From Individual Differences to Neural Substrates: The Role of Cognitive Control in Mitigating Interference in Working Memory

PhD in Neuroscience

School of Psychology and Clinical Language Sciences,

Centre for Integrative Neuroscience and Neurodynamics (CINN)

Olga Leticevscaia

September 2024

Declaration:

I, Olga Leticevscaia, confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Abstract

Working memory, the ability to hold information in mind over the short term towards a behavioural goal, plays a fundamental role in many of the cognitive processes we experience as humans. Yet distractions and interruptions can interfere with this ability. Cognitive control plays a key role in maintaining information in working memory and resolving interference, but the neural underpinnings of this process are not fully understood. This thesis explores the neural mechanisms of cognitive control in resolving interference from irrelevant information in visual working memory.

Chapter 2 explored how visual distractions impacted memory precision of naturalistic objects, demonstrating that working memory remains robust despite visually engaging distractors. Chapter 3 demonstrated that the dorsolateral prefrontal cortex (dlPFC), crucial for cognitive control, supports resilience against distractions. Suppressing dlPFC activity with transcranial magnetic stimulation (TMS) reduced memory recall precision, providing causal evidence of its role in mitigating distractor effects.

To further investigate cognitive control, Chapters 2 and 3 examined the relationship between individual differences in mind-wandering, trait worry, and distractor mitigation. Whilst worry did not impair distraction mitigation—possibly due to compensatory mechanisms—in a surprising finding, individuals with less control over mind-wandering benefited from distractors, in which they may have helped to refocus attention away from internal thoughts. In addition, pupillometry revealed that fluctuations in cognitive control, reflected by pupil dilation, predicted memory performance, and reflected cognitive efforts exerted in the presence of visual distractors.

Chapter 4 investigated how cognitive control protects memory contents from interruptions. Using multivariate pattern analysis (MVPA) of functional magnetic resonance imaging (fMRI) data, memory content was decoded within the Multiple Demand (MD) network during interruption tasks. Results showed that more challenging tasks enhanced memory representations within MD regions, reflecting the nature of cognitive control mechanisms in this context.

These findings deepen our understanding of how the MD system supports flexible behaviour, and provide valuable insights for clinical interventions, i.e. neuromodulation, aiming to improve resistance to interference in working memory in conditions such as attention deficit hyperactivity disorder, schizophrenia, and dementia.

Acknowledgements

I would like to thank the School of Psychology and Clinical Sciences of the University of Reading for providing me with the opportunity to explore the fascinating world of neuroscience by funding my research.

This research also would not be possible without the constant support, guidance and encouragement of my supervisor Dr Eva Feredoes.

I am also grateful to Dr Nicolo Biagi for his help with the TMS experiment and his wise advice. In addition, I would like to thank Theresa Wong and Zaneta Benova for their valuable assistance with fMRI data collection.

I especially would like to thank my family for their support. I am grateful to my mum who always believed in me and encouraged my curiosity, and my grandad who always inspired me to become a researcher. I would not be able to pursue my dream without the support of my husband who knows now too much about the brain and working memory. Last, but not least, I am grateful to my daughter who I hope will be proud of me when she grows up.

List of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

List of Tables

Chapter 2

Chapter 3

Chapter 4

Chapter 1. General Introduction

Background

The ability to hold information in mind referred to as working memory (WM) (Baddeley, 1992) is essential for everyday life tasks. We rely on visual WM to remember things, such as keeping in mind an image of the package of biscuits you while grocery shopping. However, it can be impacted by irrelevant information: a bright advertisement for another brand of biscuits can distract you and cause you to forget the package you intended to purchase. This type of irrelevant information is known as a distractor—an unrelated stimulus that disrupts your ability to stay on task by diverting your focus away (Lorenc et al., 2021). Another type of irrelevant information is an interruption—an event or action that temporarily halts or disrupts an ongoing task (Couffe & Michael, 2017; Foroughi et al., 2016). For example, bumping into a friend at the store or getting a message on your phone can also cause you to forget the biscuits' image you held in mind.

When faced with such distractions or interruptions, cognitive control - the ability to align actions and behaviour with internal goals- is crucial for protecting and maintaining memory representations. While cognitive neuroscience research offers valuable insights into how these control mechanisms might be implemented in the brain (Menon, & D'Esposito, 2022; Miller, 2000; Badre, 2008), the findings for cognitive control involvement in protecting WM remain mixed (D'Esposito & Postle, 2015). Compelling evidence for dorsolateral PFC (dlPFC) as a key region in eliciting control mechanisms in WM comes from studies of patients with lesions in this region, who demonstrate impairments in WM (Voytek & Knight, 2010). Importantly, some patients still can hold information in mind, but their ability to resolve interference from irrelevant information is impaired (Chao & Knight, 1995; Chao & Knight, 1998; Baldo & Shimamura, 2000). Additional support for dlPFC involvement in cognitive control in WM in healthy population comes from studies using transcranial magnetic stimulation (TMS) to perturb activity in specific regions of the prefrontal cortex and have also demonstrated the causal involvement of this region in cognitive control over memory maintenance (Oliveri et al., 2001; Postle et al., 2006), and resolving visual distraction, in particular (Feredoes et al., 2011).

However, neuroimaging studies revealed that regions outside the prefrontal cortex, such as early visual areas (Serences et al., 2009; Harrison & Tong; 2009), parietal cortex (Belger et al., 1998), medial temporal lobe (Nichols et al., 2006), are also activated during memory maintenance. Neural decoding methods applied to neuroimaging data reveal how patterns of brain activity encode information, providing insights into neural representations and how these patterns relate to cognitive processes, behaviour, and perception in a way that is more sensitive and informative than traditional analysis methods (Robinson et al., 2023). This technique has been applied to assess the impact of visual distractions on memory, revealing distractor-resistant memory codes within visual (Hallenbeck et al., 2021) and parietal regions (Lorenc et al., 2018; Rademaker et al., 2019). Despite the importance of these regions in protecting WM, their interaction with dlPFC-driven cognitive control remains unclear (Curtis & Sprague, 2021). For example, one study found that the functional connectivity of the prefrontal cortex and the visual cortex increases during external interference (Clapp et al., 2010). This finding indicates that the interaction between these areas is essential for maintaining cognitive control over memory during interference, demonstrating that cognitive control is a distributed process involving the coordinated activity of multiple brain regions. Therefore, studying the cognitive control of WM from a network perspective is crucial, as it offers a more comprehensive understanding of this process, highlighting that these connections might form a dynamic, adaptive system rather than merely linear or hierarchical (Cocchi et al., 2013).

One of the candidate networks for such a control is a set of regions in the parietal and frontal cortex, referred to as the Multiple-Demand (MD) network. It is proposed that the multiple-demand (MD) network plays a crucial role in managing cognitive control by breaking down complex tasks into more focused sub-tasks, thereby facilitating efficient problem-solving and goal-directed behaviour (Duncan, 2010). MD regions are consistently engaged during a variety of demanding tasks, such as those requiring WM, selective attention, and problem-solving (Fedorenko et al., 2013; Assem et al., 2020). This network may achieve its function by dynamically adapting neural activity to meet current task demands, selectively encoding relevant information in the face of irrelevant information. This adaptability is supported by recent studies using decoding approaches, which have provided evidence for MD network's capacity to code relevant object features in the presence of irrelevant ones (Jackson & Woolgar, 2018; Jackson et al., 2017). In addition, the activation of MD network may predict individual differences in cognitive control within WM; stronger engagement of MD network during a spatial WM task was robustly associated with more accurate and faster performance (Assem et al., 2020).

Individual differences in cognitive control could also be reflected in the ability to control interference from internal sources of distraction, i.e., you forget to buy the biscuits because you are immersed in your thoughts while thinking about the book you read recently. In this example, you were mind wandering - it is a mental state related to increased internal distraction by shifting thoughts away from a task or the external environment to internal, self-generated thoughts (Smallwood & Schooler, 2015). Although these internal thoughts can be pleasant, such as thinking about a book, they can also be negative, manifesting as worry or rumination. Worry is related to uncontrollable and excessive thoughts regarding uncertain events in the future, and it is often increased in anxiety (Borkovec et al., 1998). The research found that individuals prone to these types of internal distractions, such as mind wandering, worry or rumination, demonstrate lower cognitive control over WM (Unsworth & Robison, 2016; Robison and Unsworth, 2018; Sari et al., 2017; Gustavson & Miyake, 2015; Bruning et al., 2023). This finding highlights the crucial role of cognitive control in managing both external and internal interference within WM.

This thesis aims to contribute to the understanding of cognitive control mechanisms, specifically focusing on how they might resolve interference within WM. In the first part, I explore the mechanism for resolving external visual distractions in WM. I also test whether individuals prone to internal distractions, such as mind-wandering or worrying, exhibit reduced cognitive control over external distraction. The second part of this thesis is dedicated to the investigation of cognitive control protecting memory representations from interference caused by an interrupting task.

Throughout this thesis, I refer to cognitive control mechanisms related to external distractions as *distractor mitigation* to emphasise the active process of minimising the distracting stimulus's impact on the memorandum by reducing the amount of cognitive effort exerted for its processing. In contrast, cognitive control in the face of interrupters requires not only processing the interrupting task but also taking action to *protect the memorandum* from disruption.

I would like to note that I do not assume there is one cognitive control mechanism for both external distractions and interrupters since these types of interference result in different impacts on memory maintenance and have slightly distinct neural underpinnings (Clapp et al., 2010). However, MD network may potentially coordinate multiple distinct control mechanisms under a unified framework.

The present chapter starts with a brief discussion on the nature of WM followed by a methodological overview of the behavioural and neural measures of WM which are essential for examining the interference effects of irrelevant information on the memorandum. In addition, in this section, I review how the pupil response can be used as a proxy for cognitive effort directed to the processing of irrelevant information or for active memory maintenance, as well as how the variability of pupil response may reflect the fluctuations in cognitive effort throughout the experiment. Next, I present relevant evidence for distractor mitigation mechanisms, drawing from behavioural, neuroimaging, lesion and neural stimulation studies. Lastly, I discuss how individual differences in cognitive control over internal distractions, reflected in increased mind wandering or worrying, influence the ability to resolve external distractions in WM.

Behavioural and neural approaches to WM assessment

In this section, I provide a brief overview of the key methodological approaches used in WM assessment. This foundation is essential for understanding how interference in WM can be evaluated through behavioural and neuroimaging techniques, neural stimulation and pupillometry.

The nature of WM and behavioural approach

To assess how well information is maintained in memory over the short term, behavioural research often uses a match-to-sample approach (Anderson & Colombo, 2022): an item (*memory target*) is displayed for memorisation, then it disappears from the screen for several seconds (*delay period*), and after that, a *probe*, same or different item(s), appears on the screen. Participants must respond if the probe matches the memorised item or not. The effect of interference is calculated as the difference in recall accuracy, i.e. how many responses were correct, in conditions with or without interfering stimulus. Despite being a mainstay of WM research (Anderson & Colombo, 2022), this approach has a key limitation as it assumes that memories are discrete (Luck & Vogel, 1997). However, resource models of WM challenge this idea (Ma et al., 2014). Namely, they demonstrate that memory precision declines gradually as the number of items increases indicating that memory resources can be flexibly allocated using a continuous recall measure of WM fidelity (Ma et al., 2014). Continuous recall refers to a method of assessing WM where participants are asked to reproduce or estimate the precise value of a remembered item on a continuous scale, rather than selecting from a set of discrete options (see Figure 1 for examples).

In paradigms testing orientation memory, participants are asked to reproduce the orientation of a memorized item, such as a bar, Gabor patch, or pointer, which is displayed at recall in a random location (Burnett Heyes et al., 2012; Salmela et al., 2013). For colour recall, participants are required to select the colour they remembered from a colour wheel displaying the full spectrum (Zhang & Luck, 2008; Nemes et al., 2011; Bays, Wu, & Husain, 2011; Wilken & Ma, 2004). In studies of face memory, researchers have created a continuous (circular) space by generating a set of 80 grayscale 3D face images that vary systematically along two dimensions—age and gender (Lorenc et al., 2014). During recall, a randomly selected face is displayed on the screen, and participants must navigate through the face space to select the face they remembered. Similarly, realistic objects can be transformed into a continuous array with equal incremental changes, using a diffeomorphic algorithm (Cusack & Stojanoski, 2014), to assess memory precision for naturalistic objects (Veldsman et al., 2017; Stojanoski et al., 2019). In these paradigms, participants choose the exact image they memorized from a set of exemplars displayed on a wheel. The precision of memory recall in these continuous report tasks is calculated by measuring the angular difference between the correct item (the memory target) and the item selected by the participant during recall. The magnitude and distribution of these errors provide a measure of memory precision, with smaller errors indicating better object recall precision. The impact of irrelevant information

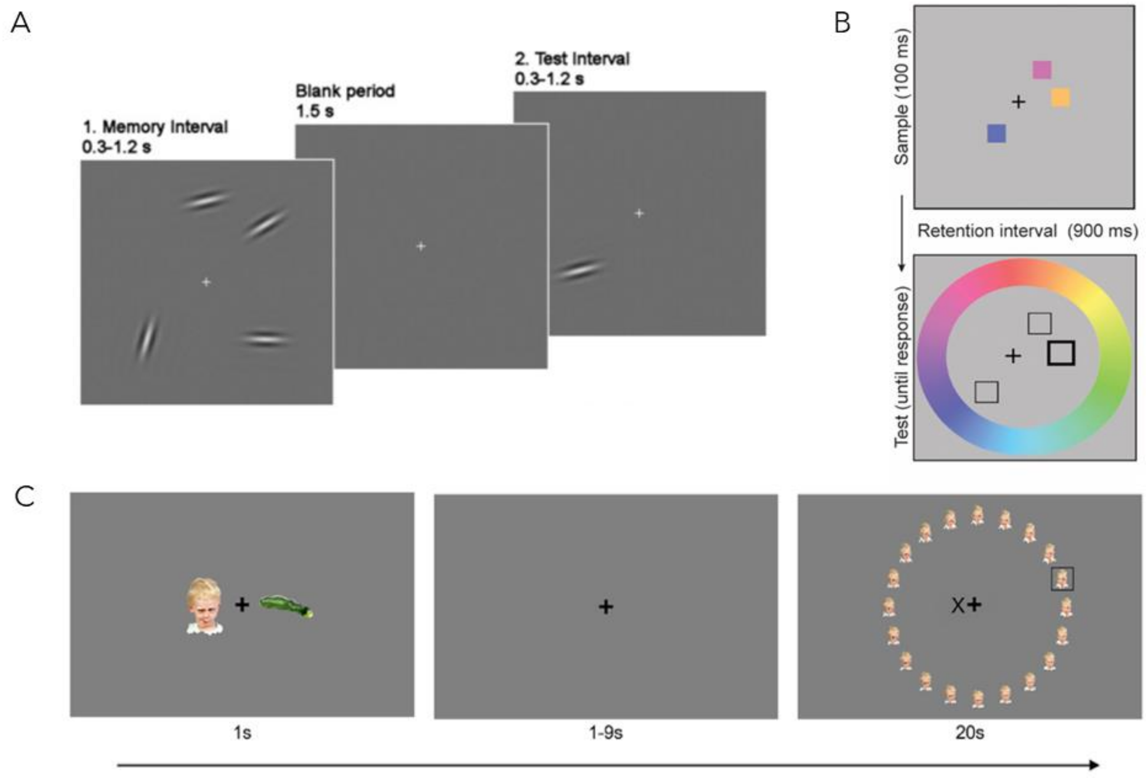


Figure 1. Examples of the continuous recall paradigms. Continuous recall was applied to study memory precision for orientation (panel A; Salmela et al., 2013), colour (panel B; Zhang & Luck, 2008) or realistic objects morphed to create a continuous measure of object precision (Panel C; Veldsman et al., 2017).

These studies suggest that continuous recall can be used to study memory quality by providing a more precise assessment of WM representations. This approach is particularly useful for exploring how memory representations are impacted by interference from irrelevant information, allowing us to measure the fine-grained changes in memory precision (Teng & Krawitz, 2019; Nemes et al. 2012; Rademaker et al., 2019; Rademaker et al., 2015; Lorenc at al., 2018; Mallett et al. in 2020).

The following sections below will provide a brief outline of some cognitive neuroscience approaches that have been most useful in studying WM and motivated the choice of methodology in this thesis.

Neuroimaging approach

Neuroimaging, broadly defined as "a set of techniques that produce images of the structure or activity of the brain or other parts of the nervous system" (APA Dictionary of Psychology, 2023) includes several approaches, but one of the most widely applied in the cognitive neuroscience is functional magnetic resonance imaging (fMRI). fMRI relies on detecting changes in blood oxygen levels. When a specific brain area is more active, it consumes more oxygen. The body responds by increasing blood flow to that region, bringing in oxygen-rich blood. fMRI detects these changes, known as the Blood Oxygen Level-Dependent (BOLD) signal, which serves as a proxy for neural activity (Huettel et al., 2014; Logothetis, 2008; Bandettini, 2009). To identify brain regions involved in WM, researchers can analyse the BOLD response obtained while participants perform a WM task during fMRI scanning. The magnitude of this response is taken as the level of involvement of a specific brain region during the task, a method commonly referred to as *univariate analysis*. However, a significant limitation of the univariate approach is its lack of sensitivity to distributed patterns of brain activity, making it less suitable for studying memory representations, which are inherently distributed across multiple brain regions (Haxby et al., 2001; Norman et al., 2006).

An alternative approach to analysing fMRI data that can effectively capture these representations is multivariate pattern analysis (MVPA). fMRI data is acquired as volumetric images whose smallest parts are called *voxels* (from 'volume pixel'). Univariate analysis detects BOLD signal values within each voxel that pass a statistical threshold, at which point it is an 'activation'. In contrast, MVPA considers the pattern of activity across multiple voxels simultaneously, identifying distributed patterns that encode specific information about cognitive states or stimuli (Robinson et al., 2023) and can be particularly useful to measure how memory-related information is distributed across various brain regions and how the processes within WM change this information.

Altogether, fMRI provides a powerful tool to explore memory-related activations with decoding techniques being a particularly useful approach to study representations in WM and to provide anatomical localisation on where memoranda are represented in the brain.

TMS approach

The necessity to causally examine multiple cognitive functions in healthy populations, rather than relying on the limited number of lesion patients, led to the development of neural stimulation techniques such as TMS(TMS), which have been applied to cognitive neuroscience questions (Walsh & Cowey, 2000).

TMS modulates neural activity by inducing electric currents in targeted brain regions, which can either enhance or inhibit neuronal firing, thereby altering synaptic plasticity and functional connectivity (Cirillo et al., 2017).

Several protocols of stimulation have been developed, but one of the widely used in cognitive neuroscience is repetitive TMS (rTMS). The protocol delivers multiple TMS pulses in sequence, resulting in prolonged modulation of brain activity and connected regions that persists even after the stimulation has ended (Klomjai et al., 2015).

A version of the rTMS protocol called continuous theta burst stimulation (cTBS) (Huang et al., 2005) is widely used to study cognitive functions and the role of the prefrontal cortex in cognition, in particular (Lowe et al., 2018). cTBS uses gamma frequency trains at theta rhythm (Huang et al., 2005) and can be applied in a brief time interval (40 seconds) for a lasting effect for ~50 min (Wischnewski and Schutter, 2015). Although the exact mechanisms of cTBS are unclear, it can produce an inhibitory effect in the stimulated region (Wischnewski and Schutter, 2015).

Furthermore, TMS can be combined with fMRI (concurrent TMS-fMRI) to observe and understand the immediate effects of TMS on brain activity with high spatial resolution (Mizutani-Tiebel et al., 2022). This sophisticated technique reveals the effect of TMS not only in the site of stimulation but also in the network of regions connected to it (Ruff et al., 2009). When the network is actively engaged in a cognitive task, the modulation of neuronal activity in the directly stimulated region (such as increased or decreased firing rates) can propagate to other connected areas. This propagation of activity is crucial and only occurs when stimulation is applied during active network engagement (Ruff et al., 2009).

In summary, TMS offers causal evidence of a brain region's involvement in a cognitive function. Importantly, this technique is non-invasive and the changes it produces are transient: single pulse effect lasts milliseconds and repetitive protocols may induce an effect lasting a few hours making it safe to use to study cognitive functions in healthy populations (Walsh & Cowey, 2000; Pascual-Leone et al., 2000).

Pupillometry approach

Pupil dilation has proven to be a useful approach to measure arousal and effort deployment during cognitive tasks (Van der Wel et al., 2018; Cronin et al., 2023). The neural basis for pupillary modulation is thought to involve the locus coeruleus-norepinephrine (LC-NE) neuromodulatory system (Aston-Jones & Cohen, 2005; Usher et al., 1999). Research indicates that the LC operates in two firing modes: tonic and phasic (Aston-Jones & Cohen, 2005; Usher et al., 1999). Baseline pupil diameter is linked to the LC's tonic firing rate, reflecting overall task engagement, while task-evoked pupil dilations correspond to LC phasic activity, indicating the effort directed toward a task stimulus (Alnaes et al., 2014). Therefore, the variability in tonic pupil response—how consistent or variable it is over time—may be indicative of the strength of cognitive control during task performance (Unsworth & Robison, 2015, 2016).

Phasic pupil response serves as an indicator of the processing effort allocated to a stimulus. For example, a larger initial pupil dilation as a response to a memory target presentation was associated with better memory recall (Kucewicz et al., 2013). Additionally, the following pupil constriction may reflect the depth of stimulus processing, as greater constriction in response to image presentation was linked to better recall of novel naturalistic images (Naber et al., 2013).

In summary, by capturing both tonic and phasic responses, pupillometry serves as a valuable tool for investigating dynamic processes within WM, offering a real-time measure of cognitive effort directed toward a task or stimulus. Therefore, it could be used to study the effects of interference in WM.

Neural mechanisms of interference mitigation in WM

To study the effect of external interference on WM, during memory maintenance researchers can introduce a stimulus to be ignored or not, i.e. distractor or an interruption, e.g. a task requiring a response that can't be ignored. Both types of interference decrease memory recall accuracy (Clapp et al., 2010) or precision (Rademaker et al., 2015), and, hence, interfere with memory maintenance process.

For example, Rademaker et al. (2015) showed that memory precision for grating orientation was reduced when a distractor grating was presented during the delay period. Moreover, the orientation of the remembered grating was biased towards the orientation of the distractor (i.e. recall error was directed towards distractor orientation). Interestingly, the distracting grating impacted the memorandum only when it was consciously perceived (Rademaker et al., 2015) suggesting that when the distractor reaches awareness, it directs the cognitive effort away from the memory maintenance resulting in interference.

Clapp et al. (2010) examined both external distraction and interruption effects on memory recall for faces or scenes: participants were asked to memorise faces or naturalistic scenes, and an irrelevant face image appeared on the screen during the maintenance period. Participants were instructed to ignore the image (distractor) or to respond if the face was a male over forty years old (interruption). Compared to the unfilled delay, both types of interference led to lower memory recall accuracy, but interruption had a larger disruptive effect (Clapp et al., 2010). In addition, there was no interference effect of a face image (interfering stimulus) presented during memory maintenance on scene recall demonstrating that perceptual stimuli that share the processing with maintained information leads to a larger interference effect (Clapp et al., 2010).

