information in a single location such as the early visual cortex, mental representation of visual information relies on a network that contains elective parts.

Despite that across all experiments of this thesis orientation is the critical feature, the association between the representation of different stimuli is not straightforward. The network necessary for the creation, maintenance, and manipulation of mental representations seems to greatly dependent on the nature of the representation itself and the presence of irrelevant features in complex stimuli appear to have varied impact on performance. Although the addition of features to stimuli provides a challenge to identify its orientation in contrast to the simple line stimulus, a complex object with monocular depth cues helps performance in comparison to a complex 2D object (see Chapter 2 – Experiment 2). Extra features in the form of depth cues may engage elective parts of the visual processing hierarchy that ultimately help bind all the features into a more coherent and tangible representation. It was speculated that the advantage that participants exhibit when memorizing complex objects, particularly with faces (Curby & Gauthier, 2007; Jackson & Raymond, 2008), is a result of a holistic process. The holistic process involved in faces perception is thought to be a fast-track mechanism that integrate several sensory features together (Taubert et al., 2011). While holistic processing has been studied mainly for face recognition (Farah et al., 1998; Tanaka & Sengco, 1997) and object of expertise (Gauthier & Tarr, 2002), it has been suggested that any realistic object may benefit from a holistic-like processing (Curby & Gauthier, 2007). As mentioned earlier, participants show distinctive memory capacity-limits for different feature dimensions (Delvenne & Bruyer, 2004; Wheeler & Treisman, 2002), which suggests that a more complex and naturalistic object may rely on several different integrated featural processes. The partial independence of the memory capacity for different features is supported by our correlations of Experiment 2 in Chapter 2, where I show greater contribution to performance from our measure of orientation capacity than colours.

Relying on a holistic process would suggest that the representation has greater consolidation by binding several features into a more complete memory representation. Although orientation is the core feature necessary for the performance of the tasks described across the entire thesis and that this feature is mainly represented in the primary visual cortex (i.e., V1), which is supported by the results of chapter 4, the 2.5D object is likely going to require the contribution of several parts higher in the visual processing hierarchy. Given that stimuli used in Chapter 2 and 3 are designed to be unfamiliar realistic objects and that the spatial relations between features are critical to these stimuli, it is likely that the components of the “where” pathway is more engaged than the “what” pathway. Reports from the imagery literature describe neuroimaging and brain stimulation evidence that the parietal cortex is engaged during mental rotation task (Harris & Miniussi, 2003; Jordan et al., 2001; Parsons, 2003; Zacks, 2008), which are part of the “where” pathway. Conversely, when the focus is on object identification in an imagery task, some portion of the “what” pathway is engaged, but the main activation is found in the primary visual cortex (Cui et al., 2007; Kosslyn et al., 1997; Kosslyn & Thompson, 2003). While performance on the recollection of their orientation is reduced, perhaps the engagement of several components of the visual processing hierarchy may offer the advantage of a more robust (i.e., less sensitive to interference) representation.

Although the cognitive representations described by the imagery and working memory fields are engaging similar mechanisms within the visual cortex, it appears that there are more differences between representations that contain contrasting features than between memory retention and imagery manipulation processes. The results from this thesis point to a non-unitary network containing elective feature-sensitive components that creates, maintains, and manipulates visual representations.

Individual Differences and Functional Purpose of Mental Representations

Even though this thesis was not set up to investigate individual differences in their cognitive representation of visual information, but rather determine how these representations are affected within the same individual, there is evidence that individual differences play a role. As discussed earlier about the BMS analysis of Chapter 2 Experiment 1 (see Mental Representations of Static Stimuli), some participants seem to have adopted different strategies when presented with simple or complex stimuli. Also, the composites scores correlation with questionnaires in Chapter 3 suggest that participants do not always rely on visual imagery strategies to perform mental rotations. Although it is not clear whether these differences were conscious, there is evidence in this thesis that, not only can participants differ from one another but also, they can differ in strategy when presented with different stimuli. The unreliable results of Chapter 4 can also be partially explained by individual difference as participants do not necessarily show the same speed or onset of mental rotation as well as rely on the same strategies to complete the rotation task. In the past couple of decades, the literature has increasingly investigated individual differences in imagery and working memory abilities, which might imply significant impact on the conception of mental representation of visual information.