The study used neuroimaging to measure the functional connectivity between the dorsolateral prefrontal cortex (dlPFC) and its connection to the visual association cortex, VAC (Clapp et al., 2010). During distraction, the connectivity between dlPFC and VAC was sustained, supporting memory maintenance. In contrast, interrupters disrupted this connectivity leading to a drop in memory recall accuracy and subsequent reinstatement of the connection between dlPFC and VAC. Stronger connectivity between dlPFC and VAC predicted better memory recall both during the distraction and followed by interruption (Clapp et al., 2010) suggesting that dlPFC controls active memory maintenance in sensory regions. These findings demonstrate that the neural mechanisms of interference between distractors and interrupters are distinct, but cognitive control elicited by dlPFC plays a role in both.

Furthermore, studies using the neural decoding approach explored the role of regions outside PFC in WM (Bettencourt & Xu, 2016; Lorenc et al., 2018; Rademaker et al., 2019; Hallenbeck et al., 2021; Vu-Cheung et al., 2021). Namely, Bettencourt and Xu (2016) found that visual distractions, such as flickering images of faces or gazebos, impaired orientation decoding in the visual cortex but not in the superior intraparietal sulcus (sIPS). This finding suggests that the IPS preserves memory representations by protecting them from incoming sensory information (Bettencourt and Xu, 2016).

Similarly, Lorenc et al. (2018) demonstrated the role of this region in protecting memoranda from distraction. Researchers used distractors and memory targets of the same type: participants reproduced the orientation of a memorised item when a distracting orientation was displayed during the delay period. The decoding analysis revealed that memory representations in visual cortex were impacted by distractors; specifically, the orientation representations were biased toward distractor orientations due to overlap in sensory areas. This finding explains why similar distractors can lead to a more detrimental impact on memory recall. The reliable decoding of orientations in the sIPS occurred only when distracting stimuli were present during the delay period, not when the delay was unfilled, pointing to the key role for this region in mitigating distraction.

Rademaker et al. (2019) further explored the IPS's role in distractor mitigation in WM. Researchers found that in the presence of visual distractors (orientations, faces, or gazebos) during memory maintenance, the representational format of memorised orientations in the IPS shifted from a sensory to a mnemonic format to resist interference, while representations in the visual cortex remained in a sensory format more susceptible to distortion. In contrast, memory representations in the visual cortex also can be robust against distractions (Hallenbeck et al., 2021; Vu-Cheung et al., 2021). For example, Hallenbeck et al. (2021) demonstrated that visual polar angles could still be reliably decoded in the sensory cortex, as well as in parietal and frontal regions, even after distracting stimuli (black and white moving dots) were presented during the delay period. Similarly, Vu-Cheung et al. (2021) found no difference in decoding accuracy for the spatial location of a memorized coloured dot in retinotopic areas, regardless of whether a distracting radial checkerboard was presented during the delay. Collectively, these findings explore the involvement of regions outside the PFC in memory maintenance and protection from distraction (Hallenbeck et al., 2021). Together, evidence for the involvement of both dlPFC and posterior cortices in memory protection indicates that these neural structures must work together to protect memoranda from external interference (Curtis & Sprague, 2021). In their review of the existing evidence for persistent activity during memory maintenance in PFC and posterior areas, Curtis & Sprague (2021) suggest that decoding studies highlight the capacity of sensory and parietal regions to maintain memory representations, while PFC-related research underscores the importance of cognitive control in managing and protecting these representations from interference. Yet, the precise interplay of PFC and posterior cortices is not known.

Earlier human fMRI studies (Sakai et al., 2002; Dolcos et al., 2007) proposed potential roles for these neural structures in distractor mitigation in visual WM. In their neuroimaging study, Sakai et al. (2002) demonstrated that IPS maintained stable activity irrespective of distraction, while preparatory activity (prior to distractor presentation) in dlPFC was associated with better performance in a spatial WM task. In addition, dlPFC activation was correlated with increased coupling of IPS and another frontal area called superior frontal sulcus (SFS), suggesting the mechanism for distractor mitigation; specifically, by dlPFC-driven active maintenance and enhancement of the stable memory representations maintained in posterior regions (Sakai et al., 2002). Using human neuroimaging to examine WM for faces during a delay period with varying levels of distraction—high (faces), low (scrambled faces), or none—Dolcos et al., (2007) found that activation in IPS remained consistent across all distraction conditions. In contrast, dlPFC activity increased in response to high distraction demonstrating the cognitive control over memorandum by its active maintenance. In addition, ventrolateral prefrontal cortex (vlPFC) activity gradually increased with increasing complexity of distractors suggesting its key role in distractor inhibition. Together, these two studies suggest a mechanism for distractor mitigation in which dlPFC exerts cognitive control over memoranda stored in posterior regions, enhancing their stability and resilience against distraction.

Furthermore, recent advances in monkey electrophysiology have offered a more granular insight into the mechanism of distractor mitigation implemented by the frontal and parietal regions (Suzuki and Gottlieb, 2013; Jacob and Nieder, 2014; Parthasarathy et al., 2017; Parthasarathy et al., 2019). Specifically, Suzuki and Gottlieb (2013) recorded electrophysiology responses in dlPFC and lateral intraparietal area (LIP) in monkeys performing a WM task in which a location of a light flash had to be memorised while a distracting flash was presented during the delay period; to respond at the recall, monkeys had to direct their gaze to the memorised location. The recordings revealed that anticipatory activity and reduced responses to distractors in dlPFC neurons were associated with better task performance. Furthermore, the inactivation of the dlPFC increased susceptibility to distractors, highlighting its critical role in inhibiting distractor processing to support action-oriented behavior (Suzuki and Gottlieb, 2013). In contrast, the association of the neural responses to distraction in LIP and distractor resistance was much weaker pointing to the role of this region in perceptual processing rather than modulating response to guide behaviour. Importantly, the target activity in LIP was enhanced throughout the delay period with a brief disruption during the distractor presentation demonstrating the complementary role of this region in protecting the memorandum from distraction by maintaining its stable representation (Suzuki and Gottlieb, 2013).

Jacob and Nieder, (2014) investigated this mechanism further by recording neural activity in dlPFC and the parietal region called the ventral intraparietal area (VIP) of monkeys performing a numerosity WM task. Animals were asked to memorise the number of items (i.e. circles of variable size) in a display while a distracting display could appear during the delay period. Contrary to the lower neural response to distractors found by Suzuki and Gottlieb, (2013), researchers demonstrated that dlPFC neurons hold information about both targets and distractors: target-related activity was decreased briefly during distractor presentation and then was restored (Jacob and Nieder, 2014). On the contrary, target representation in VIP remained enhanced in the face of distraction and predicted subsequent performance on the task. Despite conflicting patterns of results in dlPFC recordings, both these studies demonstrate that dlPFC and parietal lobe work together to protect memorandum by maintaining stable memory representation in parietal areas and flexibly modulating this representation to the goal-directed action, i.e. by suppressing distractor-processing (Suzuki and Gottlieb, 2013) or representing distractor information for adapting future behavior, e.g. distractor learning (Jacob and Nieder, 2014).

Parthasarathy et al., (2017) further explored this adaptive mechanism by testing the fate of memory code within lateral PFC (LPFC) and frontal eye field (FEF) during distraction while monkeys performed a spatial WM. Researchers found that memory code in LPFC morphed into a new stable representation when the distractor appeared, preserving memory information and mitigating the impact of distraction (Parthasarathy et al., 2017). This process is called code morphing and allows LPFC to adapt dynamically to the task at hand without losing critical information. Authors propose that adaptive code morphing is implemented by neurons with nonlinear mixed selectivity (NMS) which are found in abundance in

LPFC. NMS neurons are proposed to integrate multiple task parameters (e.g., target location and task epoch) nonlinearly, increasing the dimensionality of the population code within LPFC. In contrast, the FEF retained a consistent code but with substantial information loss after the distractor presentation, possibly due to a lower number of neurons demonstrating mixed selectivity. Parthasarathy et al. (2019) further analysed this dataset and identified distinct subspaces within the LPFC that encoded memory. Their findings revealed that despite the code-morphing induced by distractors, a low-dimensional subspace in the LPFC preserved stable memory representations. Importantly, the trials with errors showed reduced information in the identified subspace, suggesting its behavioral relevance (Parthasarathy et al., 2019). Collectively, monkey neurophysiology studies (Suzuki and Gottlieb, 2013; Jacob and Nieder, 2014; Parthasarathy et al., 2017; Parthasarathy et al., 2019) expand and offer the neural mechanism of distractor mitigation suggested by earlier human neuroimaging findings (Sakai et al., 2002; Dolcos et al., 2007); specifically, cognitive control may be implemented by morphing memoranda code within dlPFC and preserving low-dimensional memory information for guiding behavior (Parthasarathy et al., 2019); while areas posterior dlPFC such as FEF (Parthasarathy et al., 2017) and parietal lobe (Suzuki and Gottlieb, 2013; Jacob and Nieder, 2014) maintain a stable memory code which may potentially serve as a source of detailed memory information that dlPFC may flexibly recruit to optimise goal-directed behavior under varying task demands. However, further research is needed to confirm this neural mechanism in humans.

Moreover, it is important to consider that the interplay between storage mechanisms and cognitive control may be particularly crucial in scenarios in which interruptions occur. Yet, few decoding studies addressed the impact of interruptions on memory representations (Kiyonaga et al., 2017) which means that more decoding studies on interruption in WM are needed. In their study, Kiyonaga et al., (2017) asked participants to memorise either one or two items (faces or houses) for a later memory probe. During delay, participants performed a visual search (interrupter) task in which they had to identify a vertical target among distractors (tools or bodies) that were either easy or hard to differentiate from the target. The study found that decoding accuracy in sensory regions was influenced by the difficulty of the visual search, suggesting that cognitive resources are distributed between maintaining memory representations and processing external stimuli (Kiyonaga et al., 2017). Univariate fMRI Analysis revealed an interaction effect in the PFC when both WM load and visual search difficulty were manipulated. Specifically, the PFC showed heightened activity when both the WM load was high and the visual search task was difficult (Kiyonaga et al., 2017). This indicates that the PFC was particularly engaged when there were high demands for cognitive control over memoranda and interruption suggesting that this mechanism may be the key to resolving interruption-based interference in WM.

The role of the dlPFC was demonstrated in earlier studies of patients with lesions establishing a causal relationship between this region and cognitive control functions, and addressing the limitations of the correlational nature of neuroimaging studies in humans. For example, Voytek and Knight (2010) conducted a study in which healthy controls and patients with unilateral PFC lesions performed a lateralized visual WM task in which coloured squares were presented on one side of the screen). Patients performed comparably to controls when stimuli were presented unilateral to the lesion but exhibited impairments when stimuli appeared contralaterally (Voytek & Knight, 2010). Additionally, electrophysiological markers of top-down facilitation by cognitive control over sensory areas (such as alpha power and the N1 component) and memory maintenance (such as contralateral delay activity, CDA) was disrupted in PFC patients, highlighting the neural effects of lesions on memory processes.

To investigate the causal role of dlPFC in ability to protect WM from interference, Baldo and Shimamura (2000) tested patients with dlPFC lesions on spatial and colour WM using continuous scales—measuring the distance between a dot on the screen and the response or the distance between the presented colour and the response on a colour wheel spectrum. An interfering digit task was introduced during the memory delay period. Patients with PFC lesions made larger errors than healthy participants on both tasks and were more susceptible to the disruptive effects of the interfering task (Baldo & Shimamura, 2000). Notably, two influential lesion studies by Chao and Knight (1995, 1998) suggest that dlPFC plays a critical role in mitigating the effects of distraction in auditory WM. In their tasks, participants were required to memorize environmental sounds and, after a delay, determine whether a probe sound matched the memorized sound, with irrelevant tone pips serving as distractors during the delay period. Patients with dlPFC lesions exhibited lower memory recall accuracy than healthy controls, with distraction significantly impacting only the patients' performance, leading to more errors under distraction compared to healthy controls (Chao & Knight, 1998).

Lesion studies offer invaluable causal evidence for the role of dlPFC in mitigating distractions; however, a more flexible approach for causally studying brain function—one that is non-invasive, reversible, and applicable to healthy populations—is TMS. Numerous studies have explored the involvement of dlPFC in WM using TMS (Oliveri et al., 2001; Postle et al., 2006; Esslinger et al., 2014; Schicktanz et al., 2015; Vékony et al., 2018). For instance, Oliveri et al. (2001) applied single TMS pulses over dlPFC during the delay period of both visual-object and visual-spatial WM tasks, which affected both reaction times and accuracy. These findings support a process-specific model, suggesting that dlPFC is crucial for higher-level executive functions in WM, such as monitoring and manipulating information, rather than being tied to a specific type of information (spatial or object-related).

Repetitive TMS (rTMS) protocols have been also used to test the role of dlPFC in WM. Postle et al. (2006) instructed participants to memorize an array of letters that needed to be either stored in their original order ("forward" trials reflecting retention) or rearranged alphabetically ("alphabetize" trials involving manipulation) while applying rTMS over dlPFC during the delay period. The rTMS only affected response accuracy during the "alphabetize" trials, suggesting that dlPFC is more involved in manipulating information within WM rather than simply retaining it. Similarly, Esslinger et al. (2014) applied rTMS over dlPFC at a frequency of 5 Hz before blocks of a 2-back task (with digit displays) to enhance activity in this region, contrasting with most studies that use stimulation to disrupt neural activity. Following the stimulation, functional magnetic resonance imaging (fMRI) was used to acquire resting-state data to assess the connectivity between dlPFC and other regions involved in WM processing. This stimulation led to faster reaction times during the task, further emphasizing dlPFC's importance in WM, demonstrated through an enhancement approach.

In another line of research, the cTBS approach was used to explore dlPFC's role in visual WM (Schicktanz et al., 2015; Vékony et al., 2018). Schicktanz et al. (2015) applied cTBS over dlPFC while participants performed the letter n-back task. The stimulation disrupted memory accuracy in the 2-back task but not in the 0-back task, reinforcing dlPFC's role in updating and manipulating information in WM. Interestingly, the 3-back task performance was not affected, possibly due to compensatory mechanisms that mitigated the disruption caused by cTBS. Vékony et al. (2018) employed the same task and found that cTBS over dlPFC impaired the practice effect (i.e., improvement) in the task, highlighting dlPFC's role in consolidating task-specific skills and overall cognitive performance.

Although these studies offer causal evidence for dlPFC's involvement in information manipulation across various domains of WM (Oliveri et al., 2001; Postle et al., 2006; Esslinger et al., 2014; Schicktanz et al., 2015; Vékony et al., 2018), they do not directly address the issue of distraction mitigation in WM. However, evidence for dlPFC's causal role in filtering out irrelevant information in WM has been provided by using another approach—concurrent TMS-fMRI. This technique offers unique insights into the immediate effects of TMS on both the stimulated region and connected regions as measured by neuroimaging. In one study, researchers applied concurrent TMS-fMRI to perturb dlPFC activity during the memory delay period, when visual distractors might be presented (Feredoes et al., 2011). TMS over dlPFC increased activity in visual areas representing the memory items, but only in the presence of distractors. This suggests that distractor mitigation might be achieved by enhancing representations in sensory cortices, as indicated by increased connectivity between dlPFC and the visual cortex following stimulation (Feredoes et al., 2011).

However, the cognitive control of WM likely involves a network of brain regions orchestrating together to resolve interference in WM (D'Esposito & Postle, 2015) as supported by the collective findings from different techniques (i.e. dlPFC from lesion and TMS studies and IPS from studies using neural decoding). A key candidate for managing cognitive control in the brain is a set of regions in the parietal and frontal cortex known as the Multiple-Demand (MD) network.

This network has been proposed as breaking down complex tasks into manageable sub-tasks, enabling efficient problem-solving and goal-directed behaviour through the integration of cognitive operations across distributed brain regions (Duncan, 2010; Duncan, 2020). The regions identified as part of this network include the anterior inferior frontal sulcus (aIFS), posterior inferior frontal sulcus (pIFS), premotor cortex (PM), inferior frontal junction (IFJ), anterior insula/frontal operculum (AI/FO), intraparietal sulcus (IPS), and bilateral pre-supplementary motor area/anterior cingulate cortex (pre-SMA/ACC). These regions are simultaneously activated during a variety of demanding tasks, including those involving WM, selective attention, and problem-solving (Fedorenko et al., 2013; Assem et al., 2020). Recent decoding studies provided evidence for the idea that MD network supports cognitive flexibility by adaptively coding information that is relevant to the task at hand (Jackson & Woolgar, 2018; Jackson et al., 2017). For example, in an fMRI-MVPA study, Jackson et al., 2017 used abstract "spiky" objects as stimuli and asked participants to discriminate the length or the orientation of these objects in separate blocks keeping the display the same in each block. The decoding revealed that MD regions code relevant information, i.e. length or orientation, in the presence of irrelevant information, and adjust this code according to the current task (Jackson et al., 2017). Moreover, using a similar paradigm, Jackson & Woolgar (2018) found that the same voxels within MD network encode not only the information relevant to the current task but also information that was relevant in a previous task or will be relevant in an upcoming task. This finding suggests that MD network has a robust capacity for flexible, multi-task coding, allowing it to maintain and reuse neural representations across different tasks (Jackson & Woolgar, 2018). This characteristic of MD network makes it a good candidate to elicit flexible cognitive control over memory maintenance in the face of external interference, specifically, by enhancing the coding of the relevant information (memory representations) in the presence of irrelevant information (external interference).

Indeed, this mechanism is supported by the study exploring the causal role of dlPFC in MD system using concurrent TMS-fMRI (Jackson et al., 2021). Researchers applied TMS over dlPFC while participants performed a discrimination task in the scanner: they had to attend to the colour or the shape of the abstract objects in separate blocks. The stimulation was applied on each trial but of different intensity, high stimulation intensity/active stimulation and low intensity was a control condition. Comparison of neural decoding in MD network during trials with high versus low TMS intensity applied over dlPFC revealed the causal role of this region in modulating the coding of relevant information. Specifically, the application of TMS led to a decrease in the coding of relevant information across MD regions, highlighting the critical influence of dlPFC in maintaining and modulating task-relevant neural representations within MD network. Interestingly, the application of TMS did not produce a significant effect on the coding of irrelevant information, such as shape when participants were focused on colour. This finding suggests that the role of dlPFC is primarily in enhancing relevant information throughout MD network rather than actively suppressing irrelevant information (Jackson et al., 2021).

Altogether, research shows that interference from external distractions or interruptions disrupts WM maintenance resulting in lower memory recall accuracy or precision. The greater the cognitive effort required to process an interfering stimulus or task, the more detrimental the effect on memory maintenance, with interruptions— which demand cognitive efforts—causing more significant disruption than distractors that can be ignored. While neural decoding studies highlight the robustness of memory representations against distractions, causal evidence demonstrates the pivotal role of dlPFC in the cognitive control of WM. Furthermore, MD regions, including dlPFC, are crucial in maintaining cognitive flexibility by dynamically encoding task-relevant information, making MD network a strong candidate for supporting memory maintenance in the face of interference.

Role of individual differences in mind wandering and worry in cognitive control over WM

Internal distractions, such as internally generated thoughts, are powerful sources of interference in daily life (Smallwood & Schooler, 2015). These distractions may manifest as mind wandering or worry. Cognitive control may be the key mechanism to resolve internal distraction since individuals with a higher tendency to mind wander or worry demonstrate lower cognitive control in tasks requiring sustained cognitive effort, WM, in particular (Robison and Unsworth, 2018; Gustavson and Miyake, 2015).

Mind wandering is a mental state characterised by increased internal distraction, in which cognitive efforts shift from a task or the external environment to internal, self-generated thoughts (Smallwood & Schooler, 2015). Unsworth and Robison (2016) demonstrated that individuals who experienced more mind-wandering had lower WM (WM) capacity, as measured by the number of coloured squares participants could accurately recall. In another study, participants were asked to memorise one or two faces under varying cognitive loads and report their mind-wandering during the experiment (Krimsky et al., 2017). The researchers found that increased mind wandering decreased the accuracy of face recognition.

Researchers have identified two distinct types of mind-wandering (Seli et al., 2015a): spontaneous (unintentional), where focus involuntarily shifts from external stimuli to internal thoughts, and deliberate (intentional), where individuals consciously choose to focus on their thoughts. To assess these tendencies, Robison and Unsworth (2018) analysed the relationship between trait-level responses on questionnaires for spontaneous and deliberate mind wandering (Seli et al., 2015a) and performance on various WM tasks. The study found that only spontaneous mind-wandering was associated with lower performance across all tasks, whereas deliberate mind-wandering did not impair performance (Robison & Unsworth, 2018). This suggests that individuals who can intentionally direct their focus toward internal thoughts while maintaining control over their cognitive efforts are still able to perform well on tasks. The key factor for successful performance in WM appears to be the ability to control the allocation of cognitive resources between external and internal information, rather than a simple decrease in cognitive effort.

Additional evidence for this idea comes from research examining the relationship between anxiety and cognitive control. Anxious individuals tend to perform worse on tasks that require cognitive control compared to non-anxious individuals (see meta-analysis by Shi et al., 2019). Attentional Control Theory (ACT) posits that anxious individuals allocate more cognitive resources to processing worrisome thoughts, which diminishes the resources available for cognitive control and reduces their ability to focus on task-relevant information (Eysenck et al., 2007). Worry, defined as uncontrollable and excessive thoughts about uncertain future events, is often heightened in anxiety (Borkovec et al., 1998). Healthy individuals with increased worry have been shown to perform worse on WM tasks (Sari et al., 2017; Gustavson & Miyake, 2015). For instance, individuals who reported more worry during an experiment correctly recalled fewer memorized orientations than those who reported fewer instances of worry (Sari et al., 2017). Additionally, Gustavson and Miyake (2015) found that while trait worry did not affect reaction times in word-span tasks, it was associated with lower performance on tasks requiring efficient WM updating. This suggests that cognitive control is impaired in individuals with higher levels of worry, which can be measured as increased susceptibility to irrelevant information during a delay period.

Taken together, the studies on individual differences in mind wandering and worry highlight the crucial role of cognitive control. These tendencies which are common across many individuals can serve as indicators of poor cognitive control in WM and, potentially, in the ability to mitigate distractors as well.

It is important to acknowledge that other factors, such as sleep deprivation (Poh et al., 2016), stress (Crosswell et al., 2020), and lack of motivation (Seli et al., 2015b), can also contribute to increased internal distraction. However, a detailed discussion of these factors falls outside the scope of this work.

Aims

In this thesis, I aim to investigate cognitive control mechanisms of interference mitigation in visual WM. I consider two types of interference: external distractors and interrupters.

To provide a more detailed understanding of how distractors impact WM and to investigate more naturalistic memory targets beyond the simple features studied previously (Rademaker et al., 2015; Lorenc et al., 2018; Rademaker et al., 2019), I first explored how precision for realistic objects is altered by external distractors presented during memory maintenance. To address this question, in Chapter 2 I investigate how different types of visual distractors impact the precision of memory recall for objects. These findings would demonstrate how continuous measure of recall can be applied to study the effect of visual distractors on object memory.

In Chapter 3, I causally explored the role of dlPFC during the distraction of WM contents. Specifically, I used TMS to perturb activity in this region when less versus more visually engaging distractors were displayed during memory maintenance. The central question was whether dlPFC TMS would result in lower memory precision for objects when more engaging distractors were present during memory maintenance i.e., when increased cognitive control would be required. Clarifying the role of dlPFC in distractor mitigation would provide valuable insights for clinical interventions for populations with deficits in WM such as attention deficit hyperactivity disorder (ADHD), schizophrenia and various types of dementia.

In Chapters 2 and 3, I further investigated whether the tendency to mind wander or worry is linked to individual differences in cognitive control during distractor mitigation in WM. Specifically, I asked if the increased tendency to mind wandering or worry results in lower memory recall precision for objects in the presence of visual distractors. These findings can have wide implications for the educational environment in which emotional states such as a state of worry and increased mind wandering may impact the quality of learning, and, thus, need careful consideration.

Another approach to individual differences in cognitive control is to measure the fluctuation of the tonic pupil response. In Chapters 2 and 3, I additionally examined the relationship between cognitive control, as reflected in the pupillary response, and the ability to mitigate external distraction in WM. Namely, I first explore the relationship between variability in tonic pupil response (as an indicator of cognitive control) and memory recall precision for objects in the presence of distractors. Second, I tested whether cognitive effort, as reflected in phasic pupil response, is deployed differently in response to two types of distracting stimuli varied in their level of engagement (dynamic and static) presented during memory maintenance. In Chapter 2, I investigate how active memory maintenance, as reflected in pupil response during the memory delay period, is affected by different types of distracting stimuli. Furthermore, in Chapter 3, I examine the effect of dlPFC TMS on the pupillary response towards distractors.

Chapter 4 aims to explore the issue of interruptions in WM. Since interrupters appear to cause a greater disruptive effect on memory recall than distractors (Clapp et al., 2010) and significantly impact daily life and work (Foroughi et al., 2014), investigating the mechanism that protects memory representations from this type of interference is critical. Here, I investigate whether MD network elicits cognitive control over memory to protect it from interference during interruption. I use neural decoding to explore how the difficulty of the interrupting task presented during memory maintenance modulates memory representations in MD regions and the visual cortex. Specifically, does MD network enhance memorandum codes to protect from increasing task demands?

This work has important implications for our understanding of the control mechanisms that resolve interference in WM, and, more broadly, support adaptive but stable goal-directed behaviour. Crucially, the findings have practical implications for clinical applications, such as improving cognitive control in conditions like ADHD, schizophrenia, and dementia through targeted interventions like cognitive training and neuromodulation. Additionally, these insights can help enhance performance in everyday tasks by minimizing interruptions, managing distractions, and addressing the impact of worry on WM, particularly in educational settings.

References

Alnaes, D., Sneve, M. H., Espeseth, T., Endestad, T., van de Pavert, S. H.P., & Laeng, B. (2014). Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the locus coeruleus. Journal of Vision, 14, 1-20.

Anderson, C., & Colombo, M. (2022). Matching-to-sample: Comparative overview. Encyclopedia of Animal Cognition and Behavior, 4097-4102.

Assem, M., Blank, I. A., Mineroff, Z., Ademoğlu, A., & Fedorenko, E. (2020). Activity in the fronto-parietal multiple-demand network is robustly associated with individual differences in WM and fluid intelligence. Cortex, 131, 1-16.

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus–norepinephrine function: Adaptive gain and optimal performance. Annual Review of Neuroscience, 28, 403-450.

Baddeley, A. (1992). Working Memory: The interface between memory and cognition. Journal of cognitive neuroscience, 4 (3), 281-288.

Baldo, J. V., & Shimamura, A. P. (2000). Spatial and color WM in patients with lateral prefrontal cortex lesions. Psychobiology, 28 (2), 156-167.

Bandettini, P. A. (2009). What's new in neuroimaging methods? Annals of the New York Academy of Sciences, 1156 (1), 260-293.

Badre, D. (2008). Cognitive control, hierarchy, and the rostro–caudal organization of the frontal lobes. *Trends in cognitive sciences*, *12*(5), 193-200.

Belger, A., Puce, A., Krystal, J. H., Gore, J. C., Goldman-Rakic, P., & McCarthy, G. (1998). Dissociation of mnemonic and perceptual processes during spatial and nonspatial WM using fMRI. Human brain mapping, 6 (1), 14-32.

Borkovec, T. D., Ray, W. J., & Stober, J. (1998). Worry: A cognitive phenomenon intimately linked to affective, physiological, and interpersonal behavioral processes. Cognitive Therapy and Research, 22 (6), 561-576.

Bruning, A. L., Mallya, M. M., & Lewis-Peacock, J. A. (2023). Rumination burdens the updating of WM. Attention, Perception, & Psychophysics, 85, 1452–1460.

Burnett Heyes, S., Zokaei, N., van der Staaij, I., Bays, P. M., & Husain, M. (2012). Development of visual WM precision in childhood. Developmental science, 15 (4), 528-539.

Chao, L. L., & Knight, R. T. (1995). Human prefrontal lesions increase distractibility to irrelevant sensory inputs. Neuroreport, 6 (12), 1605-1610.

Chao, L. L., & Knight, R. T. (1998). Contribution of human prefrontal cortex to delay performance. Journal of Cognitive Neuroscience, 10 (2), 167-177.

Cirillo, G. D. P. G., Di Pino, G., Capone, F., Ranieri, F., Florio, L., Todisco, V., ... & Di Lazzaro, V. (2017). Neurobiological after-effects of non-invasive brain stimulation. Brain stimulation, 10 (1), 1-18.

Clapp, W. C., Rubens, M. T., & Gazzaley, A. (2010). Mechanisms of WM disruption by external interference. Cerebral cortex, 20 (4), 859-872.

Cocchi, L., Zalesky, A., Fornito, A., & Mattingley, J. B. (2013). Dynamic cooperation and competition between brain systems during cognitive control. Trends in cognitive sciences, 17 (10), 493-501.

Couffe, C., & Michael, G. A. (2017). Failures due to interruptions or distractions: A review and a new framework. The American journal of psychology, 130 (2), 163-181.

Cronin, M. J., Robbins, R. A., Wessling, C. M., & Braver, T. S. (2023). Phasic pupil dynamics track adaptive memory strategies in WM. Cortex, 156, 1-14.

Curtis, C. E., & Sprague, T. C. (2021). Persistent activity during WM from front to back. Frontiers in Neural Circuits, 15, 696060.

D'Esposito, M., & Postle, B. R. (2015). The cognitive neuroscience of WM. Annual review of psychology, 66 (1), 115-142.

Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. Trends in cognitive sciences, 14 (4), 172-179.

Duncan, J., Assem, M., & Shashidhara, S. (2020). Integrated intelligence from distributed brain activity. Trends in Cognitive Sciences, 24 (10), 838-852.

Dolcos, F., Miller, B., Kragel, P., Jha, A., & McCarthy, G. (2007). Regional brain differences in the effect of distraction during the delay interval of a working memory task. *Brain research*, *1152*, 171-181.

Esslinger, C., Zwipp, J., Schad, L. R., & Meyer-Lindenberg, A. (2014). Repetitive TMSto the left dorsolateral prefrontal cortex in WM tasks: effects of task load. Neuropsychologia, 53 , 66-71.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: attentional control theory. Emotion, 7 (2), 336.

Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. Proceedings of the National Academy of Sciences, 110 (41), 16616-16621.

Feredoes, E., Heinen, K., Weiskopf, N., Ruff, C., & Driver, J. (2011). Causal evidence for frontal involvement in memory target maintenance by posterior brain areas during distracter interference of visual WM. Proceedings of the National Academy of Sciences, 108 (42), 17510-17515.

Foroughi, C. K., Malihi, P., & Boehm-Davis, D. A. (2016). WM capacity and errors following interruptions. Journal of Applied Research in Memory and Cognition, 5 (4), 410-414.

Foroughi, C. K., Werner, N. E., Nelson, E. T., & Boehm-Davis, D. A. (2014). Do interruptions affect quality of work? Human factors, 56 (7), 1262-1271.

Gustavson, D. E., & Miyake, A. (2016). Trait worry is associated with difficulties in WM updating. Cognition and Emotion, 30 (7), 1289-1303.

Hallenbeck, G. E., Sprague, T. C., Rahmati, M., Sreenivasan, K. K., & Curtis, C. E. (2021). WM representations in visual cortex mediate distraction effects. Nature Communications, 12 (1), 4714.

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual WM in early visual areas. Nature, 458 (7238), 632-635.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science, 293 (5539), 2425-2430.

Huettel, S. A., Song, A. W., & McCarthy, G. (2014). Functional magnetic resonance imaging (3rd ed.). Sinauer Associates.

Jackson, J. B., & Woolgar, A. (2018). Adaptive coding in the human brain: Distinct object features are encoded by overlapping voxels in frontoparietal cortex. Cortex, 108 , 25-34.

Jackson, J. B., Feredoes, E., Rich, A. N., Lindner, M., & Woolgar, A. (2021). Concurrent neuroimaging and neurostimulation reveals a causal role for dlPFC in coding of task-relevant information. Communications Biology, 4 (1), 588.

Jackson, J., Rich, A. N., Williams, M. A., & Woolgar, A. (2017). Feature-selective attention in frontoparietal cortex: multivoxel codes adjust to prioritize task-relevant information. Journal of Cognitive Neuroscience, 29 (2), 310-321.

Jacob, S. N., & Nieder, A. (2014). Complementary roles for primate frontal and parietal cortex in guarding working memory from distractor stimuli. Neuron, 83(1), 226-237.

Kucewicz, M. T., Dolezal, J., Kremen, V., Berry, B. M., Miller, L. R., Magee, A. L., ... & Worrell, G. A. (2018). Pupil size reflects successful encoding and recall of memory in humans. Scientific reports, 8(1), 4949.

Kiyonaga, A., Dowd, E. W., & Egner, T. (2017). Neural representation of WM content is modulated by visual attentional demand. Journal of Cognitive Neuroscience, 29 (12), 2011-2024.

Klomjai, W., Katz, R., & Lackmy-Vallée, A. (2015). Basic principles of TMS(TMS) and repetitive TMS (rTMS). Annals of physical and rehabilitation medicine, 58 (4), 208-213.

Krimsky, M., Forster, D. E., Llabre, M. M., & Jha, A. P. (2017). The influence of time on task on mind wandering and visual WM. Cognition, 169, 84-90.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279-281.

Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. Nature, 453 (7197), 869-878.

Lorenc, E. S., Mallett, R., & Lewis-Peacock, J. A. (2021). Distraction in visual WM: Resistance is not futile. Trends in cognitive sciences, 25 (3), 228-239.

Lorenc, E. S., Sreenivasan, K. K., Nee, D. E., Vandenbroucke, A. R., & D'Esposito, M. (2018). Flexible coding of visual WM representations during distraction. Journal of Neuroscience, 38 (23), 5267-5276.

Lowe, C. J., Manocchio, F., Safati, A. B., & Hall, P. A. (2018). The effects of theta burst stimulation (TBS) targeting the prefrontal cortex on executive functioning: A systematic review and meta-analysis. Neuropsychologia, 111 , 344-359.

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of WM. Nature neuroscience, 17 (3), 347-356.

Mallett, R., & Lewis-Peacock, J. A. (2019). Working memory prioritization impacts neural recovery from distraction. Cortex, 121, 225-238.

Menon, V., & D'Esposito, M. (2022). The role of PFC networks in cognitive control and executive function. Neuropsychopharmacology, 47(1), 90-103.

Miller, E. K. (2000). The prefontral cortex and cognitive control. Nature reviews neuroscience, 1(1), 59-65.

Mizutani-Tiebel, Y., Tik, M., Chang, K. Y., Padberg, F., Soldini, A., Wilkinson, Z., ... & Keeser, D. (2022). Concurrent TMS-fMRI: technical challenges, developments, and overview of previous studies. Frontiers in Psychiatry, 13, 825205.

Naber, M., & Nakayama, K. (2013). Pupil responses to high-level image content. Journal of Vision, 13 (6), 7.

Nemes, V. A., Parry, N. R., Whitaker, D., & McKeefry, D. J. (2012). The retention and disruption of color information in human short-term visual memory. *Journal of Vision*, *12*(1), 26-26.

Neuroimaging. (2023). In APA Dictionary of Psychology. American Psychological Association. Retrieved from https://dictionary.apa.org/neuroimaging

Nichols, E. A., Kao, Y. C., Verfaellie, M., & Gabrieli, J. D. (2006). WM and long-term memory for faces: Evidence from fMRI and global amnesia for involvement of the medial temporal lobes. Hippocampus, 16 (7), 604-616.

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. Trends in Cognitive Sciences, 10 (9), 424-430.

Oliveri, M., Turriziani, P., Carlesimo, G. A., Koch, G., Tomaiuolo, F., Panella, M., & Caltagirone, C. (2001). Parieto-frontal interactions in visual-object and visual-spatial WM: evidence from transcranial magnetic stimulation. Cerebral Cortex, 11 , 606–618.

Pascual-Leone, A., Walsh, V., & Rothwell, J. (2000). TMSin cognitive neuroscience–virtual lesion, chronometry, and functional connectivity. Current opinion in neurobiology, 10 (2), 232-237.

Parthasarathy, A., Herikstad, R., Bong, J. H., Medina, F. S., Libedinsky, C., & Yen, S. C. (2017). Mixed selectivity morphs population codes in prefrontal cortex. Nature neuroscience, 20(12), 1770-1779.

Parthasarathy, A., Tang, C., Herikstad, R., Cheong, L. F., Yen, S. C., & Libedinsky, C. (2019). Time-invariant working memory representations in the presence of code-morphing in the lateral prefrontal cortex. Nature communications, 10(1), 4995.

Postle, B. R., Ferrarelli, F., Hamidi, M., Feredoes, E., Massimini, M., Peterson, M., Alexander, A., & Tononi, G. (2006). Repetitive TMSdissociates WM manipulation from retention functions in the prefrontal, but not posterior parietal, cortex. Journal of Cognitive Neuroscience, 18, 1712–1722. doi:10.1162/jocn.2006.18.10.1712.

Rademaker, R. L., Chunharas, C., & Serences, J. T. (2019). Coexisting representations of sensory and mnemonic information in human visual cortex. Nature neuroscience, 22 (8), 1336-1344.

Rademaker et al. (2015). The impact of interference on short-term memory for visual orientation. Journal of Experimental Psychology: Human Perception and Performance, 41 , 1650–1665.

Robinson, A. K., Quek, G. L., & Carlson, T. A. (2023). Visual representations: Insights from neural decoding. Annual Review of Vision Science, 9 (1), 313-335.

Ruff, C. C., Driver, J., & Bestmann, S. (2009). Combining TMS and fMRI: from 'virtual lesions' to functional-network accounts of cognition. Cortex, 45 (9), 1043-1049.

Sari, B. A., Koster, E. H., & Derakshan, N. (2017). The effects of active worrying on working memory capacity. Cognition and Emotion, 31(5), 995-1003.

Sakai, K., Rowe, J. B., & Passingham, R. E. (2002). Active maintenance in prefrontal area 46 creates distractor-resistant memory. *Nature neuroscience*, *5*(5), 479-484.

Seli, P., Carriere, J. S., & Smilek, D. (2015a). Not all mind wandering is created equal: Dissociating deliberate from spontaneous mind wandering. Psychological Research, 79, 750-758.

Seli, P., Cheyne, J. A., Xu, M., Purdon, C., & Smilek, D. (2015b). Motivation, intentionality, and mind wandering: Implications for assessments of task-unrelated thought. Journal of Experimental Psychology: Learning, Memory, and Cognition, 41 (5), 1417.

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.

Shi, R., Sharpe, L., & Abbott, M. (2019). A meta-analysis of the relationship between anxiety and attentional control. Clinical psychology review, 72, 101754.

Suzuki, M., & Gottlieb, J. (2013). Distinct neural mechanisms of distractor suppression in the frontal and parietal lobe. *Nature neuroscience*, *16*(1), 98-104.

Smallwood, J., & Schooler, J. W. (2015). The science of mind wandering: Empirically navigating the stream of consciousness. Annual review of psychology, 66 (1), 487-518.

Teng, C., & Kravitz, D. J. (2019). Visual working memory directly alters perception. Nature Human Behaviour, 3(8), 827-836.

Unsworth, N., & Robison, M. K. (2016). The influence of lapses of attention on WM capacity. Memory & Cognition, 44, 188-196.

Usher, M., Cohen, J. D., Servan-Schreiber, D., Rajkowski, J., & Aston-Jones, G. (1999). The role of locus coeruleus in the regulation of cognitive performance. Science, 283, 549-554.

Van der Wel, P., & Van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. Psychonomic bulletin & review, 25, 2005-2015.

Voytek, B., & Knight, R. T. (2010). Prefrontal cortex and basal ganglia contributions to visual WM. Proceedings of the National Academy of Sciences, 107 (42), 18167-18172.

Walsh, V., & Cowey, A. (2000). Transcranial Magnetic Stimulation and cognitive neuroscience. Nature Reviews Neuroscience, 1 (1), 73-80.

Watkins, E. R. (2008). Constructive and unconstructive repetitive thought. Psychological Bulletin, 134 (2), 163.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual WM. Nature, 453 (7192), 233-235.

Zokaei, N., Board, A. G., Manohar, S. G., & Nobre, A. C. (2019). Modulation of the pupillary response by the content of visual WM. Proceedings of the National Academy of Sciences, 116 (45), 22802-22810.

Chapter 2. Investigation of Visual Distraction Effects on Memory Precision for Naturalistic Objects

Abstract

Irrelevant visual information can distract us and impair our performance in everyday life, making it essential to investigate how working memory (WM) mitigates these distractions. While recent WM studies use continuous memory recall measures to assess the impact of distractors on precision, they typically focus on simple features like colours or orientations. In contrast, real-life situations often require us to remember more complex information, such as objects.

In the present study, we asked how visual distractors displayed during memory maintenance impact memory recall precision for naturalistic objects. We generated a continuous array of exemplars of each object; participants had to memorise objects in detail and choose the correct exemplar at the recall. We used vibrant colourful shapes as distractors—static (less engaging) and dynamic (more engaging)—and hypothesized that these distractors, compared to a no-distractor (fixation cross) condition, would disrupt cognitive control over memory, leading to lower recall precision.

Pupil size can indicate the effort exerted on a stimulus. Therefore, we used pupillometry to measure responses to three types of distractors, predicting that the more effort a distractor demands, the less effort would be available for maintaining the memory. We used pupillometry to track active memory maintenance and tested whether distractors disrupted this process by comparing pupil responses before and after distractor presentation during the delay period. We assessed how individual differences in cognitive control reflected in pre-trial pupil size variability and tendencies to mind wander or worry, influenced memory recall precision in trials with and without distractors. We found no effect of distracting stimuli on object memory recall precision. Pupil response during distraction indicated cognitive effort to the distractor rather than visual processing of stimuli. Memory maintenance was reflected by a continuous increase in pupil size until recall, unaffected by the distractor. Individual differences in mind wandering or worry as well as pupil variability did not correlate with susceptibility to distraction or memory recall precision.

This study contributes to the research of mechanisms of distractor mitigation by showing that memory for objects is robust to visual distraction and encourages further exploration of the impact of distractors dissimilar to memory targets.

Introduction

To perform successfully in everyday tasks, we rely on our ability to hold information in mind referred to as working memory (WM). Yet, external sensory information (i.e. distractors) from our surroundings or our internal, self-generated thoughts can make us forget what has just been on our minds. Thus, it is critical to protect WM from different types of irrelevant information that could lead to increased chances of forgetting and higher susceptibility to memory distortions or errors.

Recent studies have used the continuous recall approach, which offers a more detailed measure of WM fidelity (Ma et al., 2014), to investigate the effects of external distractions on memory representations of simple stimuli such as colour (Teng & Krawitz, 2019; Nemes et al. 2012), orientations (Rademaker et al., 2019; Rademaker et al., 2015; Teng & Krawitz, 2019; Lorenc et al., 2018) and faces (Mallett et al., 2020). In these paradigms, the response error (i.e., the difference between the correct and recalled item) reflects recall precision, which declines when visual distractors are introduced during the delay period (Nemes et al. 2012; Rademaker et al., 2019; Rademaker et al., 2015; Teng & Krawitz, 2019; Lorenc et al., 2018; Mallett et al., 2020).

Importantly, most of these studies explore the effects of perceptual similarity of distracting information on the memorandum (Nemes et al., 2012; Teng & Krawitz, 2019; Lorenc et al., 2018; Mallett et al., 2020); and, thus, choose the distractors from the same modality as memory targets. The distractors that are similar to the memory lead to larger memory recall errors (Rademaker et al., 2019), possibly by biasing memory representations in the sensory cortex (Teng & Krawitz, 2019; Lorenc et al., 2018). However, the research on the impact of irrelevant visual information dissimilar to the memorandum received less attention in recent years since the effect of this distractor type is often modest or absent (Rademaker et al., 2019; Yoon et al., 2006; Postle et al., 2004). Yet, in real life, we are often distracted by irrelevant information that differs from the type of information we hold in mind; for example, while reading an article on the internet, a pop-up with unrelated content can distract you from the article (i.e. verbal information is disrupted by visual information and motion). Therefore, further research on the impact of dissimilar distractors on visual WM is needed to understand how we mitigate the effects of irrelevant information on our memory in daily life.

One such mechanism can potentially be cognitive control; namely, distractor mitigation could be achieved by effectively eliciting control over relevant information (memorandum) and restricting the deployment of cognitive efforts toward irrelevant information (distractors). This mechanism is supported by the strong connection between cognitive control and WM. For example, young individuals with Attention-deficit/hyperactivity disorder (ADHD) show lower WM performance compared to healthy controls as reviewed by Alderson et al., 2013; in particular, the meta-analysis revealed that children with ADHD usually get worse on tasks that involve temporary storage and manipulation of visual and spatial information (Alderson et al., 2013). In addition, poor cognitive control leads to failures to stay on a task observed in healthy individuals reflected in fluctuations in cognitive control, which in turn are related to lower visual WM performance (Unsworth & Robison, 2015, 2018).

Specifically, studies used the pupil response during the pre-trial phase as a measure of fluctuations of cognitive control to provide support for its role in memory maintenance (Unsworth & Robison, 2015, 2018). This approach is based on the evidence that the Locus Coeruleus-Norepinephrine (LC-NE) neuromodulatory system plays a critical role in the regulation of arousal and stress responses in the brain (Aston-Jones & Cohen, 2005; Usher et al., 1999). Research indicates that the Locus Coeruleus (LC) operates in two distinct firing modes: tonic and phasic (Aston-Jones & Cohen, 2005; Usher et al., 1999). The baseline diameter of the pupil is linked to the LC's tonic firing rate, representing overall engagement with a task, while task-induced pupil dilations are associated with the LC's phasic activity, which reflects the allocation of cognitive efforts to specific task stimuli (Alnaes et al., 2014). Using the tonic pupil activity as a measure of fluctuations in cognitive control, researchers demonstrated that higher pre-trial pupil variability (i.e. how stable the pupil size was) leads to a lower number of successfully recalled colours of the squares (Unsworth & Robison, 2015, 2018). Yet, it is unclear if this measure of cognitive control could also predict the precision of maintained memories.