While early work showed little association between imagery ability and working memory (Heuer et al., 1986; Reisberg & Chambers, 1991; Reisberg & Leak, 1987), more recent reports show not only that the capacity to create mental images is positively associated with working memory (Albers et al., 2013; Keogh & Pearson, 2011, 2014), but also with several cognitive functions including spatial navigation (Ghaem et al., 1997), language comprehension (Bergen et al., 2007; Zwaan et al., 2002), and ethical decision and altruism (Amit & Greene, 2012; Gaesser & Schacter, 2014) to name a few. Although strong imagers were shown more sensitive to visual interference when presented with task-irrelevant stimuli during working memory and imagery tasks (Baddeley & Andrade, 2000; Keogh & Pearson, 2011, 2014), greater imagery ability often shows a facilitatory effect on perception through priming (Chang et al., 2013; Maróthi & Kéri, 2018; Pearson et al., 2008). Therefore, being a strong imager may appear to positively facilitate the processing of a wide range of cognitive functions.

One's ability to generate internal images, however, might not be a requirement for performing visuo-cognitive task. Jacobs et al. (2018) suggested that, while a significant portion of participants appear to employ visual imagery as a cognitive strategy to complete working memory tasks, imagery is not a critical component of working memory ability – in other words, imagery is of little functional relevance. This conclusion was driven by the fact that participants with varying degree of visual imagery ability can emulate the same performance on working memory (Jacobs et al., 2018). Generally, aphantasic individuals who report not experiencing mental images do not tend to show visuo-cognitive impairments (Zeman et al., 2015). In a case study, however, they found very subtle differences between an aphatasic and a group of matched-control non-aphantasic participants (Jacobs et al., 2018). Indeed, the aphatasic individual was assessed on a battery of working memory tasks and showed that they only performed significantly less on tasks that required high degree of precision than a control group, while also reporting lower meta-cognitive abilities. These reports suggest that the role of visual imagery is mainly as a fine-tuning mechanism, but is not necessary for the adequate performance of working memory tasks (Jacobs et al., 2018). Conversely, a study showed that a self-rated aphantasic group showed above average performance on visual imagery tasks, suggesting that they lacked meta-cognition (Keogh & Pearson, 2018). In other words, they were blind to their own mental representation instead of lacking imagery ability.

While this thesis does not address individual differences in the experience of mental representation of visual information, there is some evidence that participants are experiencing these cognitive representations differently and may adopt different visuo-cognitive strategies. Recent reports from the literature demonstrating that individual differences present important cognitive distinctions in how representations are expressed and can help shine a light on the functional purpose of this mental ability. Therefore, it appears important to consider individual difference for future research.

References

Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & de Lange, F. P. (2013). Shared Representations for Working Memory and Mental Imagery in Early Visual Cortex. Current Biology, 23(15), 1427–1431. https://doi.org/10.1016/j.cub.2013.05.065

Alvarez, G. A., & Cavanagh, P. (2004). The Capacity of Visual Short-Term Memory is Set Both by Visual Information Load and by Number of Objects. Psychological Science, 15(2), 106–111. https://doi.org/10.1111/j.0963-7214.2004.01502006.x

Amedi, A., Malach, R., & Pascual-Leone, A. (2005). Negative BOLD differentiates visual imagery and perception. Neuron, 48(5), 859–872. https://doi.org/10.1016/j.neuron.2005.10.032

Amit, E., & Greene, J. D. (2012). You See, the Ends Don't Justify the Means: Visual Imagery and Moral Judgment. Psychological Science, 23(8), 861–868. https://doi.org/10.1177/0956797611434965

Baddeley, A. D., & Andrade, J. (2000). Working memory and the vividness of imagery. Journal of Experimental Psychology: General, 129(1), 126–145. https://doi.org/10.1037/0096-3445.129.1.126

Baddeley, A. D., & Hitch, G. (1974). Working Memory. In G. H. Bower (Ed.), Psychology of Learning and Motivation (Vol. 8, pp. 47–89). Academic Press. https://doi.org/10.1016/S0079-7421(08)60452-1

Bays, P. M. (2018). Reassessing the Evidence for Capacity Limits in Neural Signals Related to Working Memory. Cerebral Cortex (New York, N.Y.: 1991), 28(4), 1432–1438. https://doi.org/10.1093/cercor/bhx351

Bays, P. M., & Husain, M. (2008). Dynamic Shifts of Limited Working Memory Resources in Human Vision. Science (New York, N.Y.), 321(5890), 851–854. https://doi.org/10.1126/science.1158023

Bergen, B. K., Lindsay, S., Matlock, T., & Narayanan, S. (2007). Spatial and Linguistic Aspects of Visual Imagery in Sentence Comprehension. Cognitive Science, 31(5), 733–764. https://doi.org/10.1080/03640210701530748

Campos, A. (2011). Internal Consistency and Construct Validity of Two Versions of the Revised Vividness of Visual Imagery Questionnaire. Perceptual and Motor Skills, 113(2), 454–460. https://doi.org/10.2466/04.22.PMS.113.5.454-460

Chang, S., Lewis, D. E., & Pearson, J. (2013). The functional effects of color perception and color imagery. Journal of Vision, 13(10), 4. https://doi.org/10.1167/13.10.4

Chaudhuri, R., & Fiete, I. (2016). Computational principles of memory. Nature Neuroscience, 19(3), 394–403. https://doi.org/10.1038/nn.4237

Christophel, T. B., Cichy, R. M., Hebart, M. N., & Haynes, J.-D. (2015). Parietal and early visual cortices encode working memory content across mental transformations. NeuroImage, 106, 198–206. https://doi.org/10.1016/j.neuroimage.2014.11.018

Cichy, R. M., Heinzle, J., & Haynes, J.-D. (2012). Imagery and perception share cortical representations of content and location. Cerebral Cortex (New York, N.Y.: 1991), 22(2), 372–380. https://doi.org/10.1093/cercor/bhr106

Costa, V. D., Tran, V. L., Turchi, J., & Averbeck, B. B. (2015). Reversal learning and dopamine: A bayesian perspective. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 35(6), 2407–2416. https://doi.org/10.1523/JNEUROSCI.1989-14.2015

Cui, X., Jeter, C. B., Yang, D., Montague, P. R., & Eagleman, D. M. (2007). Vividness of mental imagery: Individual variability can be measured objectively. Vision Research, 47(4), 474–478. https://doi.org/10.1016/j.visres.2006.11.013

Curby, K. M., & Gauthier, I. (2007). A visual short-term memory advantage for faces. Psychonomic Bulletin & Review, 14(4), 620–628.

Curtis, C. E., & D’Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. Trends in Cognitive Sciences, 7(9), 415–423. https://doi.org/10.1016/S1364-6613(03)00197-9

Delvenne, J.-F., & Bruyer, R. (2004). Does visual short - term memory store bound features? Visual Cognition, 11(1), 1 - 27. https://doi.org/10.1080/13506280344000167

D’Esposito, M., & Postle, B. R. (2015). The cognitive neuroscience of working memory. Annual Review of Psychology, 66, 115–142. https://doi.org/10.1146/annurev-psych-010814-015031

Emrich, S. M., Riggall, A. C., Larocque, J. J., & Postle, B. R. (2013). Distributed patterns of activity in sensory cortex reflect the precision of multiple items maintained in visual short-term memory. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 33(15), 6516–6523. https://doi.org/10.1523/JNEUROSCI.5732-12.2013