Using the phasic pupil response as a proxy for the cognitive effort, studies demonstrated that pupil response to a stimulus at encoding could predict the success of its subsequent recall, suggesting that this phasic response reflects the depth of the stimulus processing (Cronin et al., 2023; Kucewicz et al., 2018; Miller & Unsworth, 2020; Naber et al., 2013). Hence, the depth of processing of distractors could potentially be monitored through the phasic pupil response to distractor stimuli. Additionally, phasic response during the delay period can actively track the number of maintained memory items (Unsworth & Robison, 2018; Robison & Unsworth, 2019). Interestingly, Unsworth & Robison (2018) also demonstrated that when participants were presented with distractors along with target items, the pupillary response was similar to that seen when only target items were presented, indicating effective storage of only relevant items. Zokaei et al. (2019) used the retro-cue approach and further provided evidence for the involvement of cognitive control in memory maintenance by demonstrating that the pupil response during the delay period reflected the cued item in memory even in the absence of any visual input (Zokaei et al., 2019). Altogether, these findings indicate that pupil response during the delay period can actively track the maintenance of the memorandum; thus, by comparing the pupil size before and after a distractor presentation during memory maintenance, we could potentially evaluate the impact of distraction on visual memory contents.

Furthermore, in their study, Unsworth and Robison (2018) investigated the relationship between phasic pupil response and self-reported distraction. Namely, during the WM task participants were presented with *thought probes* asking them to self-report their current state of alertness, allowing the researchers to assess whether participants were focused on the task, mind-wandering, or distracted by external factors; when participants reported being off-task, their pupil dilated less reflecting the decrease in engagement to the task (Unsworth & Robison, 2018). Researchers suggest that the pupil response follows the distraction (internal or external): Distractions interfere with the work of the LC-NE system, and disrupt cognitive control resulting in fluctuations in arousal. Indeed, internal distractors, along with external ones, are a significant source of interference in WM in our daily lives.

Mind wandering is a mental state related to increased internal distraction by shifting thoughts away from a task or the external environment to internal, self-generated thoughts (Smallwood & Schooler, 2015). Unsworth & Robison (2016) demonstrated that individuals who experienced more mind-wandering had lower WM capacity, i.e. how many coloured squares participants could accurately recall. In another study, Researchers found that mind wandering reduced the memory accuracy of face recognition (Krimsky et al., 2017). Moreover, researchers identified two distinct types of mind-wandering based on volitional control (Seli et al., 2015): spontaneous (unintentional), where focus involuntarily shifts from the external environment to one's thoughts, and deliberate (intentional), where one consciously chooses to immerse in their thoughts. Using the trait-level questionnaires for spontaneous and deliberate mind wandering (Seli et al., 2015), researchers examined the link between these tendencies in a series of tasks requiring cognitive control, such as WM, and found that only spontaneous mind-wandering was linked to poorer performance, while deliberate mind-wandering did not impair task performance (Robison and Unsworth, 2018). This finding suggests that controlling where one directs cognitive effort is crucial for WM performance rather than the amount of cognitive effort deployed to the task per se (Robison and Unsworth, 2018). Both deliberate and spontaneous mind wanderers spent equivalent time engaging in their thoughts during the WM task, implying similar cognitive effort levels. However, the time spent mind wandering did not negatively affect memory recall for individuals who deliberately chose to mind wander, as they could easily refocus on the task without any impact on their overall performance. Therefore, findings from the mind-wandering research highlight the critical role of cognitive control in WM.

Another support for this idea comes from the large body of work exploring the relationship between anxiety and cognitive control: anxious individuals show lower performance in various tasks requiring cognitive control compared to non-anxious controls (see meta-analysis by Shi et al., 2019 for a comprehensive overview). Attentional Control Theory (ACT) explains this relationship by positing that processing of worrisome thoughts consumes the cognitive resources that would otherwise be available for cognitive control (Eysenck et al., 2007); anxiety impairs attentional control by increasing the influence of the stimulus-driven attentional system while reducing the influence of the goal-directed attentional system, making individuals more susceptible to distraction by both internal (i.e. worries) and external stimuli (i.e. distractors).

Worry can be defined as uncontrollable and excessive thoughts regarding uncertain events in the future, and it is often increased in anxiety (Borkovec et al., 1998). Healthy individuals with increased worry demonstrate impaired performance in WM tasks (Sari et al., 2017; Gustavson & Miyake, 2015). Individuals who worried more during the experiment correctly recalled fewer memorised orientations than those who reported fewer instances of active worry (Sari et al., 2017). Gustavson and Miyake (2015) demonstrated that while trait worry—a personality trait characterized by a persistent tendency to worry—did not affect reaction times in word-span tasks, it was associated with lower performance on tasks requiring efficient WM updating. This suggests that cognitive control is diminished in individuals with higher levels of worry. Hence, increased worry could also lead to increased susceptibility to irrelevant information presented during the delay period.

In the current study, we asked how distractors presented during the delay period impact WM precision for naturalistic objects. We tested the precision of recall for morphed images of tools and animals (Cusack & Stojanoski, 2014; Veldsman et al., 2017; Stojanoski et al., 2019), offering a more detailed measure of object memory fidelity using realistic memory targets, rather than simple features like orientations or colours commonly used in WM research. We compared memory precision for objects between trials with vibrant static or dynamic shapes served as distractors presented during the delay period, and unfilled delay (i.e. fixation cross display). We hypothesized that more visually engaging distractors would reduce memory recall precision: fixation cross would not interfere with memory, while static images with vibrant colours would shift cognitive efforts from maintaining memorandum and lead to lower recall accuracy. Dynamic distractors, which involve movement, were expected to cause the greatest decrease in recall precision. We specifically tested the impact of distractors that differ from memory targets, as recent research has primarily explored the memory-distractor relationship within the same feature space, and dissimilar distractions are more ubiquitous in everyday life (e.g. a pet running around can distract you from a thought you were going to write in a document).

There is a strong connection between individual differences in the tendency to mind-wander or worry and WM, which may be attributed to reduced cognitive control in individuals prone to spontaneous mind-wandering (Robison & Unsworth, 2018; Unsworth & Robison, 2018) and worry (Eysenck et al., 2007). We propose that cognitive control is a likely mechanism for distractor mitigation in WM. Specifically, this mechanism could be achieved by eliciting cognitive control over the memorandum in the face of external distractors. Individuals who are prone to mind wander or worry show difficulties in cognitive control during WM tasks (Robison & Unsworth, 2018; Unsworth & Robison, 2018; Eysenck et al., 2007) because external distractors disrupt cognitive control, thus impacting memory maintenance. We hypothesised that the increased tendency to mind wander or worry would be related to a larger distractor effect, i.e. lower memory recall precision for objects.

Since tonic pupil size may also reflect fluctuations in cognitive control in WM (Unsworth and Robison, 2015, 2018), we applied this approach to investigate the relationship between cognitive control and object recall precision in the present study. Following pupillometry evidence for tracking the ongoing processing of visual stimuli (Nieuwenhuis et al., 2011; Galeano-Keiner et al., 2023; Koevoet et al., 2023), in this study, we compared pupil responses to three distractor types (fixation, static, dynamic) to examine if phasic pupil response will reflect visual processing of cognitive efforts allocated for a stimulus. To examine pupil dynamics during distractor presentation, we analysed distinct processing stages as indicated by the previous studies: initial dilation (~220 ms) and subsequent constriction (~700-1000 ms), which reflect the depth of stimulus processing (Nieuwenhuis et al., 2011; Koevoet et al., 2023). Greater dilation and constriction would indicate deeper distractor processing. We also identified the recovery phase (1000-2500 ms), reflecting the reallocation of cognitive resources, and the maintenance stage, where the pupil remains stable, indicating active memory retention (Beatty, 1982; Robison & Unsworth, 2019; Strauch et al., 2022). Furthermore, we investigated the impact of distraction on memory maintenance by comparing pupil size before and after the distractor presentation. We predicted that the fixation distractor would not affect memory maintenance, leading to no change in pupil size. However, we expected static and dynamic distractors to interfere with cognitive control over memorandum reducing memory quality and pupil size. Alternatively, the distractor might be stored alongside the memory target, increasing pupil size post-distraction, as shown by Unsworth & Robison (2019).

Methods

Participants

Thirty-five participants (33 females, mean age=21.36, age range: 18-45) were recruited for the study. Participants reported normal colour vision and had normal or corrected-to-normal visual acuity. The study was approved by the research ethics boards of School of Psychology and Clinical Language Sciences and University Research and Ethics Committee of the University of Reading. Participants gave informed consent prior to all study procedures and were reimbursed for their time.

Stimuli

Memory targets

The stimulus set for the memory targets consisted of colour photos of real-life objects (animals and tools). To construct a circular space for each object, we used the diffeomorphic transformation method (Stojanoski & Cusack, 2014). This method generates scrambled images by applying smooth, continuous, and reversible spatial transformations that distort the original image while preserving essential perceptual properties, such as spatial frequency and figure-ground organization. These transformations effectively remove recognizable content, ensuring the scrambled image retains the fundamental visual characteristics of the original (Stojanoski & Cusack, 2014). This algorithm was used to create a parametric sequence of 15 distortions of each object transitioning the image from fully recognizable to maximally distorted (but still recognisable), with the midpoint representing the peak distortion. The continuum was closed by smoothly mapping the maximally distorted state back to the original recognizable image, ensuring seamless continuity and maintaining a consistent magnitude of transformation at each step.

Importantly, all morphed images in the present study were recognisable; previous neuroimaging studies that used these types of stimuli have shown that semantic information improves memory recall (Veldsman et al., 2017). Namely, researchers found that recognisable morphs could be recalled better than the exemplars of the same object, but morphed to be unrecognisable, i.e. with a larger transformation degree (Veldsman et al., 2017).

On each trial, one randomly chosen exemplar was displayed on the screen for memorising. At the memory recall, the display of all 15 exemplars of that object appeared as a wheel (i.e. placed in an invisible circle) representing the continuous measure of memory recall precision, analogical to the recall in paradigms testing the short-term memory for orientation (Rademaker et al., 2015) or colour (Teng et al., 2019; Nemes et al. 2012).

Visual Distractors

The main aim of distracting stimuli was to shift the allocation of cognitive resources from memory maintenance during the delay period. Thus, we have chosen stimuli that are attractive to the eye - a series of abstract, geometric designs with vibrant, multi-coloured gradients. There were two types of distractors – static and dynamic. The static distractor set included images of abstract shapes designed to engage with vibrant colours. The dynamic distractors were animated versions of these shapes, displayed at 30 frames per second, aimed to engage through both colour and movement. The shapes were abstract to reduce the semantic interference effect (i.e. when one object category interacts with another resulting in interference).

Each image/clip contained nearly the entire colour spectrum to minimise interference with the colour of memory targets. The goal was to prevent any single colour from influencing the recall of maintained information, as the presence of all colours during the distractor phase eliminates specific colour bias.

Static stimulus set consisted of four spherical and four cubic images (.JPG), and dynamic stimulus sets consisted of four spherical and four cubic video clips (.MP4).

Delayed Continuous Recall Task

Each trial began with a fixation cross (0.4°×0.4° of visual angle) presented in the centre of the screen for 500 ms followed by a memory target (8°×8° visual angle) presented on the screen for 2500 ms (see Figure 1). Participants were asked to memorise the memory target displayed on the screen and retain it in memory until the recall wheel appeared.

The delay period consisted of a fixation cross (0.4°×0.4°) presented for 750, 1000 or 1250 ms (delay 1), followed by a distractor stimulus presented for 5000 ms and followed by another fixation cross (0.4°×0.4°) presented for 0.75, 1 or 1.25 seconds (delay 2).

Distractor stimulus could be dynamic, static or a fixation cross; the trials with each distractor type were counterbalanced.

In the dynamic distractor condition, a video clip (14.58°× 14.58°) of spherical or cubic abstract colourful shapes (4 of each type) appeared in the centre of the screen. All video clips were fixed at 30 frames per second. In the static distractor condition, a static image (14.58°× 14.58°) of the same shapes as in the dynamic condition (4 of each type) was displayed as a distractor stimulus. In the fixation distractor condition, a fixation cross (0.4°×0.4°) was displayed in the centre of the screen. This type of distractor served as a control. Note, that because of the frame refreshing during the experiment presentation, the fixation did flick once when the fixation from the delay 1 was replaced by the fixation cross during the distractor phase.

Following the delay, a recall wheel (with a radius of 11.19°) appeared on the screen with 15 morphed exemplars of the same object. The wheel was randomly rotated on each trial to prevent participants from memorising the positions of any exemplars. To respond, participants used their right hand to click on the memorised image with the left mouse button and then pressed the space key with their left hand to submit their response. There was no time limit, and participants could press the space bar to submit a 'no response'; no response was selected by default. After pressing the space bar, a gray screen was displayed for 500ms (inter-trial interval, ITI) before the next trial began. All stimuli were presented at the centre of the screen on a gray background.

The experiment consisted of 144 trials of the delayed recall task divided into three blocks (46 trials each). The experiment lasted about 65 minutes in total.

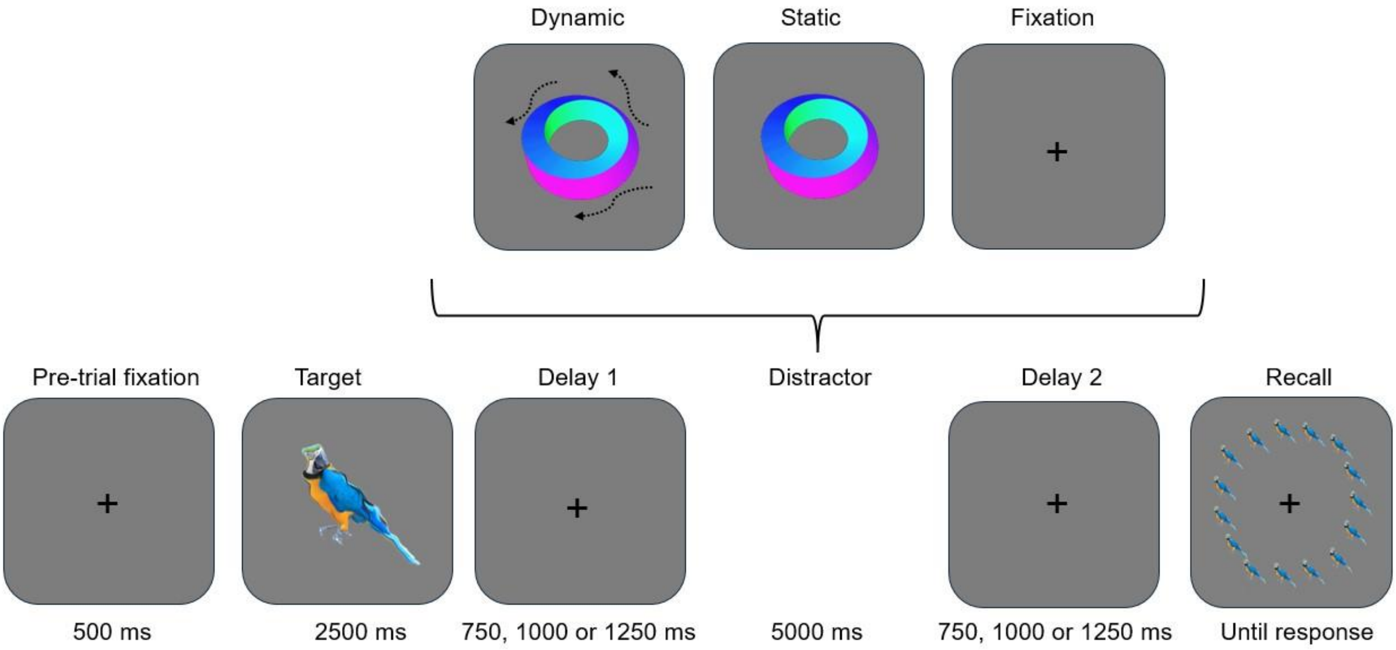


Figure 1. Experimental paradigm. In each trial, participants were first asked to focus on the fixation cross, then memorise the image presented and keep it in mind for the duration of the delay period. Then after a jittered delay, one of three types of distractors appeared: a dynamic clip of vibrant abstract shapes (spheric or cubic), a static image of similar shapes or a fixation cross. After another jittered delay, a recall wheel with fifteen morphed exemplars of the probed object appeared on the screen asking participants to choose the image they memorised with the mouse and to press the space key to send their response. Pupil size was tracked throughout the experiment. Pre-trial fixation phase was used for calculating tonic (baseline) pupil response. Pupil size during delay 1 and delay 2 were used to compare the memory maintenance before and after the distraction. Pupil response to distractors was calculated for the duration of the distractor phase.

Eye-tracking

Eye data was recorded from twenty-five participants. Eye movement and pupil size were monitored with an EyeLink 1000 tracker (SR Research, Ottawa, Canada) with the eye data sampled at 1000 Hz. The eye tracker camera was placed underneath the presentation monitor, 85 cm from the participant. The participant's head was maintained fixed on a chinrest throughout the experiment. Nine-point calibration and validation were performed before each of the experiment blocks (48 trials each) of eye-data recording. Eye data was only collected from one eye. Eye movement data was not analysed as part of this study since all stimuli were presented centrally and participants were instructed to maintain central fixation.

Procedure

Each participant underwent a short training (3-4 runs of 12 trials each) of the Delayed Recall Task before the main experiment to achieve 60% accuracy. During training, accuracy was measured by the number of correct or adjacent responses to the current item. (i.e. errors 0°, -24° or 24° were calculated as 'correct' but only in the practice runs). The main experiment consisted of 3 blocks (48 trials each) of the Delayed Recall Task.

After the memory experiment, participants were asked to complete three questionnaires testing their tendency to mind wander (Carriere et al., 2013; Vannucci et al., 2020) spontaneously (MW-S), deliberately (MW-D) and their trait worry using the Penn State Worry Questionnaire, PSWQ (Meyer et al., 1990). The order of the questionnaires was intermixed for each participant to reduce any interference of responses on one questionnaire over another.

Custom-written Matlab script (The MathWorks Inc., 2021) using Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997; Kleiner et al, 2007) were used to present the task, collect participants' responses, control the eye-tracker and collect the eye-movement data.

Behavioural data analysis

Memory recall precision was defined as the angular difference between the selected and target stimuli. To assess the distraction mitigation ability, we analysed mean absolute recall error and mean reaction times (RTs) across three distractor conditions using a three-way Analysis of Variance (ANOVA) with the main factor as distractor type and with the error term accounting for within-participant variability.

The asymptotic two-sample Kolmogorov-Smirnov test was used to compare the raw error distributions between each pair of conditions: dynamic and static, dynamic and fixation, and static and fixation.

One participant was excluded from the error analysis because of 12% of no responses (17 out of 144). Thirty-two participants were included in the reaction time analysis since another two participants were excluded as outliers following the Interquartile Range analysis: statistical dispersion, representing the range within which the central 50% of the data in a dataset lies was calculated as the difference between the first quartile (Q1) and the third quartile (Q3), where Q1 marks the 25th percentile and Q3 marks the 75th percentile of the data.

In addition, we examined how the level of object morphing impacted memory recall accuracy and RTs by plotting averaged values for each of the fifteen transformation stages and compared these values across three distraction conditions using a (non-parametric test) paired version of the Wilcoxon signed rank test with continuity correction.

Relation of Mind-wandering and Trait Worry questionnaires' scores and precision for objects with and without distraction

We used the Pearson correlation to assess the relationship between the questionnaires' scores testing the tendency to mind wander spontaneously (MW-S) or deliberately (MW-D) and worry (PSWQ) with the ability to mitigate distractors in working memory. Two measures of this ability were calculated using mean absolute error (MAE) as follows: dynamic condition measure = MAE(dynamic) - MAE(fixation); static condition measure = MAE(static) - MAE(fixation). We then correlated these differences with the scores in the questionnaires.

In addition, we asked if general performance on the task (i.e. memory recall precision) was related to tendencies to mind wander or worry. For this purpose, we calculated the mean recall error and mean reaction times across all distractor conditions, and correlated them with the scores of the questionnaires in question.

To test if distractors impacted the speed of response in individuals prone to mind wander or worry, we correlated the scores in the questionnaires (MW-S, MW-D, PSWQ) with mean reaction times from the trials where the error was zero, i.e. the correct item was chosen at the recall.

Furthermore, to assess the interaction between tendencies to worry and to mind-wander, both deliberately and spontaneously, we conducted pairwise correlations of the questionnaire scores (MW-S, MW-D, PSWQ).

Pearson correlation was chosen for the analysis since each variable followed the normal distribution, as tested by the Shapiro-Wilk test.

All statistical analyses and plotting were performed using R Statistical Software (v4.3.2; R Core Team 2023).

Pupillometry

The pupil size was recorded throughout the experiment. All pupil data is reported in the units of measure by the EyeLink systems (arbitrary units). All pupil size analysis was performed in R Statistical Software (v4.3.2; R Core Team 2023). Original .edf files were converted to .ascii format and analysed using a custom-written script in R using the 'eyelinker' package.

To preprocess the pupil data, we first removed all missing data. One subject had 30% of missing data points and thus was excluded from further analysis (another participant had 20% of data removed and the rest of the participants had less than 16% of data removed).

To evaluate the fluctuation of cognitive effort throughout the experiment, we calculated the pre-trial variability (at fixation) reflecting the tonic (baseline) pupil response as in previous work (Robison & Unsworth, 2019); for each participant, the standard deviation of the mean pupil size across trials was divided by the mean pupil size at pre-trial fixation. Then, a linear regression ('lm' function) was used to test the relationship between the pre-trial pupil variability and mean memory recall precision. Distractor type was included as a regressor in a separate model.

We further focused on the analysis of pupil dynamics during the delay period. Specifically, we asked two questions. First, to investigate the effect of distraction on memory maintenance reflected in pupil response during the delay period, we asked if the pupil size changed after the distraction (second delay) compared to the period before the distractor presentation (first delay). For this analysis, we have chosen 500 ms of the first delay period before the distractor phase (*pre-distractor period*) and 500 ms of the second delay period after the distractor was displayed on the screen (*post-distractor period*).

For this analysis, we re-sampled raw pupil size every 50 ms from each trial resulting in 10 bins for the pre- and post-distractor phases each. In addition, we re-sampled pupil response at the pre-trial fixation to calculate the baseline pupil activity for each trial. The baseline was measured as the mean of 10 pre-trial (fixation) phase bins and then subtracted from each bin of the pre-and post-distractor phases (i.e. baseline-corrected). To ensure the quality of the pupil data, the trials with more than 3 bins (150 ms) with missing data in either the pre- or post-distractor phase were excluded from the analysis; 76.3% of trials were included in this analysis. All data points per trial were paired. To fill in the missing data points in the trials with less than 3 bins with missing data, the spline interpolation ('spline' function) was applied ensuring a good fit for non-linear data.

To check if the pupil size differed between the pre- or post-distractor phases in each distractor condition, we used a (non-parametric test) paired version of the Wilcoxon signed rank test with continuity correction. The effect size (r) of the Wilcoxon test was calculated by dividing the Z-value (obtained by the Wilcox test) by the square root of the total number of observations (N).

We also were interested if the difference (or absence of it) between pre- or post-distractor phases varies across distractor conditions. To test that, we used Analysis of Variance (ANOVA) of the median pupil sizes calculated for each condition and participant, with the distractor condition as a main factor.

Another question we asked was if different distractors evoked the same or different pupil responses. We resampled the pupil data every 50 ms during the distractor phase resulting in 100 bins. The exclusion criterion for trials in this analysis was more than 6 bins (300 ms) with missing data. Participants who had more than one-third of their trials rejected (out of a total of 48 trials) were excluded from the analysis. As a result, thirteen participants were included in the final analysis. To address the missing data points in trials with fewer than 6 bins of missing data, spline interpolation (using the 'spline' function) was applied. This method ensures a good fit for non-linear data, effectively filling in the missing data points. The final data set was slightly unbalanced across distractor conditions: the aggregated dataset contained 2356 data points (i.e. one data point corresponded to a bin of a trial of a participant) for dynamic, 2332- for static and 2328- for fixation trials.

To look more closely at the pupil dynamic during the delay presentation, we divided the distractor phase into the processing stages: initial dilation (0-250 ms), subsequent constriction (250-1000 ms), recovery (1000-2500 ms) and maintenance (2500-5000 ms).

Median values were calculated for the binned pupil sizes for each distractor processing stage and distractor condition. ANOVA was used to test the effect of distractor type of each processing stage reflected by the pupil (median value). The t-test for pairwise comparisons we used a paired t-test for each pair of conditions.

Results

Effect of distractor on recall precision for realistic objects

To examine the impact of distractors on memory precision, we analysed mean absolute recall error and reaction times across three distractor conditions. Analysis revealed that distractors did not affect the memory recall error ($F(2, 66) = 0.287$, $p = .751$), as well as reaction times ($F(2,62) = 0.811$, $p = .44$) as illustrated in the top panel of Figure 2 and 3, respectively.

Error distributions (see Figure 2, bottom panel) also did not differ between the distractor conditions: dynamic and static ($D = .0183$, p-value = 0.947), dynamic and fixation(D = 0.016, p-value = 0.983), and static and fixation(D = 0.022, p-value = 0.815).

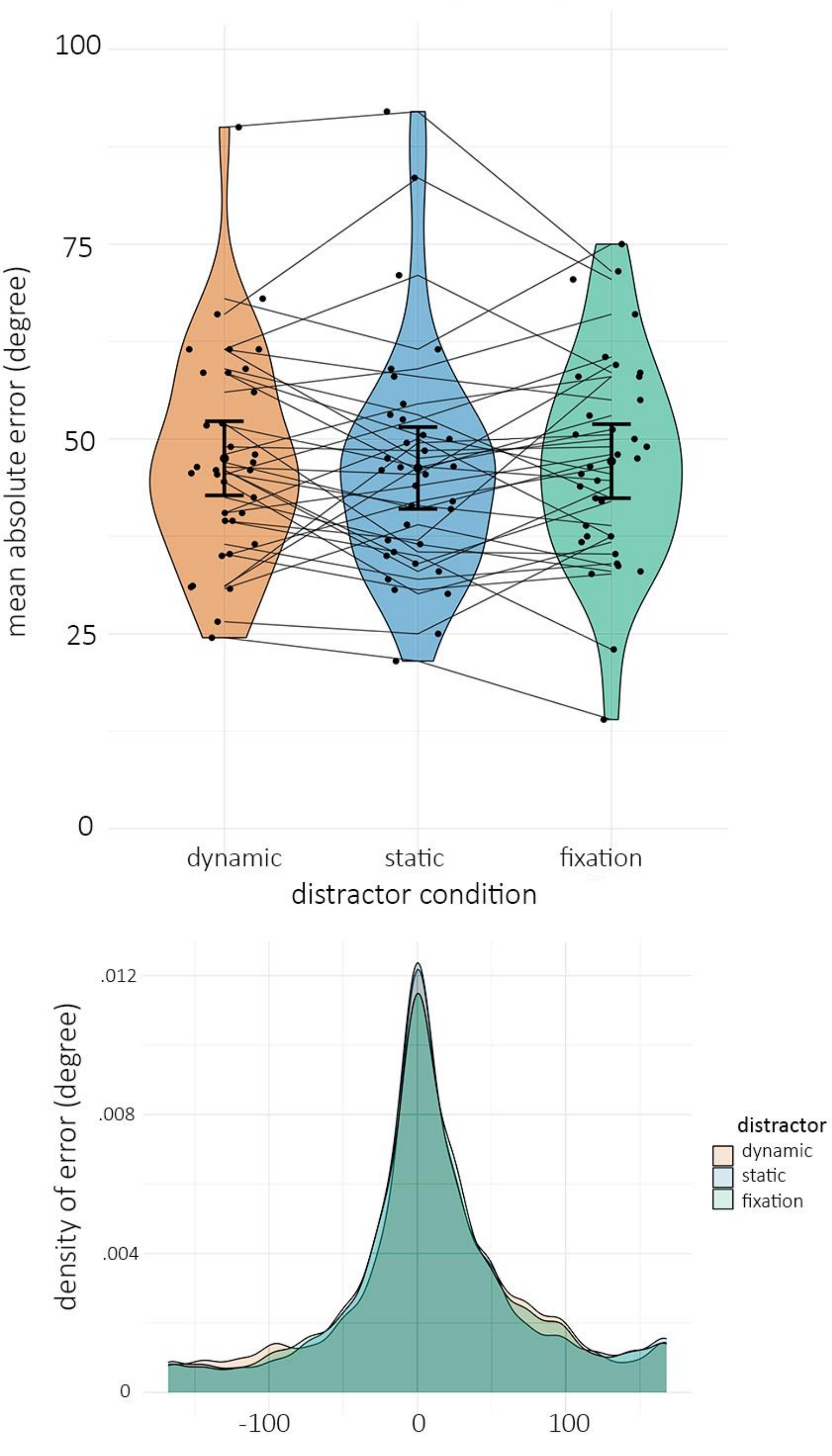


Figure 2. Precision of memory recall for naturalistic objects reflected in recall error, N=34. (Top) The mean absolute recall errors (reverse to precision) in the three distractor conditions

were the same. Bars indicate the standard error of the mean. Lines connect the mean responses from each participant. (Bottom) The raw error distribution in the three distractor conditions did not differ.

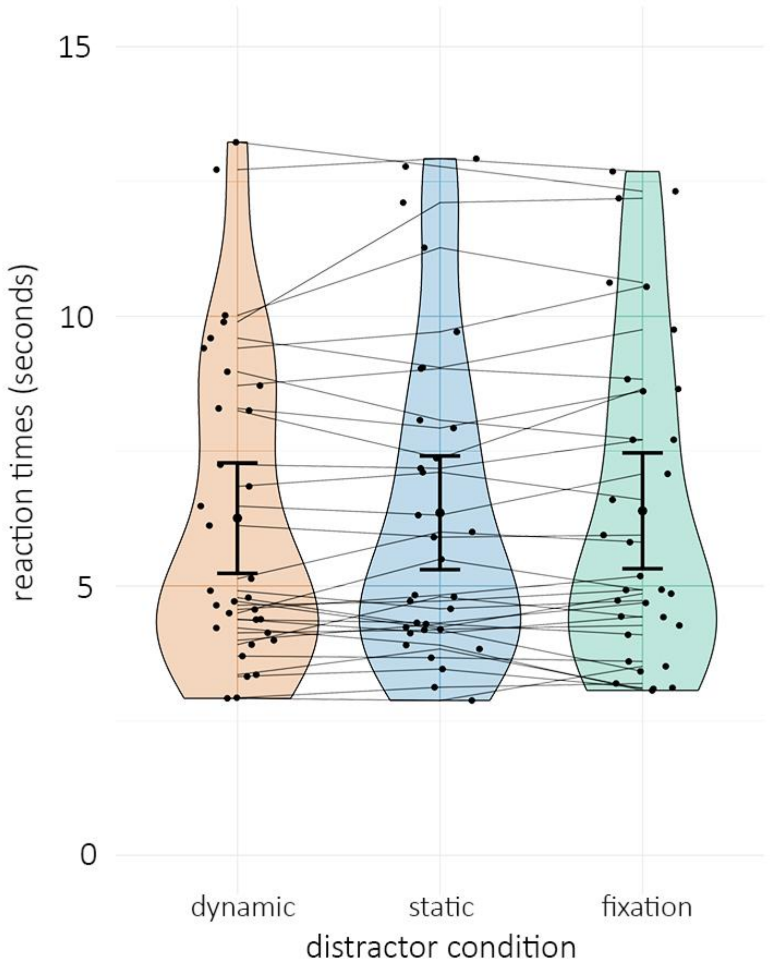


Figure 3. Reaction times in three distractor conditions (N=33). The response time did not differ between the conditions. Bars indicate the standard error of the mean. Lines connect the mean responses from each participant.

Did the level of morphing impact recall performance or susceptibility to distraction?

To examine how object recall was influenced by morph levels and whether this effect interacted with distraction, we analysed the average absolute error and RTs across all transformation steps within each distractor condition. As shown in Figure 4A, images with lower transformation levels (stages 1, 2, and 15) were recalled more accurately, resulting in lower recall errors. However, there was no clear trend of increasing error as the transformation level approached the highest distortion at stage 8 (since the transformation space was designed in a circular manner, where maximum distortion occurs at stage 8, after which the image gradually transforms back to its original image).

However, the same response pattern was observed across all distraction conditions. For absolute error, the medians were dynamic (52.23°), static (49.66°), and fixation (45.21°). Statistical comparisons revealed no significant differences between distractor conditions: dynamic vs. static (V=60, p=1), dynamic vs. fixation (V=83, p=0.2078), and static vs. fixation (V=56, p=0.8469).

As shown in Figure 4B, the analysis of RTs revealed no evidence of a dependency between recall performance and the transformation stage. For RTs, the medians were dynamic (6.17 s.), static (6.32 s.), and fixation (6.48 s.) and, similarly, comparisons showed no significant differences: static vs. fixation (V=67, p=0.719), dynamic vs. fixation (V=33, p=0.1354), and dynamic vs. static (V=42, p=0.33).

Altogether, despite some differences in memory recall for different transformation stages, these differences were irrelevant for the current analysis that focused on the effect of distractors on object memory recall and thus were not analysed further.

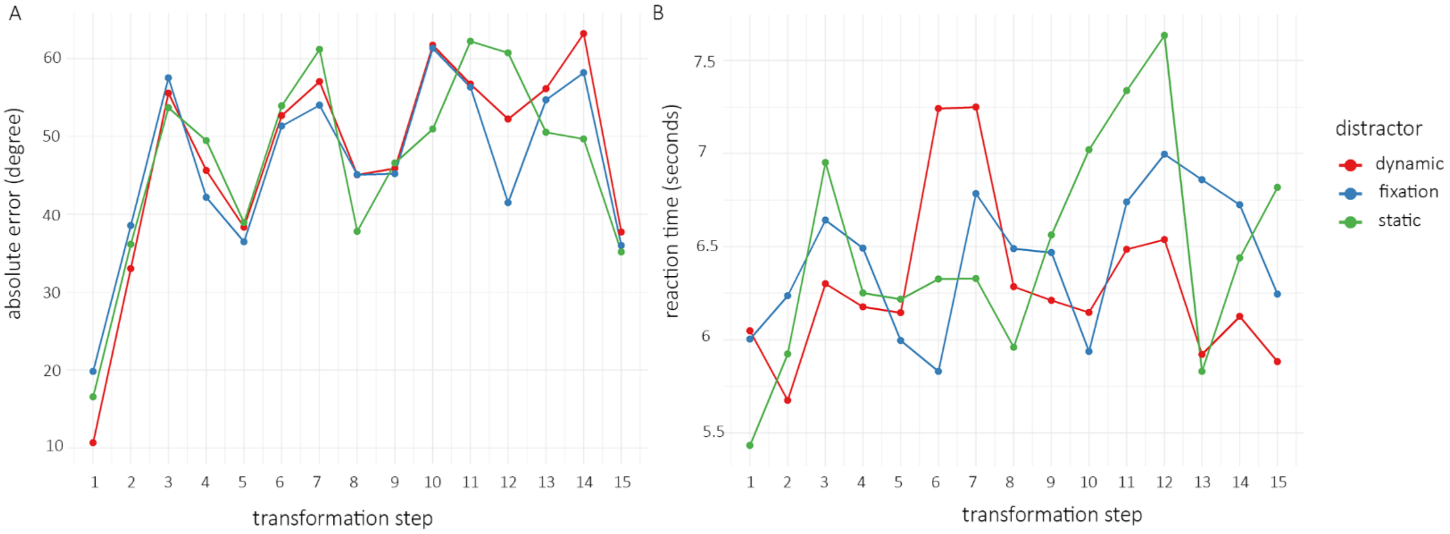


Figure 4. Average absolute errors (A) and RTs (B) at each object transformation step. Although, on average, the images more similar to the intact object – at steps 1,2 and 15- were recalled with lower absolute error (A), this effect was found irrespective of distraction. In addition, there was no such effect on RTs (B). Note: Transformation steps denote the steps for morphing an original image (step 1) to the maximum distortion (step 8) and back to the original image; thus, creating the circular space of object transformation.

Pupil variability did not predict the memory recall precision for objects

The results showed that fixation variability was not a significant predictor of the square root of the absolute error (see Table 1). The overall model was not statistically significant, F(1,73)=2.33, p=0.132. A model with a distractor condition regressor included did not reveal any relationships (see Table 2), and was statistically nonsignificant as well (F(3,61)=0.882, p=0.456).

Table 1. Regression on mean absolute error by pupil variability

Predictor	Estimate (β)	Std. Error	t-value	p-value
Intercept	6.165	0.549	11.230	< 0.001 ***
Fixation Variability	6.443	4.216	1.528	0.132

Table 2. Regression on mean absolute error by pupil variability and distractor conditions

Predictor	Estimate (β)	Std. Error	t-value	p-value
Intercept	6.118	0.599	10.217	< 0.001 ***
Fixation Variability	6.243	4.320	1.445	0.154
Distractor Type (fixation)	0.182	0.316	0.577	0.566
Distractor Type (static)	0.035	0.317	0.111	0.912

Phasic pupil response was not modulated by the type of distracting stimuli

All three types of distractions evoked the same initial dilation (F(2, 24)=0.1, p= .905) as illustrated in Figure 5. This result is surprising since we expected the initial pupil response to reflect the visual quality of a distracting stimulus which varied across the conditions: fixation cross is a less visually attractive stimulus than vibrant colour in the case of static distractors, or colour with movement in the case of dynamic distractors.

Subsequent constriction was also the same across distractor conditions (F(2,24)=0.77,p= .473), suggesting a similar depth of processing of all three types of distracting stimuli.

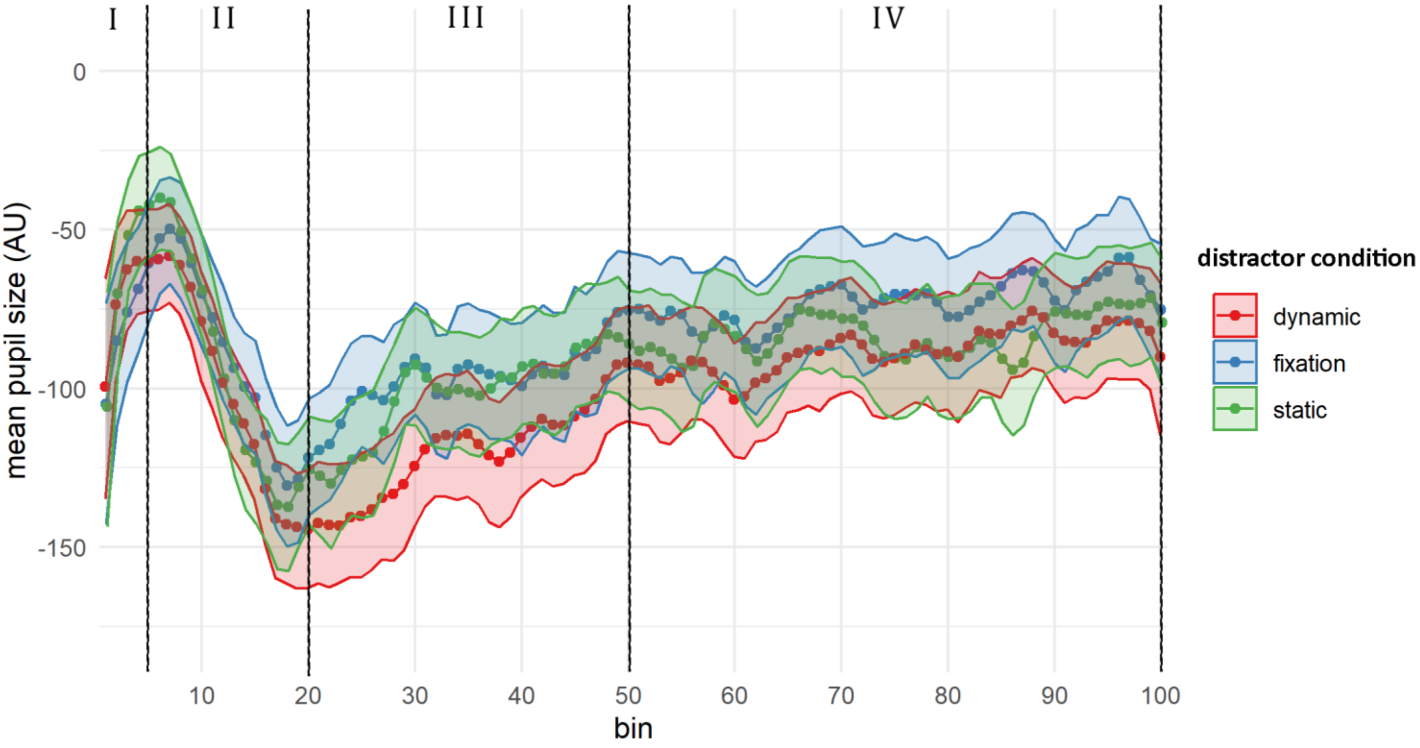


Figure 5. Phasic pupil response during the distractor phase (N=13). Pupil size is resampled to 50 ms bins. Roman numerals denote the stages of processing: I is the initial dilation, II – subsequent constriction, III – recovery, and IV – maintenance. All three types of distractors elicited the same phasic response at all stages. However, during the recovery phase, the dynamic distractors resulted in a smaller pupil size compared to the fixation cross.

Table 3. Median pupil size values for each processing stage during the distractor phase

	static	dynamic	fixation
spike	-18.78579	-13.84267	-29.12911
dip	-56.27398	-70.6435	-56.58058
recovery	-71.89898	-80.80966	-35.15708
maintenance	-65.8169	-90.72768	-43.8048

There was a slight trend towards the difference in pupil size during the recovery stage (F(2,24)=2.65, p= .0912); specifically, pupil response to the dynamic distractors was lower (or pupil size rose slower) than when fixation was displayed during the delay period (t = -2.6956, df = 12, p-value = 0.0194). No difference in pupil response at the recovery stage was found between static distractors and fixation (t = -0.578, df = 12, p = .573) or static and dynamic distractors (t = 1.528, df = 12, p = .1522).

The pupil size during the memory maintenance stage was not affected by the distractor type (F (2,24) =0.349, p=0.709).

See Table 3 for the median values for each step in each distractor condition.

Note that the analysis included only trials with less than 150 ms of missing data due to blinks or offscreen gaze suggesting that participants did attend to the distracting stimuli.

Increase in pupil dilation during memory maintenance is not impacted by distractors

To test if distractors impacted memory maintenance reflected by the pupil response, we compared pupil size 500 ms pre- and post-distractor presentation. The pupil size increased in the post-distractor period compared to the pre-distractor period when a dynamic (V = 117044, p < .0001, r= .31) or static (V = 118083, p < .0001, r=.32) distractor, or fixation (V = 120577, p < .0001, r= .31) was shown on the screen as illustrated in Figure 6. In addition, this increase was the same in all three distracting conditions(F(2,72)=0.06, partial η^2=0.001). All median pupil size values are presented in Table 4.

Table 4. Pupil size medians before (pre-) and after (post-) distractor presentation, and their differences in each distractor condition.

Condition	Pre-distractor	Post-distractor	Pre minus Post
Fixation	915.94	1021.57	-30.47
Static	910.59	1021.88	-44.29
Dynamic	918.14	1021.57	-36

These findings suggest that the pupil kept dilating during the whole delay period reflecting the active maintenance of the memorandum; importantly, the presentation of irrelevant vibrant shapes (static or dynamic) on the screen did not alter this process.

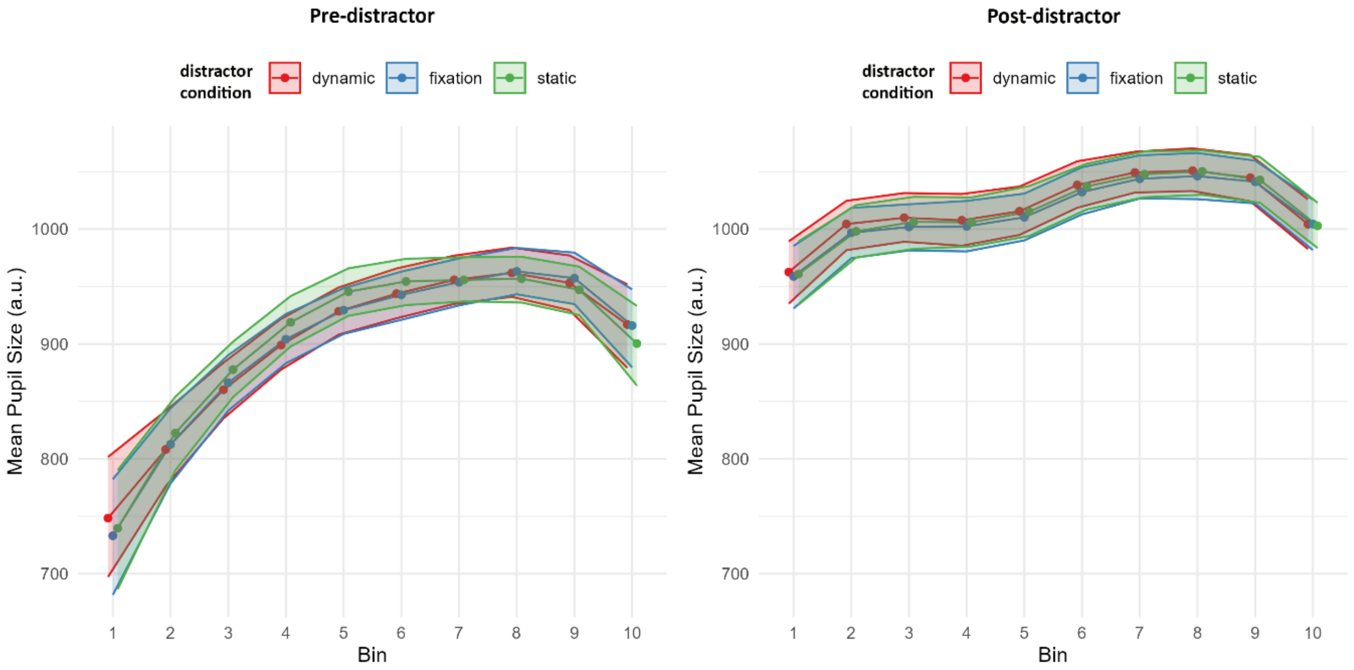


Figure 6. Pupil size dynamic pre- and post-distractor phase in three distractor conditions (N=25). Pupil response sampled to 50 ms bins, shaded plot areas correspond to the confidence intervals calculated using a bootstrapping approach (1000 samples) for mean value calculation. First, the pupil size increased during the post-distractor period compared to the pre-distractor period irrespective of distractor presence (or type). Second, as expected, there was no difference between pupil response pre-distractor phase (left) in all trials. After the distractor presentation, the pupil size dilated equally in all distractor conditions (right). Note, that the pupil size rises across both phases irrespective of the distracting stimuli.