Eng, H. Y., Chen, D., & Jiang, Y. (2005). Visual working memory for simple and complex visual stimuli. Journal of Vision, 5(8), 611–611. https://doi.org/10.1167/5.8.611

Eriksson, J., Vogel, E. K., Lansner, A., Bergström, F., & Nyberg, L. (2015). Neurocognitive Architecture of Working Memory. Neuron, 88(1), 33–46. https://doi.org/10.1016/j.neuron.2015.09.020

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is “special” about face perception? Psychological Review, 105(3), 482–498. https://doi.org/10.1037/0033-295X.105.3.482

Foster, J. J., Sutterer, D. W., Serences, J. T., Vogel, E. K., & Awh, E. (2015). The topography of alpha-band activity tracks the content of spatial working memory. Journal of Neurophysiology, jn.00860.2015. https://doi.org/10.1152/jn.00860.2015

Gaesser, B., & Schacter, D. L. (2014). Episodic simulation and episodic memory can increase intentions to help others. Proceedings of the National Academy of Sciences, 111(12), 4415–4420.

Garcia, J. O., Srinivasan, R., & Serences, J. T. (2013). Near-Real-Time Feature-Selective Modulations in Human Cortex. Current Biology, 23(6), 515–522. https://doi.org/10.1016/j.cub.2013.02.013

Gauthier, I., & Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: Bridging brain activity and behavior. Journal of Experimental Psychology: Human Perception and Performance, 28(2), 431–446. https://doi.org/10.1037/0096-1523.28.2.431

Ghaem, O., Mellet, E., Crivello, F., Tzourio, N., Mazoyer, B., Berthoz, A., & Denis, M. (1997). Mental navigation along memorized routes activates the hippocampus, precuneus, and insula. NeuroReport, 8(3), 739–744.

Hajonides, J. E., Nobre, A. C., van Ede, F., & Stokes, M. G. (2021). Decoding visual colour from scalp electroencephalography measurements. NeuroImage, 237, 118030. https://doi.org/10.1016/j.neuroimage.2021.118030

Harris, I. M., & Miniussi, C. (2003). Parietal lobe contribution to mental rotation demonstrated with rTMS. Journal of Cognitive Neuroscience, 15(3), 315–323. https://doi.org/10.1162/089892903321593054

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. Nature, 458(7238), 632–635. https://doi.org/10.1038/nature07832

Haynes, J.-D., & Rees, G. (2005). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. Nature Neuroscience, 8(5), 686–691. https://doi.org/10.1038/nn1445

Heuer, F., Fischman, D., & Reisberg, D. (1986). Why does vivid imagery hurt colour memory? Canadian Journal of Psychology/Revue Canadienne de Psychologie, 40(2), 161–175. https://doi.org/10.1037/h0080090

Ishai, A., & Sagi, D. (1995). Common mechanisms of visual imagery and perception. Science (New York, N.Y.), 268(5218), 1772–1774.

Jackson, M. C., & Raymond, J. E. (2008). Familiarity enhances visual working memory for faces. Journal of Experimental Psychology: Human Perception and Performance, 34(3), 556–568. https://doi.org/10.1037/0096-1523.34.3.556

Jacobs, C., Schwarzkopf, D. S., & Silvanto, J. (2018). Visual working memory performance in aphantasia. Cortex, 105, 61–73. https://doi.org/10.1016/j.cortex.2017.10.014

Jordan, K., Heinze, H.-J., Lutz, K., Kanowski, M., & Jäncke, L. (2001). Cortical Activations during the Mental Rotation of Different Visual Objects. NeuroImage, 13(1), 143–152. https://doi.org/10.1006/nimg.2000.0677

Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. Nature Neuroscience, 8(5), 679–685. https://doi.org/10.1038/nn1444

Keogh, R., & Pearson, J. (2011). Mental Imagery and Visual Working Memory. PLoS ONE, 6(12). https://doi.org/10.1371/journal.pone.0029221