How do individual differences in mind wandering or worry impact memory recall for objects?

Correlation analysis of differences and average mean absolute error with questionnaire scores (see Table 5) did not reveal any significant relationships (p's >.1). Only the tendency to worry and the difference in mean absolute error between trials with static versus fixation distractors revealed a trend (r(32)= .32, p = 0.0568).

Table 5. Correlation (r values) of questionnaires' scores with differences in mean absolute error, and averaged mean absolute error across conditions (MW – mind wandering, PSWQ – Penn State Worry Questionnaire).

	Dynamic – Fixation error difference	Static – Fixation Error difference	Average Error
Spontaneous MW	.14	.26	.08
Deliberate MW	.17	.11	-.04
PSWQ (Worry)	.09	.32	.30

Analysis of reaction times in correct trials (see Table 7) revealed only one statistically significant negative correlation between the average reaction times (across distractor conditions) and spontaneous mind wandering (r (30) = -.37, p=0.0326). The rest of the correlations of scores with RTs were statistically insignificant (p's>.1).

Table 6. Correlation of questionnaires' scores with differences in mean reaction time in correct trials, and averaged RTs across conditions (MW – mind wandering, PSWQ – Penn State Worry Questionnaire).

	Dynamic – Fixation RT difference	Static – Fixation RT difference	Average RT
Spontaneous MW	-.14	-.07	-.37*
Deliberate MW	-.01	-.23	.03
PSWQ (Worry)	-.03	-.21	-.09

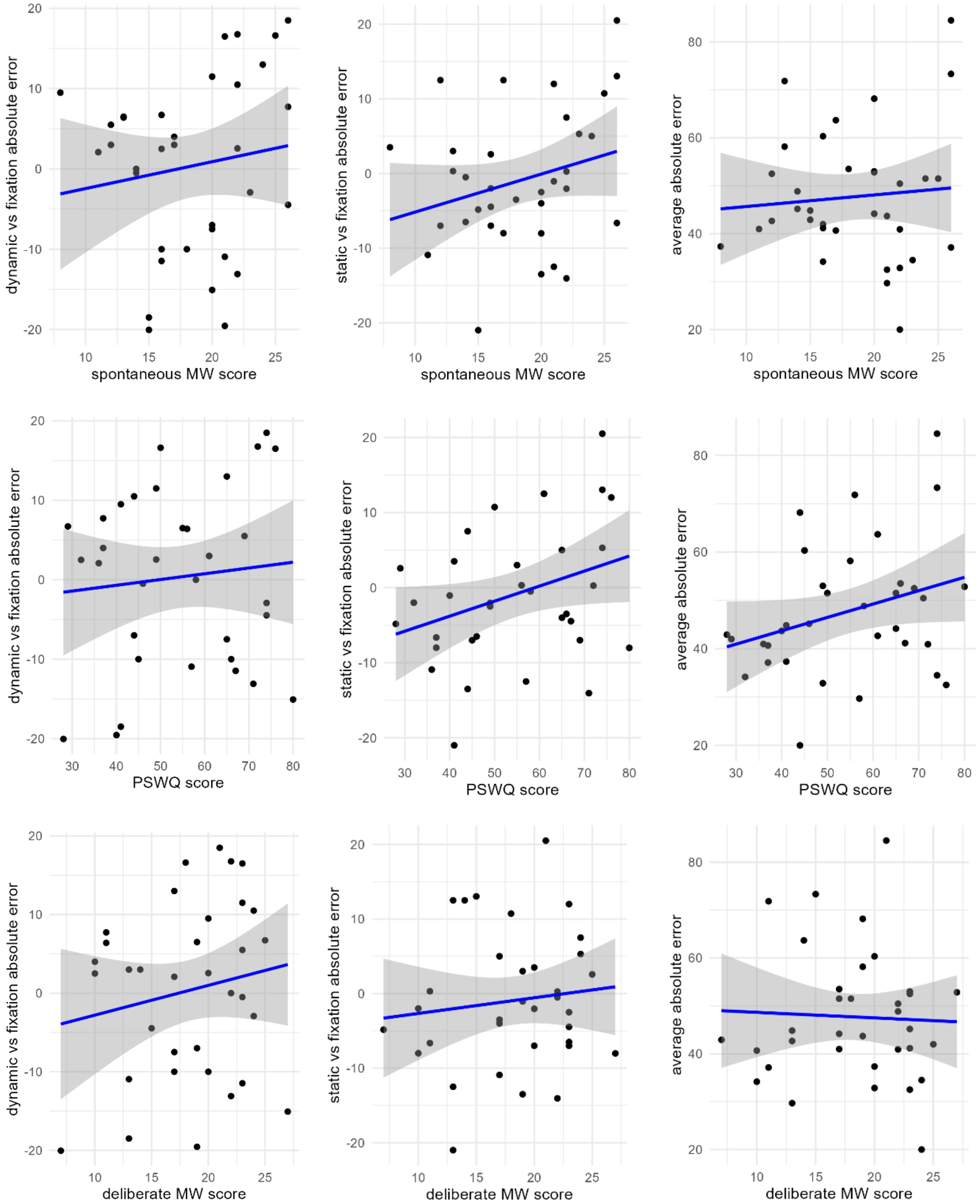


Figure 7. Correlations of tendencies to mind wander (spontaneously or deliberately) and worry (PSWQ) with absolute recall error. X-axis denotes the score on a questionnaire (see labels for details), Y-axis denotes the difference of mean absolute recall errors between conditions or averaged absolute error across all distractor conditions.

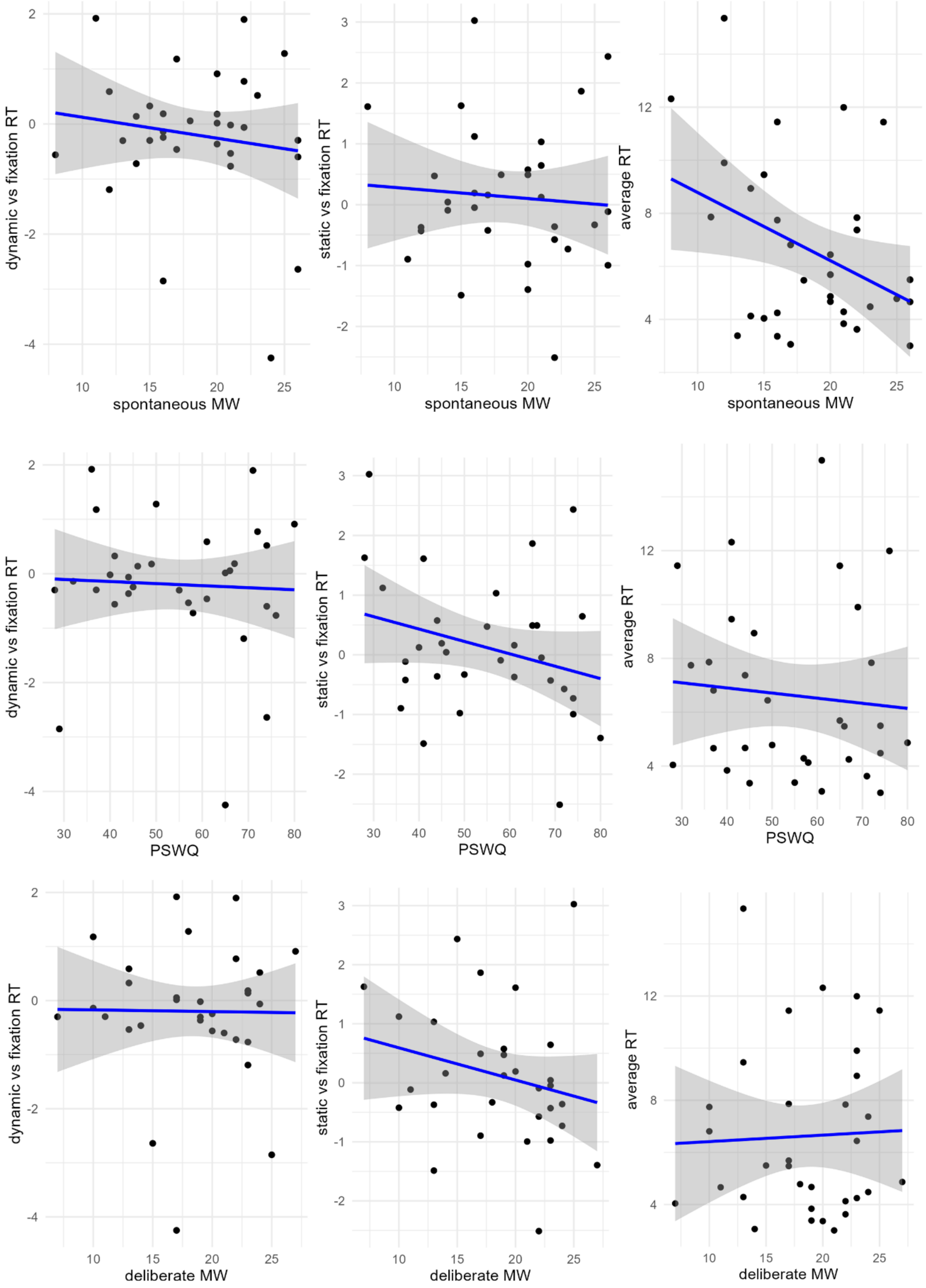


Figure 8. *Correlations of tendencies to mind wander (spontaneously or deliberately) and worry (PSWQ) with RTs.* X-axis denotes the score on a questionnaire (see labels for details),

Y-axis denotes the difference of mean RTs between conditions or averaged absolute error across all distractor conditions.

Relationships between tendencies to mind-wander and worry

In addition to the main analysis, we examined the relationship between worry and both deliberate and spontaneous mind-wandering. We also tested whether these two types of mind-wandering were correlated or represented distinct aspects of the mind-wandering state.

First, we found no significant relationship between spontaneous and deliberate mind-wandering (r(32) = .13, p = 0.5), confirming that these two questionnaires assess distinct types of mind-wandering as displayed in Figure 9. Second, the tendency to worry was not significantly associated with spontaneous mind-wandering (r(32) = .22, p = 0.26). In contrast, worry was positively correlated with deliberate mind-wandering (r(32) = .52, p = 0.0058), suggesting that individuals who tend to worry are more likely to engage in deliberate mind-wandering.

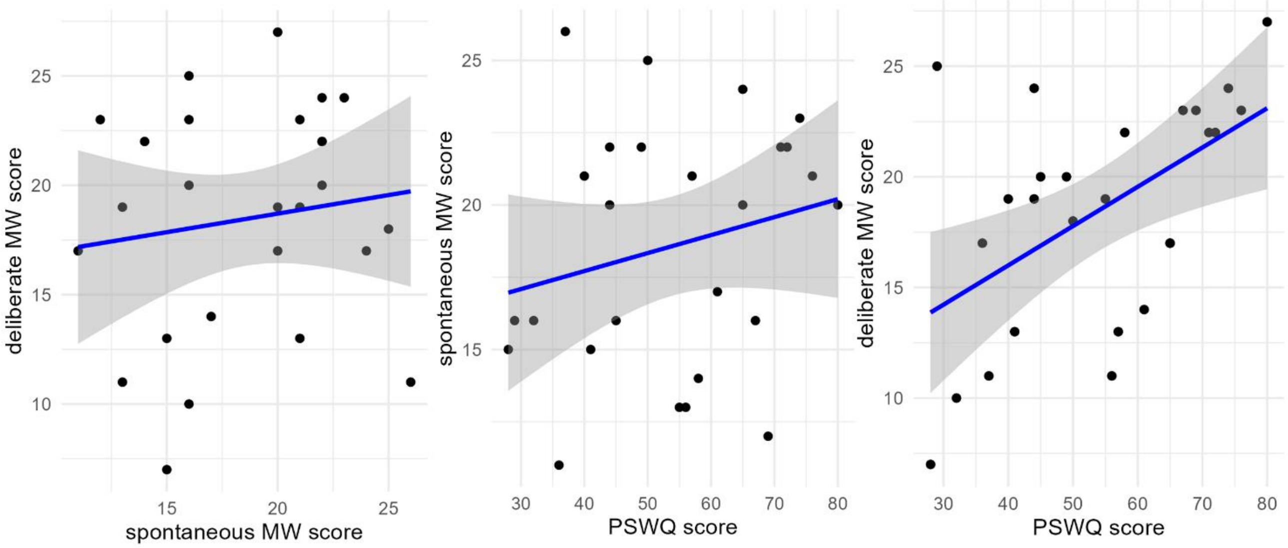


Figure 9. Correlation plots of spontaneous and deliberate mind-wandering tendencies (left panel) and their relationship with worry (PSWQ) in the middle and right panels, respectively. No relationship was found between both types of MW (left) and between worry and spontaneous MW (middle). Deliberate MW correlated positively with worry (right). MW stands for Mind-Wandering, PSWQ stands for Penn State Worry Questionnaire.

Discussion

In the present study, we investigated how irrelevant information, in the form of static or dynamic images with vibrant colourful shapes presented during the delay period, impacts working memory (WM) precision for naturalistic objects. We investigated how individual differences in mind-wandering or worrying, linked to lower cognitive control, might increase susceptibility to distractors, thereby reducing memory precision. We also used pupillometry to track cognitive effort for processing different distractor stimuli. By comparing pupil responses before and after distractor presentation, we aimed to test how different types of distractors impact memory maintenance reflected in pupil response during the delay period.

In our study using a continuous report approach, we first investigated whether the memory precision for naturalistic objects is changed by the presence of vibrant colours and dynamic movements of stimuli designed to engage participants and consume the cognitive resources deployed for memory maintenance. We did not find a detectable difference between the precision of memory recall for objects when a fixation cross (no-distractor), static or dynamic distractor was presented during the delay period. This finding suggests that memory for objects has a degree of robustness against the impact of irrelevant information presented during the maintenance phase. This aligns with previous studies that have found no significant effect of delay distractors on WM (Rademaker et al., 2019; Yoon et al., 2006; Postle et al., 2004). In their first experiment, Rademaker et al. (2019) tested the memory for orientations and presented participants with contrast-reversing visual distractors (e.g., noise distractors), which did not lead to a decrease in memory recall precision. In contrast, in Experiment 2, when more naturalistic and unpredictable distractors such as flickering images of faces and gazebos were used, there was a significant drop in behavioural performance (Rademaker et al., 2019). Yet, in our study, the movement of the vibrant shapes did not lead to a distraction effect. Yoon et al. (2006) found no effect of a scene distractor on face recognition. However, they observed a reduction in recall accuracy when a face was presented during the delay period, emphasising the effect of more similar distractors (Yoon et al., 2006). Postle et al., (2004) tested spatial memory in the presence of a flickering checkerboard as a delay distractor and found no behavioural effect of such a strong visual stimulus. Importantly, the lack of a distractor effect in these studies indicates the effective mitigation of visual distractions in working memory.

Indeed, these studies also used neuroimaging to demonstrate the neural mechanism of distractor mitigation. Rademaker et al. (2019) used evidence from neural decoding to argue that memory robustness to distractors is achieved by changing memory codes in sensory areas; memory representations change their format to mnemonic format, less prone to interference with external sensory information. Postle et al. (2004) proposed that memories are replayed in the sensory areas during the delay period as evidenced by the elevated activity in the sensory regions during the delay period that corresponded to the memorised location. Thus, active rehearsal mitigated the effect of the flickering checkerboard presented during the delay period. Yoon et al. (2006) found that activity in the prefrontal cortex (PFC) was reduced when distractors impacted behavioural performance suggesting that this region is important for maintaining memorandum in sensory areas in the face of distraction. Altogether, neuroimaging provides important insights into the mechanisms of distractor mitigation, especially when the behavioural effect of distraction is not observed.

Yet, apart from participants' ability to effectively mitigate distraction, several factors may have contributed to the absence of the distractor effect in the present study. First, we used recognisable naturalistic objects as memory targets carry semantic information which enriches the memorandum. Brady et al. (2022) demonstrated that WM capacity for real-life objects is higher than for simple features like coloured squares or shapes. Researchers suggest that real-world objects carry more detailed and conceptually rich information compared to simple stimuli; this additional information produces a more distributed, robust representation, permitting the reliance on multiple codes to maintain information over time in working memory (Postle & Hamidi, 2007; Brady et al., 2022). Veldsman et al. (2017) provided further support for this idea by using neural decoding to compare the memorandum of recognisable versus non-recognisable morphed objects. The study found that recognisable objects were recalled with higher precision and produced more varied neural patterns across trials compared to unrecognisable objects (Veldsman et al., 2017). This indicates that recognisable objects might be supported by a richer and more diverse set of neural representations, even though these differences do not result in stronger or additional brain region activity (Veldsman et al., 2017). Therefore, it is possible that the naturalistic memory targets in our experiment could be encoded with a richer and more diverse set of neural representations, making them more resistant to interference.

Another factor could be the expectation of distraction. A recent study demonstrated that when participants could predict the onset of interference (distractor or interrupter) during the memory delay period, i.e. when interference onset was fixed compared to the variable onset, the impact of this interference was reduced (Gresch et al., 2021). Although we slightly jittered the distractor presentation, participants knew that irrelevant information might appear during the delay period (and they were aware the experiment was testing distraction), so they could engage a strategy to avoid delay period distractors.

One way to overcome these constraints and explore the role of distractors dissimilar to memory targets (similar to the type of distraction we encounter in real life) is to study distractor mitigation in WM in more naturalistic environments. For example, virtual reality (VR) may provide a medium to study distraction as we face it in real life. Stokes et al., (2022) used a VR classroom setting with eye-tracking to study how children with Attention-Deficit Hyperactivity Disorder (ADHD) perform tasks such as math, Stroop, and continuous performance test (CPT), with various classroom-like distractions introduced (e.g. a pedestrian walks by the window, a student avatar sneezes, a phone rings on a student’s desk). The presence of distractors significantly disrupted on-task eye-gaze and lowered task performance, as children with ADHD had difficulty returning to tasks once interrupted, leading to sustained disengagement even when the distractors did not shift their focus for long periods (Stokes et al., 2022). Another study investigated how visual distractions affect various cognitive processes including encoding, visual search, WM usage, and decision-making while performing a task in a VR environment: participants were asked to copy a model display by selecting objects from a resource pool, i.e. an arrangement of 24 cubes (Kumle et al., 2024). Eight of these cubes displayed images of objects that matched the targets in the model display and the remaining16 cubes in the resource pool were distractor objects that varied in distractibility: distracting cubes could be more transparent, i.e. easy to distinguish from targets, or they could have the same opacity as the target objects (Kumle et al., 2024). Researchers found that the difficulty in managing distractions resulted in reduced reliance on working memory, ultimately slowing overall task performance and requiring more physical effort to complete the object-copying task (Kumle et al., 2024). Notably, cognitive tests performed in VR environments with immersive, colourful backgrounds and the use of 3D depth provide similar results to the simple cognitive tasks done on a computer (Redlinger et al., 2022), and, hence, could be used for valid evaluation of WM and distraction mitigation in future studies.

Pupillometry is an effective method for measuring the processing effort directed toward a stimulus (Nieuwenhuis et al., 2011; Galeano-Keiner et al., 2023; Koevoet et al., 2023). It can reveal whether participants attended to a distracting stimulus and reflect the engagement level elicited by different types of distractors. In this study, we tested this idea by comparing pupil responses during the presentation of vibrant static or dynamic distractors, as well as a fixation cross on the screen. Our analysis focused on four processing stages: initial dilation, subsequent constriction, recovery, and maintenance. Despite the distinct visual features of the distractors (such as colour and movement, compared to the fixation cross), all three types elicited similar pupil responses, except for the recovery phase. Specifically, the recovery phase showed a lower pupil size for dynamic distractors than for the fixation cross.

Provided that the distractors used in the study did not lead to lower memory precision for objects, it is unsurprising that the pupil responses they elicited were similar. It is also important to note that the fixation distractor flickered at the start of the distractor phase, evoking a pupil response similar to the other two types of distractors. This finding aligns with evidence suggesting that pupil size does not simply reflect the processing of a visual stimulus but rather the cognitive effort allocated to it (Kahneman & Beatty, 1966).

The observed difference in the recovery phase between dynamic and fixation conditions may suggest that the recovery process involves greater cognitive effort when returning to the baseline maintenance level after engaging with a dynamic stimulus, compared to the relatively lower effort required for processing a fixation cross. This implies that the dynamic stimulus demands more cognitive resources, making it more challenging for the system to revert to its resting state, as opposed to the simpler task of refocusing after attending to a static fixation point. That response error was not different between distractor difficulty levels adds further to the argument that in our experimental paradigm, participants were able to effectively deal with visual distractors, regardless of their different visual features.

We also used pupillometry to investigate the impact of distracting stimuli on the memorandum, building on recent studies that demonstrated that pupil size tracks the active maintenance of working memory (Unsworth & Robison, 2018; Zokaei et al., 2019). Specifically, comparing pupil size before and after the distractor phase revealed that memory quality remained unaffected by both dynamic and static distracting stimuli, as indicated by a consistent pupil size difference across all distractor conditions. This finding aligns with the lack of observable behavioural effects of distraction in the present study. Interestingly, pupil size increased after the distractor phase regardless of the type of distracting input, possibly indicating the growing cognitive effort required to maintain working memory over time.

Low cognitive control can manifest as higher fluctuations of cognitive control over time that could be measured by pupil variability before trial onset (Unsworth & Robison, 2015, 2018). To examine the relationship between cognitive control and memory fidelity in our paradigm, we applied this approach and found that pupil variability did not predict the precision of object memory recall, irrespective of distractor condition. One possible explanation could be that the measure of memory performance used in the present study—memory precision for objects—differs from the memory capacity measure (K) used in previous work by Unsworth & Robison (2015, 2018). Therefore, while pupil variability might be related to the number of items one can hold in memory (i.e., K), it may not necessarily relate to how precisely those items are remembered. Since the distractors did not impair memory in our experiment, it is possible that participants did not need to exert a high level of cognitive control to maintain the memoranda. Consequently, the fluctuations in cognitive control, as indicated by pupil variability, may not have been crucial for the successful performance of our task.

To investigate further the role of cognitive control, we examined if individual differences in the tendency to mind wander or worry can be attributed to the ability to mitigate distraction in WM. Previous studies found that individuals prone to spontaneous (but not deliberate) mind-wandering (Robison & Unsworth, 2018) or worry (Eysenck et al., 2007) perform worse in WM tasks due to the reduced control: external distractors disrupt the attentive state and internal distraction such as internally generated thoughts or worries consume resources from the task at hand. Since in the present study, we did not find the effect of distraction, it is unsurprising that individual differences in mind wandering or worry were not correlated with memory precision difference in trials with dynamic or static distractors compared to the fixation distractor condition. There was a trend for a possible positive correlation between trait worry and the difference in memory errors between static and fixation distractors, suggesting that static distractors could induce worry and lead to larger errors (i.e. lower memory recall precision). In addition, general memory recall precision (averaged across conditions) was also unrelated to tendencies to mind wander or worry. This result may relate to differences in memory measures between our study and previous ones. Whereas earlier studies focused on the success or speed of memory recall, we examined how the precision or quality of WM is associated with tendencies to mind wander or worry. While mind wandering or worry may affect recall efficiency, they do not seem to degrade the accuracy or quality of information in working memory. This interpretation points to a nuanced understanding of cognitive performance, where different aspects of memory—speed, accuracy, and precision—can be differentially affected by factors such as mind wandering or worry.

We also analysed average reaction times (RTs) and differences in RTs between distractor conditions in correct trials and found no correlations between tendencies to mind wander or worry except for the negative relationship between spontaneous mind wandering and average reaction time. Individuals who reported being more prone to spontaneous mind wandering recalled memory targets more quickly than those who reported less mind wandering. Although this finding is counterintuitive, there may be several explanations. First, individuals whose mind wander spontaneously have lower cognitive control and could potentially rely on a more automatic, less deliberate retrieval process without overthinking at the recall and responding faster. Second, distractors appeared in two-thirds of the trials and may have helped keep individuals, whose minds tend to wander spontaneously, engaged with the task. In other words, an increased load of the task caused by distractors could potentially consume cognitive efforts leaving less for internally generated thoughts. Indeed, some participants reported that they enjoyed watching the colourful shapes during the delay period and found them helpful in keeping memory targets in mind.

Furthermore, we examined the relationship between worry and spontaneous and deliberate mind-wandering. We found a positive link between tendencies to worry and to mind wander deliberately. Deliberate mind wandering occurs when individuals consciously choose to engage in off-task thinking, often directed toward specific goals, such as problem-solving or planning (Smallwood & Schooler, 2015). Individuals prone to worry may intentionally engage in deliberate mind wandering as a coping strategy to work through their worrisome thoughts (Baird et al., 2011). However, if the worry remains unresolved, this process can lead to rumination, ultimately intensifying the worry (Seli et al., 2015). The present finding may help explain why we did not observe a relationship between distractibility or performance in the working memory task (i.e., memory precision and RT) and the tendency to worry. Individuals prone to worry likely possess sufficient cognitive control to intentionally engage in mind wandering (off-task thinking), suggesting they also may have sufficient cognitive control over the task itself.

To provide further evidence for this link, future studies could incorporate a more detailed questionnaire on mind wandering, including a broader range of questions that specifically assess the nature of deliberate mind wandering; namely, questions should aim to determine whether mind wandering is used intentionally to resolve worrisome thoughts. Moreover, the current assessment relied on self-report questionnaires, which are subject to bias (Podsakoff et al., 2003). Therefore, to examine more closely the interactions between mind wandering and worry in WM and distractor mitigation, future studies may use Experience Sampling Methods (ESM), prompting individuals to report their thoughts at random intervals, providing real-time measurement of active worrying (Sari et al., 2017) or state of mind-wandering (Unsworth & Robison, 2015) during the WM task performance. This approach may help to assess the link between the current internal state of an individual and their immediate responses to the task at hand; namely, future studies may ask whether increased active worry leads to more spontaneous or deliberate mind-wandering and whether it results in lower performance or higher susceptibility for distraction in WM.

In addition, future studies on individual differences in mind wandering or worry in WM protection must address their interaction using paradigms in which distractors disrupt behaviour. It is also essential to expand our knowledge of these relationships by exploring various types of memory targets, and measurements such as continuous recall or assessments within realistic environments such as VR.

To conclude, this study provides further evidence for memory robustness to distraction by demonstrating that memory for naturalistic objects is not impacted by unrelated vibrant static or dynamic stimuli presented during maintenance. Pupil response during distraction presentation reflected the cognitive effort and not visual processing per se. Memory maintenance was reflected by the pupil response which kept increasing until the recall phase and was not changed by the distractor presentation. Individual differences in cognitive control as reflected in tendencies to mind wander or worry were unrelated to the susceptibility to external distraction or overall memory recall precision.

References

Alderson, R. M., Kasper, L. J., Hudec, K. L., & Patros, C. H. (2013). Attention-deficit/hyperactivity disorder (ADHD) and working memory in adults: a meta-analytic review. Neuropsychology, 27(3), 287.

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus–norepinephrine function: Adaptive gain and optimal performance. Annual Review of Neuroscience, 28, 403-450.

Baird, B., Smallwood, J., & Schooler, J. W. (2011). Back to the future: Autobiographical planning and the functionality of mind-wandering. Consciousness and cognition, 20(4), 1604-1611.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. Psychological Bulletin, 91(2), 276–292.

Blom, T., Mathôt, S., Olivers, C. N. L., & Van der Stigchel, S. (2016). The pupillary light response reflects encoding, but not maintenance, in visual working memory. Journal of Experimental Psychology: Human Perception and Performance, 42(11), 1716–1723.

Borkovec, T. D., Ray, W. J., & Stober, J. (1998). Worry: A cognitive phenomenon intimately linked to affective, physiological, and interpersonal behavioral processes. Cognitive Therapy and Research, 22(6), 561-576.

Bays, P. M., Wu, E. Y., & Husain, M. (2011). Storage and binding of object features in visual working memory. Neuropsychologia, 49, 1622–1631.

Brady, T. F., & Störmer, V. S. (2022). The role of meaning in visual working memory: Real-world objects, but not simple features, benefit from deeper processing. Journal of Experimental Psychology: Learning, Memory, and Cognition, 48(7), 942.

Clapp, W. C., Rubens, M. T., & Gazzaley, A. (2010). Mechanisms of working memory disruption by external interference. Cerebral Cortex, 20(4), 859-872.

Cronin, S. L., Lipp, O. V., & Marinovic, W. (2023). Pupil dilation during encoding, but not type of auditory stimulation, predicts recognition success in face memory. Biological Psychology, 178, 108547.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. Emotion, 7, 336–353.

Galeano-Keiner, E. M., Pakzad, S., Brod, G., & Bunge, S. A. (2023). Examining the role of attentional allocation in working memory precision with pupillometry in children and adults. Journal of Experimental Child Psychology, 231, 105655.

Gustavson, D. E., & Miyake, A. (2016). Trait worry is associated with difficulties in working memory updating. Cognition and Emotion, 30(7), 1289-1303.

Gresch, D., Boettcher, S. E., Van Ede, F., & Nobre, A. C. (2021). Shielding working-memory representations from temporally predictable external interference. Cognition, 217, 104915.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. Science, 154(3756), 1583-1585.

Koevoet, D., Naber, M., Strauch, C., Somai, R. S., & Van der Stigchel, S. (2023). Differential aspects of attention predict the depth of visual working memory encoding: Evidence from pupillometry. Journal of Vision, 23(6), 9.

Krimsky, M., Forster, D. E., Llabre, M. M., & Jha, A. P. (2017). The influence of time on task on mind wandering and visual working memory. Cognition, 169, 84-90.

Kumle, L., Võ, M. L. H., Nobre, A. C., & Draschkow, D. (2024). Multifaceted consequences of visual distraction during natural behaviour. Communications Psychology, 2(1), 49.

Kucewicz, M. T., Dolezal, J., Kremen, V., Berry, B. M., Miller, L. R., Magee, A. L., ... & Worrell, G. A. (2018). Pupil size reflects successful encoding and recall of memory in humans. Scientific reports, 8 (1), 4949.

Lorenc, E. S., Pratte, M. S., Angeloni, C. F., & Tong, F. (2014). Expertise for upright faces improves the precision but not the capacity of visual working memory. Attention, Perception, & Psychophysics, 76, 1975-1984.

Mallett, R., & Lewis-Peacock, J. A. (2019). Working memory prioritization impacts neural recovery from distraction. Cortex, 121, 225-238.

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. Nature Neuroscience, 17(3), 347-356.

Miller, A. L., & Unsworth, N. (2020). Variation in attention at encoding: Insights from pupillometry and eye gaze fixations. Journal of Experimental Psychology: Learning, Memory, and Cognition, 46(12), 2277.

Naber, M., & Nakayama, K. (2013). Pupil responses to high-level image content. Journal of Vision, 13(6), 7.

Nemes, V. A., Parry, N. R., Whitaker, D., & McKeefry, D. J. (2012). The retention and disruption of color information in human short-term visual memory. Journal of Vision, 12(1), 26-26.

Nieuwenhuis, S., De Geus, E. J., & Aston-Jones, G. (2011). The anatomical and functional relationship between the P3 and autonomic components of the orienting response. Psychophysiology, 48(2), 162–175.

Rademaker, R. L., Bloem, I. M., De Weerd, P., & Sack, A. T. (2015). The impact of interference on short-term memory for visual orientation. Journal of Experimental Psychology: Human Perception and Performance, 41(6), 1650.

Rademaker, R. L., Chunharas, C., & Serences, J. T. (2019). Coexisting representations of sensory and mnemonic information in human visual cortex. Nature Neuroscience, 22(8), 1336-1344.

Redlinger, E., Glas, B., & Rong, Y. (2022). Impact of visual game-like features on cognitive performance in a virtual reality working memory task: within-subjects experiment. JMIR Serious Games, 10(2), e35295.

Robison, M. K., & Unsworth, N. (2018). Cognitive and contextual correlates of spontaneous and deliberate mind-wandering. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(1), 85.

Robison, M. K., & Unsworth, N. (2019). Pupillometry tracks fluctuations in working memory performance. Attention, Perception, & Psychophysics, 81, 407-419.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. Journal of applied psychology, 88(5), 879.

Postle, B. R., & Hamidi, M. (2007). Nonvisual codes and nonvisual brain areas support visual working memory. Cerebral cortex, 17(9), 2151-2162.

Salmela, V. R., & Saarinen, J. (2013). Detection of small orientation changes and the precision of visual working memory. Vision Research, 76, 17-24.

Sari, B. A., Koster, E. H., & Derakshan, N. (2017). The effects of active worrying on working memory capacity. Cognition and Emotion, 31(5), 995-1003.

Seli, P., Carriere, J. S., & Smilek, D. (2015). Not all mind wandering is created equal: Dissociating deliberate from spontaneous mind wandering. Psychological Research, 79, 750-758.

Starc, M., Anticevic, A., & Repovˇs, G. (2017). Fine-grained versus categorical: Pupil size differentiates between strategies for spatial working memory performance. Psychophysiology, 54(5), 724–735.

Stokes, J. D., Rizzo, A., Geng, J. J., & Schweitzer, J. B. (2022). Measuring attentional distraction in children with ADHD using virtual reality technology with eye-tracking. Frontiers in virtual reality, 3, 855895.

Strauch, C., Wang, C. A., Einhäuser, W., Van der Stigchel, S., & Naber, M. (2022). Pupillometry as an integrated readout of distinct attentional networks. Trends in Neurosciences, 45(8), 635-647.

Stojanoski, B., & Cusack, R. (2014). Time to wave good-bye to phase scrambling: Creating controlled scrambled images using diffeomorphic transformations. Journal of Vision, 14(12), 6-6.

Stojanoski, B., Emrich, S. M., & Cusack, R. (2019). Representation of semantic information in ventral areas during encoding is associated with improved visual short-term memory. BioRxiv, 2019-12.

Teng, C., & Kravitz, D. J. (2019). Visual working memory directly alters perception. Nature Human Behaviour, 3(8), 827-836.

Unsworth, N., & Robison, M. K. (2015). Individual differences in the allocation of attention to items in working memory: Evidence from pupillometry. *Psychonomic bulletin & review*, *22*, 757-765.

Unsworth, N., & Robison, M. K. (2018). Tracking working memory maintenance with pupillometry. Attention, Perception, & Psychophysics, 80, 461-484.

Usher, M., Cohen, J. D., Servan-Schreiber, D., Rajkowski, J., & Aston-Jones, G. (1999). The role of locus coeruleus in the regulation of cognitive performance. Science, 283, 549-554.

Veldsman, M., Mitchell, D. J., & Cusack, R. (2017). The neural basis of precise visual short-term memory for complex recognisable objects. NeuroImage, 159, 131-145.

Wilken, P., & Ma, W. J. (2004). A detection theory account of visual short-term memory for color. Journal of Vision, 4(8), 150-150.

Yoon, J. H., Curtis, C. E., & D'Esposito, M. (2006). Differential effects of distraction during working memory on delay-period activity in the prefrontal cortex and the visual association cortex. Neuroimage, 29(4), 1117-1126.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. Nature, 453, 233–235.

Zhou, C., Lorist, M. M., & Mathôt, S. (2022). Is categorization in visual working memory a way to reduce mental effort? A pupillometry study. Cognitive Science, 46(9), e13194.

Chapter 3. Causal Involvement of Dorsolateral Prefrontal Cortex in Distractor Mitigation During Working Memory for Objects

Abstract

The ability to mitigate distractors is a key component of successful WM (WM) performance. Causal evidence suggests dorsolateral prefrontal cortex (dlPFC) plays a key role in resolving distraction in WM by eliciting cognitive control over memoranda but the number of causal studies in healthy humans is limited, and many studies have not used sensitive enough measures of WM performance. In the present study, we test memory precision for realistic objects following offline continuous theta-burst transcranial magnetic stimulation (cTBS-TMS) over right dlPFC or primary somatosensory cortex (SI, control site). In our task, participants were presented with realistic visual stimuli to retain in WM, along with dynamic (clips of moving colorful abstract shapes) or static (a white noise image) delay period distractors. The prediction was that TMS targeting dlPFC rather than the SI would reduce recall precision for memoranda when dynamic distractors are presented compared to static distractors. We measured fluctuations in baseline pupil activity and distractor-triggered pupil response to assess changes in cognitive effort allocated to distracting stimuli due to TMS and varying distraction levels. Additionally, we administered mind-wandering and trait worry questionnaires to determine if these measures could predict individual susceptibility to distraction. TMS over dlPFC (versus SI) resulted in a lower precision of memory recall in trials with dynamic versus static distractors. The performance of individuals prone to mind-wander unintentionally (but not intentionally) benefited from visually engaging distractors suggesting that they helped these individuals to stay on task; however, this effect was disrupted in the session with dlPFC TMS. Fluctuations in baseline pupil response measured by pupillometry predicted memory recall precision irrespective of the distractor type or stimulation site. Distractor-triggered pupil response was not impacted by dlPFC TMS, but the memory maintenance-related response was decreased. These findings highlight the significance of individual differences in WM performance, and underscore the link between attentional states, physiological responses, and WM processes.

Introduction

An ongoing aspect of everyday life is the moment-to-moment need to focus on relevant visual information in working memory (WM) and mitigate the impact of irrelevant information, whether it originates from external distractions (Allen et al., 2017; Berry et al., 2009) or internal sources such as unwanted or unrelated thoughts (Unsworth & Robison, 2016; Soemer, 2019). Prior research suggests that external distractions interfere with the retention of sensory information in WM, as recently reviewed by Lorenc, et al. (2021). Specifically, the effect of this interference is reflected in the reduction in accuracy of the memory recall distractors present during the memory delay period (Clapp et al., 2010).

To provide a more detailed understanding of the effects of visual distractors, recent studies have shifted to using a continuous report method (as reviewed by Ma et al., 2014) instead of the traditional match-to-sample approach. This method has been used to examine distraction effects on simple stimuli like color and orientation (Teng & Kravitz, 2019; Nemes et al., 2012; Rademaker et al., 2015) as well as more complex stimuli like faces (Mallett & Lewis-Peacock, 2020). In these paradigms, distraction reduces memory precision, as reflected by a wider error distribution: correct responses decrease, while the number of larger errors increases.

To examine the impact of distractors on neural representations, recent neuroimaging studies combined this approach with neural decoding, showing that visual distractors during memory retention disrupt representations of orientations, leading to reduced decoding accuracy and decreased recall precision (Rademaker et al., 2019; Lorenc et al., 2018). Lorenc et al. (2018) highlight the adaptive role of the superior intraparietal cortex (sIPS), where reliable decoding of orientations occurs only in the presence of distractors, emphasizing its role in mitigating interference. Rademaker et al. (2019) found that memory representations in the IPS shift from sensory to mnemonic formats to resist interference, while those in the visual cortex remain more vulnerable to distortion. While neuroimaging studies emphasize the role of sensory and parietal cortices in mitigating distractions, lesion studies provide causal evidence that the prefrontal cortex (PFC) plays crucial role in

WM (Voytek & Knight, 2010), namely, by exerting cognitive control which can be defined as the ability to regulate and coordinate thought and behaviour in accordance with internally generated goals (Miller & Cohen, 2001). Specifically, the dorsolateral PFC (dlPFC) was found to play a key role in resolving distraction, as demonstrated by studies on patients with lesions in this region (Chao & Knight, 1995, 1998; Baldo & Shimamura, 2000) highlighting the role of cognitive control in this process.

For example, in two influential studies, Chao & Knight (1995, 1998) demonstrated that dorsolateral PFC lesions disrupt the ability to resolve distractions in auditory WM tasks, as patients exhibited lower memory recall accuracy in the presence of distractors compared to controls. Using a continuous recall, Baldo & Shimamura (2000) observed that dlPFC lesion patients had greater errors in spatial and colour WM tasks compared to healthy individuals, particularly when faced with an interfering task (i.e. monitoring digits that were appearing under fixation cross). Despite their limited number, lesion studies provide valuable evidence for the pivotal role of dlPFC in mitigating distraction in WM.

The neural stimulation studies corroborate further the causal role of dlPFC in WM by examining healthy individuals. Namely, several studies demonstrated the role of dlPFC in WM by using transcranial magnetic stimulation (TMS) (Oliveri et al., 2001; Postle et al., 2006; Esslinger et al., 2014; Schicktanz et al., 2015; Vékony et al., 2018). For example, Oliveri et al. (2001) applied single TMS pulses over dlPFC during the delay period of visual-object and visual-spatial WM tasks. dlPFC TMS reduced both response times and accuracy in both tasks, suggesting that this region plays a key role in cognitive control over various types of information in WM (Oliveri et al., 2001).

An alternative approach is repetitive transcranial magnetic stimulation (rTMS): this protocol delivers multiple TMS pulses in a sequence and produces a prolonged (i.e. lasting after stimulation is produced) modulation of brain activity and connected regions (Klomjai et al., 2015). For instance, Postle et al. (2006) asked participants to memorise an array of letters which had to be stored in memory in their order (“forward” trials reflecting retention) or retained and then sorted in the alphabetic order (“alphabetize’’ trials corresponding to manipulation) while applying the rTMS over dlPFC during the delay period; rTMS dlPFC only affected the response accuracy in the “alphabetize” trials suggesting that this region is primarily involved in manipulating information in WM rather than its short-term retention (Postle et al., 2006). These findings highlight the key role of dlPFC in maintaining and manipulating information in WM.

A variant of the repetitive TMS protocol, continuous theta burst stimulation (cTBS) TMS (Huang et al., 2005), uses gamma frequency trains at a theta rhythm and can be applied briefly (40 seconds) with effects lasting up to 50 minutes (Wischnewski & Schutter, 2015). While the mechanisms remain unclear, cTBS often produces inhibitory effects in the stimulated region. Due to its shorter duration and prolonged effect, many studies have explored the role of dlPFC in cognition (Lowe et al., 2018; Ngetich et al., 2020), but few have focused on its role in visual WM (Schicktanz et al., 2015; Vékony et al., 2018). For instance, Schicktanz et al. (2015) applied cTBS over dlPFC during a letter n-back task, finding that stimulation disrupted memory accuracy in the 2-back task but not in the 0-back task, confirming the role of dlPFC in updating and manipulating information in WM. Interestingly, performance in the 3-back task was unaffected, possibly due to compensatory mechanisms counteracting the disruption caused by cTBS. Similarly, Vékony et al. (2018) found that dlPFC cTBS impaired the practice effect in the n-back task, highlighting the role of this region in consolidating task-specific skills and overall cognitive performance. This result suggests that dlPFC not only supports the manipulation of information in WM but also plays a crucial part in long-term learning and skill improvement, further emphasizing its multifaceted contribution to cognitive control.

Although these studies provide causal evidence for the role of dlPFC in eliciting control in various aspects of WM such as manipulation and monitoring (Oliveri et al., 2001; Postle et al., 2006; Esslinger et al., 2014; Schicktanz et al., 2015; Vékony et al., 2018), they did not address the issue of distraction mitigation in WM. However, this issue was investigated by a study using a more sophisticated approach to stimulation - concurrent TMS-fMRI; this technique provides unique insights into the immediate effects of TMS on both the stimulated region and connected regions as measured by neuroimaging. Researchers applied concurrent TMS-fMRI to perturb activity in dlPFC during the memory delay period when visual distractors could be presented (Feredoes et al., 2011). TMS over dlPFC increased activity in sensory areas representing memorandum only in the presence of distractors. Thus, distractor mitigation might be achieved through enhancing representations in sensory cortices as revealed by increased connectivity between dlPFC and visual cortex as a result of stimulation (Feredoes et al., 2011). However, more causal evidence is needed to expand upon this finding and to examine the impact of visual distractors on memory precision using a more fine-grained measure of memory fidelity.

The role of cognitive control in WM can be also addressed through the lens of individual differences in mind-wandering (Krimsky et al., 2017, Unsworth & Robison, 2016). Mind wandering is a mental state characterised by increased internal distraction, in which cognitive efforts shift from a task or the external environment to internal, self-generated thoughts (Smallwood & Schooler, 2015). For instance, Unsworth and Robison (2016) found that individuals who experienced more mind-wandering had lower WM capacity, measured by their ability to recall coloured squares. Similarly, Krimsky et al. (2017) showed that increased mind-wandering reduced face recognition accuracy under varying cognitive loads.

Interestingly, the degree of volitional control over the mind-wandering state results in two dissociated types of mind-wandering (Seli et al., 2015): spontaneous (or unintentional) occurring when focus involuntarily shifts from the external environment to one's thoughts, and deliberate (or intentional) when one chooses to direct their attention internally to immerse in their thoughts. These tendencies can be assessed using trait-level questionnaires (Seli et al., 2015), which were used to measure the impact of different types of mind wandering on performance in a series of tasks on attention and WM (Robison and Unsworth, 2018). Specifically, only the tendency to mind-wander spontaneously but not deliberately was associated with lower performance (Robison and Unsworth, 2018) pointing to the importance of cognitive control over this state rather than the number of lapses in focus per se. It remains unclear if this mechanism is also involved in mitigating external distractions in WM.

Although mind wandering can be exacerbated by multiple factors such as sleep deprivation (Poh et al., 2016), stress (Crosswell et al., 2020), and lack of motivation (Seli et al., 2009), negative thoughts like worry and rumination are powerful sources of mind wandering (Robison et al., 2017), negatively impacting WM (Sari et al., 2017; Bruning et al., 2023; Gustavson & Miyake, 2015). Sari et al., (2017) demonstrated that active worrying (i.e. having uncontrollable and intrusive negative thoughts about the future) decreased the number of correctly recalled orientations of the rectangles after a memory delay highlighting the role of worry in memory capacity. Bruning et al. (2023) examined the impact of rumination—fixation on intrusive, negative thoughts—on the ability to update images in a memory task. The study found that individuals with low WM capacity and a high tendency to ruminate were less accurate in the updating task. These findings suggest that rumination may impair WM by hindering the ability to distinguish between relevant and irrelevant information (Bruning et al., 2023). Similarly, Gustavson and Miyake (2015) found that while trait worry (a personality trait involving a tendency to worry) did not affect word-span performance (reaction times), it was associated with poorer performance on tasks requiring efficient WM updating. Together, these findings suggest that negative thoughts impair the ability to prioritize task-relevant (e.g. memorandum) over task-irrelevant information (e.g. worry, rumination) in WM; however, their impact on mitigating external distractions in visual WM remains unexplored.