Keogh, R., & Pearson, J. (2014). The sensory strength of voluntary visual imagery predicts visual working memory capacity. Journal of Vision, 14(12), 7–7. https://doi.org/10.1167/14.12.7

Keogh, R., & Pearson, J. (2018). The blind mind: No sensory visual imagery in aphantasia. Cortex, 105, 53–60. https://doi.org/10.1016/j.cortex.2017.10.012

Kosslyn, S. M. (2005). Mental images and the Brain. Cognitive Neuropsychology, 22(3–4), 333–347. https://doi.org/10.1080/02643290442000130

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural Foundations of Imagery. Nature Reviews Neuroscience, 2(9), 635–642. https://doi.org/10.1038/35090055

Kosslyn, S. M., & Thompson, W. L. (2003). When is early visual cortex activated during visual mental imagery? Psychological Bulletin, 129(5), 723–746. https://doi.org/10.1037/0033-2909.129.5.723

Kosslyn, S. M., Thompson, W. L., & Alpert, N. M. (1997). Neural Systems Shared by Visual Imagery and Visual Perception: A Positron Emission Tomography Study. NeuroImage, 6(4), 320–334. https://doi.org/10.1006/nimg.1997.0295

Liesefeld, H. R., & Zimmer, H. D. (2011). The advantage of mentally rotating clockwise. Brain and Cognition, 75(2), 101–110. https://doi.org/10.1016/j.bandc.2010.10.012

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. Nature, 390(6657), 279–281. https://doi.org/10.1038/36846

Luria, R., Sessa, P., Gotler, A., Jolicoeur, P., & Dell'Acqua, R. (2010). Visual short-term memory capacity for simple and complex objects. Journal of Cognitive Neuroscience, 22(3), 496–512. https://doi.org/10.1162/jocn.2009.21214

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

Marks, D. F. (1973). Visual imagery differences in the recall of pictures. British Journal of Psychology (London, England: 1953), 64(1), 17–24.

Maróthi, R., & Kéri, S. (2018). Enhanced mental imagery and intact perceptual organization in schizotypal personality disorder. Psychiatry Research, 259, 433–438. https://doi.org/10.1016/j.psychres.2017.11.015

McKelvie, S. J. (1995). The VVIQ as a psychometric test of individual differences in visual imagery vividness: A critical quantitative review and plea for direction. Journal of Mental Imagery, 19(3–4), 1–106.

Parsons, L. M. (2003). Superior parietal cortices and varieties of mental rotation. Trends in Cognitive Sciences, 7(12), 515–517. https://doi.org/10.1016/j.tics.2003.10.002

Pearson, J., Clifford, C. W. G., & Tong, F. (2008). The functional impact of mental imagery on conscious perception. Current Biology: CB, 18(13), 982–986. https://doi.org/10.1016/j.cub.2008.05.048

Pearson, J., Naselaris, T., Holmes, E. A., & Kosslyn, S. M. (2015). Mental Imagery: Functional Mechanisms and Clinical Applications. Trends in Cognitive Sciences, 19(10), 590–602. https://doi.org/10.1016/j.tics.2015.08.003

Pylyshyn, Z. (2003). Return of the mental image: Are there really pictures in the brain? Trends in Cognitive Sciences, 7(3), 113–118. https://doi.org/10.1016/S1364-6613(03)00003-2

Pylyshyn, Z. W. (1979). The rate of “mental rotation” of images: A test of a holistic analogue hypothesis. Memory & Cognition, 7(1), 19–28. https://doi.org/10.3758/BF03196930

Reisberg, D., & Chambers, D. (1991). Neither pictures nor propositions: What can we learn from a mental image? Canadian Journal of Psychology, 45(3), 336–352. https://doi.org/10.1037/h0084297

Reisberg, D., & Leak, S. (1987). Visual imagery and memory for appearance: Does Clark Gable or George C. Scott have bushier eyebrows? Canadian Journal of Psychology/Revue Canadienne de Psychologie, 41(4), 521–526. https://doi.org/10.1037/h0084167

Samaha, J., Sprague, T. C., & Postle, B. R. (2016). Decoding and Reconstructing the Focus of Spatial Attention from the Topography of Alpha-band Oscillations. Journal of Cognitive Neuroscience, 28(8), 1090–1097. https://doi.org/10.1162/jocn_a_00955

Schurgin, M. W., Wixted, J. T., & Brady, T. F. (2020). Psychophysical scaling reveals a unified theory of visual memory strength. Nature Human Behaviour, 4(11), 1156–1172. https://doi.org/10.1038/s41562-020-00938-0

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

Shepard, R. N., & Cooper, L. A. (1986). Mental Images and Their Transformations (New edition edition). The MIT Press.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. Science (New York, N.Y.), 171(3972), 701–703.

Snow, J. C., Skiba, R. M., Coleman, T. L., & Berryhill, M. E. (2014). Real-world objects are more memorable than photographs of objects. Frontiers in Human Neuroscience, 8. https://doi.org/10.3389/fnhum.2014.00837

Song, J.-H., & Jiang, Y. (2006). Visual working memory for simple and complex features: An fMRI study. NeuroImage, 30(3), 963–972. https://doi.org/10.1016/j.neuroimage.2005.10.006

Sprague, T. C., Ester, E. F., & Serences, J. T. (2016). Restoring Latent Visual Working Memory Representations in Human Cortex. Neuron, 91(3), 694–707. https://doi.org/10.1016/j.neuron.2016.07.006

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. Memory & Cognition, 25(5), 583–592. https://doi.org/10.3758/BF03211301

Taubert, J., Apthorp, D., Aagten-Murphy, D., & Alais, D. (2011). The role of holistic processing in face perception: Evidence from the face inversion effect. Vision Research, 51(11), 1273–1278. https://doi.org/10.1016/j.visres.2011.04.002

Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. Nature, 428(6984), 751–754. https://doi.org/10.1038/nature02466

Tong, F. (2013). Imagery and visual working memory: One and the same? Trends in Cognitive Sciences, 17(10), 489–490. https://doi.org/10.1016/j.tics.2013.08.005

van den Berg, R., Awh, E., & Ma, W. J. (2014). Factorial Comparison of Working Memory Models. Psychological Review, 121(1), 124–149. https://doi.org/10.1037/a0035234

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. Nature, 428(6984), 748–751. https://doi.org/10.1038/nature02447

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

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. Journal of Vision, 4(12), 11. https://doi.org/10.1167/4.12.11

Xie, W., & Zhang, W. (2017). Familiarity increases the number of remembered Pokémon in visual short-term memory. Memory & Cognition, 45(4), 677–689. https://doi.org/10.3758/s13421-016-0679-7

Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. Nature, 440(7080), 91–95. https://doi.org/10.1038/nature04262

Zacks, J. M. (2008). Neuroimaging Studies of Mental Rotation: A Meta-analysis and Review. Journal of Cognitive Neuroscience, 20(1), 1–19. https://doi.org/10.1162/jocn.2008.20013

Zeman, A., Dewar, M., & Della Sala, S. (2015). Lives without imagery – Congenital aphantasia. Cortex, 73, 378–380. https://doi.org/10.1016/j.cortex.2015.05.019

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. Nature, 453(7192), 233–235. https://doi.org/10.1038/nature06860

Zhou, X., Mondloch, C. J., & Emrich, S. M. (2018). Encoding differences affect the number and precision of own-race versus other-race faces stored in visual working memory. Attention, Perception, & Psychophysics, 80(3), 702–712. https://doi.org/10.3758/s13414-017-1467-6

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language Comprehenders Mentally Represent the Shapes of Objects. Psychological Science, 13(2), 168–171. https://doi.org/10.1111/1467-9280.00430