Pupillometry also can be used to investigate the role of cognitive control in WM since pupil dilation serves as a proxy for measuring arousal and attentional state in WM tasks (Robison & Unsworth, 2019; Unsworth & Robison, 2015). The proposed neural basis for pupillary modulation is the locus coeruleus-norepinephrine (LC-NE) neuromodulatory system (Aston-Jones & Cohen, 2005; Usher et al., 1999). Research suggests that the LC has two firing modes: tonic and phasic (Aston-Jones & Cohen, 2005; Usher et al., 1999). Baseline pupil diameter corresponds to LC tonic firing rate indicating overall task engagement and task-evoked dilations correspond to LC phasic activity reflecting attention allocation to task stimuli (Alnaes et al., 2014). Thus, fluctuations in attention can be measured by analysing the baseline pupil diameter (phasic pupillary response). Researchers asked participants to memorise coloured squares while monitoring the pupil response before the start of each trial (at pre-trial fixation) as a baseline (Robison & Unsworth, 2019; Unsworth & Robison, 2018). Findings demonstrated that smaller pre-trial pupil size and its fluctuation were related to the lower number of memory items recalled on a trial (Robison & Unsworth, 2019; Unsworth & Robison, 2018). Whereas phasic pupil response can be used as a proxy for attentional allocation to a stimulus: larger initial dilation (increase of pupil size) is related to better (memory) stimulus encoding (Griffin & Nobre, 2003). Also, the subsequent constriction (decrease of pupil size) may indicate the depth of the attentional processing of a stimulus: novel naturalistic images were recalled better when the pupil constricted more as a response to image presentation (Naber et al., 2013). Active memory maintenance can also be tracked using phasic pupillary response: Robison & Unsworth (2019) demonstrated that pupil diameter during the delay period increased with the number of items held in memory and then successfully recalled. Hence, the depth of processing of irrelevant information may potentially be tracked by the (phasic) pupil response to distractor stimuli; also, tonic pupil response during the delay period may reflect the content of maintained memorandum; while overall fluctuations in attention in WM may be monitored by the (pre-trial) tonic pupil response.

In the present study, we investigated causally the necessity of dlPFC for distractor mitigation in WM using the cTBS TMS to disrupt activity in this region since this approach has been found to produce a robust and long-lasting suppression (Wischnewski & Schutter, 2015). To examine how distraction impacts memory for real-life objects, we used a more precise measure of WM performance than a match-to-sample approach. Specifically, we used a pseudo-continuous assessment of the object memory recall (Veldsman et al., 2017, Stojanoski et al., 2019). We compared memory recall in the trials with more (dynamic distractors) versus less attentionally engaging distractors (static noise image) presented during the delay period to reveal the impact of distraction on memory precision for objects. To test the effect of dlPFC stimulation on the ability to mitigate distraction, we compared memory precision in two sessions: one with cTBS applied over the region of interest (i.e., dlPFC) and the control session with stimulation applied over the primary somatosensory cortex (SI, control site). Choosing SI as the control stimulation region is advantageous for several reasons: SI TMS does not produce immediate output (like motor areas), applies the stimulation over the cortical tissue (in comparison to vertex TMS), and is relatively easy to locate once the primary motor area (M1) is identified in the motor-threshold procedure (see Holmes et al., 2018 for the localisation guidelines).

To explore the role of cognitive control through the lens of individual differences in mind-wandering or worry, we correlated the trait worry and mind-wandering (deliberate and spontaneous) with the difference in memory precision in trials with dynamic versus static noise distractors. We also examined the relationship between tonic pupil variability measured at pre-trial fixation and memory recall precision in all distractor and stimulation conditions. To evaluate possible stimulation effects on the pupil response to distracting stimuli, we analysed the phasic pupil responses during distractor presentation in trials with static noise and dynamic distractors separately, comparing pupil responses between stimulation sessions. Additionally, to assess the impact of dlPFC TMS on the subprocesses reflected in the phasic pupil response during the distractor phase, we divided the response into four stages: initial dilation, subsequent constriction, recovery, and maintenance. The recovery and maintenance phases were identified after the visual inspection of the pupil response plots. The maintenance phase was defined following the evidence that stable phasic response during the delay period reflects active memory maintenance (Unsworth & Robison, 2018; Zokaei et al., 2019). The recovery phase was defined as a phase between subsequent constriction and maintenance, potentially demonstrating the amount of cognitive effort needed to recover after the stimulus processing back to the memory maintenance. We then compared these stages between sessions for each distractor condition.

Methods

Online immediate memory recall experiment

Participants

To ensure that the memory targets in the in-lab experiment could be memorized equally, we recruited a separate cohort of participants, and ran an online memory experiment without distractors, and compared the immediate memory recall (i.e., no delay period) performance.

23 participants were recruited via Prolific platform (https://www.prolific.com/). Participants were paid at the rate of 11 pounds/hour. The study was approved by the research ethics boards of School of Psychology and Clinical Language Sciences and University Research and Ethics Committee of the University of Reading.

Procedure

The immediate memory recall of the 10 objects (5 animals and 5 tools) was tested in the online study. 16 morphed exemplars of each object were produced using the diffeomorphic algorithm (Stojanoski & Cusack, 2014).

The experiment started with the task description followed by the attention check (4 true/false statements testing participants' attention to the task description and understanding of the task). Participants were asked to complete a short training run (8 trials) before moving to the main part of the experiment. The main part consisted of 160 trials, so each morphed exemplar of each object was presented once as a memory target.

On each trial, a fixation cross (10%×10%) was presented on the screen for half second followed by a memory target presented (33%×33%) for 3 seconds (see *Figure 1*). Right after the memory target disappeared, a recall wheel (radius =48%) was presented on the screen with sixteen morphed exemplars of the same object; the wheel was rotated each trial, so participants could not memorize the positions of any exemplars on a wheel. To respond, participants were instructed to choose the image they have just seen by clicking on it with a computer mouse or a touchpad and pressing the space key to submit their response. After response submission or after 8 seconds passed (the recall timeout), the next trial started immediately.

All stimuli were presented at the centre of the screen on gray background. The stimulus sizes are shown in percentage of the screen height of participants' individual computers; hence, the sizes were varied across the participants.

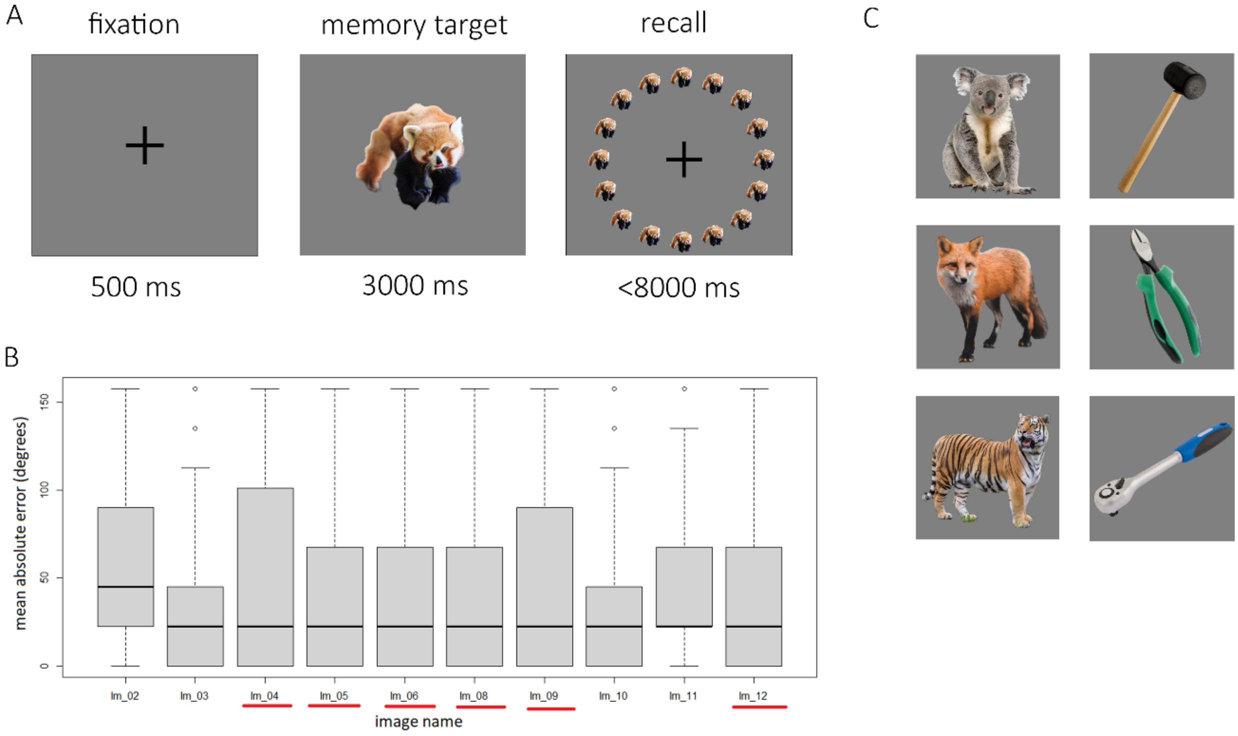


Figure 1. Online immediate memory recall experiment: paradigm (A), resulting recall precision per image tested (B), and the images chosen for the main experiment (C). (A) In each trial, participants saw a fixation cross, followed by the morphed image of an animal or a tool. Then, immediately after the image, a recall wheel appeared on the screen and participants had to choose the image they had just seen. Mean memory precision (B) for the 10 images (5 tools and 5 animals) tested in the immediate recall online experiment. We used visual inspection of the bar plots of the mean absolute error to choose images with the most similar recall precision (highlighted red) to ensure that all memory targets in the main experiment can be memorised equally. (C) Three images of (wild) animals and three images of tools were chosen for the main experiment. The image of the red panda shown in the paradigm description was used in the practice run in the main experiment.

Results

As in the delayed recall task, the memory recall precision was defined as the angular difference between the responded and correct morphed images. For each object, the mean recall precision and mean reaction time were calculated by averaging responses to all the exemplars of an object across participants.

We used the distributions of the memory precision (by assessing the boxplots displayed in *Figure 1B*) to visually pinpoint the most similarly recalled objects within each category (i.e., by choosing objects with a similar mean precision distribution, separately for the tools and animals); 3 objects per category were chosen (*Figure 1C*, also highlighted with red in B panel).

Main experiment

Participants

Twenty-seven participants (19 females, mean age 21.7, age range: 18-33) were recruited for the study. Participants were screened for TMS (Rossi et al., 2009) and MRI contraindications. The study was approved by the research ethics boards of School of Psychology and Clinical Language Sciences and University Research and Ethics Committee of the University of Reading. Participants gave informed consent prior to all study procedures and were reimbursed for their time.

Stimuli

Colour photographic images of real-life objects (animals and tools) were chosen as memory targets. The diffeomorphic algorithm (Stojanoski and Cusack, 2014) was applied to each image to create a pseudo-circular space of object transformation consisting of 16 morphed exemplars of the same object; the same amount of transformation was created at each step/for each exemplar of an object (i.e., each object exemplar had the same amount of spatial transformation in comparison to the next or previous exemplars in the continuum space). One randomly chosen exemplar was presented as a memory target on each trial. The display of all 16 exemplars of each object appeared at the memory recall wheel, similar to the continuous recall paradigms testing the short-term memory for orientation (Rademaker et al., 2015) or colour (Teng & Kravitz, 2019; Nemes et al. 2012).

To choose memory target stimuli for the in-lab experiment, we ran a separate online study (Online Immediate Memory Recall Experiment) ensuring that all the objects could be equally stored in WM i.e., that some stimuli were not easier to remember than others. To compare the memory encoding of objects, this study tested the *immediate* memory (i.e., no delay period) recall of the morphed objects (Figure 1A). As a result, 3 objects per category (animals, tools) were chosen for the in-lab experiment (Figure 1B).

Delayed Continuous Recall Task

As displayed in Figure 2, each trial began with a fixation cross (0.38°×0.38° of visual angle) presented in the centre of the screen for 500 ms followed by a memory target (7°×7° visual angle) presented on the screen for 3000 ms (Figure 2). Participants were instructed to memorize the memory target image as it appeared on the screen and hold it in mind for an upcoming memory test.

The delay period consisted of a fixation cross (0.38°×0.38°) presented for 750, 1000 or 1250 ms (randomised order across trials), followed by a distractor stimulus (14.58°× 14.58°) presented for ms and another fixation cross (0.38°×0.38°) presented for 0.75, 1 or 1.25 s (randomly jittered).

In 50% of trials, a distractor stimulus was a video clip (dynamic distractor condition). Video clips (.MP4) of moving round or angular abstract colourful shapes (3 of each type) were chosen as distracting stimuli (*dynamic* distractors). All video clips were fixed at 30 frames per second. In the in-lab experiment, an image of the white static noise served as a (non/less distracting) control stimulus (*static* distractor) shown during the delay period; it was displayed to ensure visual input without engaging exogenous attention to the same extent as the video clips. The static distractor was displayed as a distractor stimulus (static distractor condition) in the other half of the trials. Both distractor conditions were randomly intermixed throughout the experiment.

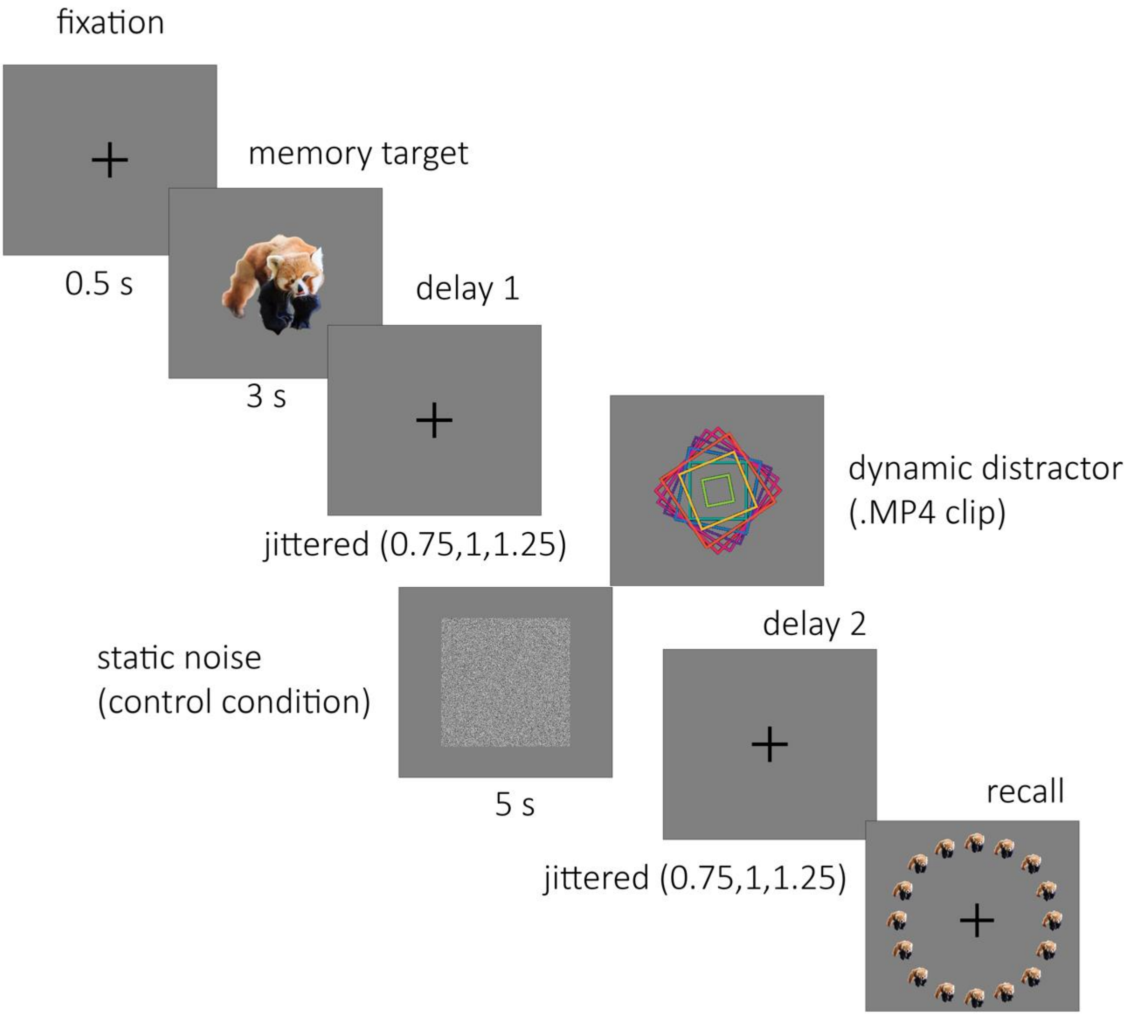


Figure 2. Trial scheme of the delayed recall task. On each trial, participants saw a fixation followed by the memory target – morphed image of an animal or a tool. During the delay period, a fixation cross was displayed on the screen before and after distractor presentation which appeared at the jittered time point. Two types of distractors could appear on the screen – static noise image served as a control condition or a dynamic clip of abstract colourful shapes (round or square). After the delay period, a recall wheel displaying 16 exemplars of the same object were shown on the screen. To provide their response, participants had to click on the image with a computer mouse and press the 'Tab' key on the keyboard. A black frame indicating the choice appeared around the image before moving to the next trial. The recall wheel was rotated on each trial. All stimuli were presented against a gray background.

After the delay, a recall wheel (radius: 11.19°) was presented on the screen with 16 morphed exemplars of the same object; the wheel was randomly rotated on each trial, so participants could not memorize the positions of any exemplars on the wheel. To respond, participants were asked to choose the memorised image by clicking on it with the left button of the computer mouse using the right hand and then to submit their response by pressing the space key with their left hand. After the response submission or after 7 seconds passed (the recall timeout), a gray screen was presented for 500ms (inter-trial interval, ITI) before the start of the next trial. All stimuli were presented at the centre of the screen on a gray background.

The in-lab experiment consisted of 132 trials of the delayed recall task divided into three blocks (44 trials each). The experiment lasted about 50 minutes in total.

Online experiment and questionnaires

In the online version of the delayed recall task, a separate set of stimuli was used: 2 memory target stimuli per category (animals, tools) and 2 distractor video clips per type (round, angular). The online task was almost identical with the no-distractor condition showing only a fixation cross on the screen (rather than a white noise image). The relative stimuli sizes in the online experiment were 10%×10% (of screen height of each participant's display size), for the fixation cross, 33%×33% for the memory target, 66%×66% for the distractor stimulus and 48% for the recall wheel radius. The online delayed recall task consisted of 48 trials in total. After the delayed recall task, participants were asked to complete 3 questionnaires assessing their trait-like tendencies to engage in worry (Penn State Worry Questionnaire, PSWQ (Meyer et al., 1990), and deliberate and spontaneous mind-wandering (Carriere et al., 2013; Vannucci et al., 2020). The online study lasted about 25 minutes.

MRI scan

Prior to the first stimulation session, all participants underwent an anatomical MRI scanning session. High-resolution anatomical T1-weighted magnetic resonance imaging (MRI) was acquired with a 3T Siemens MAGNETOM Prisma scanner (Siemens, Erlangen, Germany). The scanner parameters were as the folowing: TR = 2300 ms, TE = 2.29 ms, TI=900 ms, flip angle = 8◦, FOV = 240 mm, 192 slices and 0.94 mm thickness. The scanning session lasted ~6 minutes.

Stimulation target localization

TMS stimulation sites were localized in individual participants using a frameless stereotaxic procedure (BrainSight, Rogue Research, Montreal, Canada). Using this system, participants' head landmarks were co-registered to the same landmarks identified on their T1 scan. Using the 3D surface rendering of the brain, and the three planes of the T1-weighted MRI scan, the right DLPFC was identified based on individual gyral morphology as the midpoint of the middle frontal gyrus (Mylius et al., 2013).

The left primary somatosensory area (S1) served as a control site for stimulation. SI was marked relative to the individually found M1 according to the method previously confirmed to be precise in identifying the SI location (Holmes et al., 2019). Briefly, the TMS coil needs to be positioned 2 cm lateral and 0.5 cm posterior to the M1-hand area based on evidence that the S1-hand area is more lateral than traditionally assumed (Holmes et al., 2019).

TMS Procedure

All TMS procedures were performed using a 60mm diameter figure-eight coil (C-B60) and a MagProX100 stimulator (Magventure, Farum, Denmark) set up to deliver biphasic pulses in the standard pulse mode current direction (AP/PA). The coil was placed over the stimulation site, positioned tangentially to the skull with the coil-handle pointing posteriorly at an angle of 45 degrees to induce a current in the posterolateral-to-anteromedial direction (Chen et al., 2021; Rossini et al., 2005). If dlPFC stimulation site was uncomfortable or painful for a participant, the coil was moved posteriorly up to 1cm, or its angle was slightly rotated (away from midline).

Each participant attended two TMS sessions followed by the in-lab experiment. The sessions were held on different days with an interval of no more than 3 days between each session. Each session started with the localization of the stimulation target followed by TMS application. Either the right DLPFC or primary somatosensory area (SI, control site) was stimulated. The order of stimulation sites was counterbalanced across participants.

Motor threshold procedure

All participants underwent an active motor threshold (AMT) determination procedure. AMT was defined as the lowest TMS intensity applied over the left primary motor cortex (M1) eliciting visually detectable motor-evoked potential in 5 out of 10 TMS pulses during voluntary contraction of the abductor pollicis brevis muscle (APB) using standard methods and direct visualisation of APB movement described by Fitzgerald et al., 2002.

Theta-burst stimulation protocol (cTBS)

cTBS stimulation was delivered to the right DLPFC or left SI (control site) at 80% of individual AMT. Stimulation parameters were adopted from Huang et al., (2005). The protocol consisted of 50 Hz trains of 3 TMS pulses repeated every 200 ms continuously over a period of 40 s (600 pulses in total).

Eye-tracking

Eye data was recorded from twenty participants.

During the in-lab experiments, the eye movements were monitored with an EyeLink 1000 tracker (SR Research, Ottawa, Canada) with the eye data sampled at 1000 Hz. The eye tracker was placed underneath the presentation monitor, 85 cm away from the eyes of participants. The participants' heads were maintained fixed on a chinrest throughout the experiment. A nine-point calibration and validation was performed prior to each of the 3 in-lab experiment blocks of eye-data recording. Eye-data was only collected from one eye (same for each participant across both sessions).

Procedure

Figure 3 summarises the experimental procedure that was applied to each participant in the main experiment. Each participant completed the online study before the first laboratory session. The online experiment aimed to train participants to perform the task, so the effect of learning during two sessions in the laboratory would be minimised. Since the main purpose of the online experiment was to train participants and obtain the scores of mind wandering and trait worry questionnaires, we did not analyse the recall error of the data acquired online. The custom-written script based on JSPsych (de Leeuw, J. R., 2015) was used to run the online study. It was hosted on the Pavlovia platform (https://pavlovia.org/).

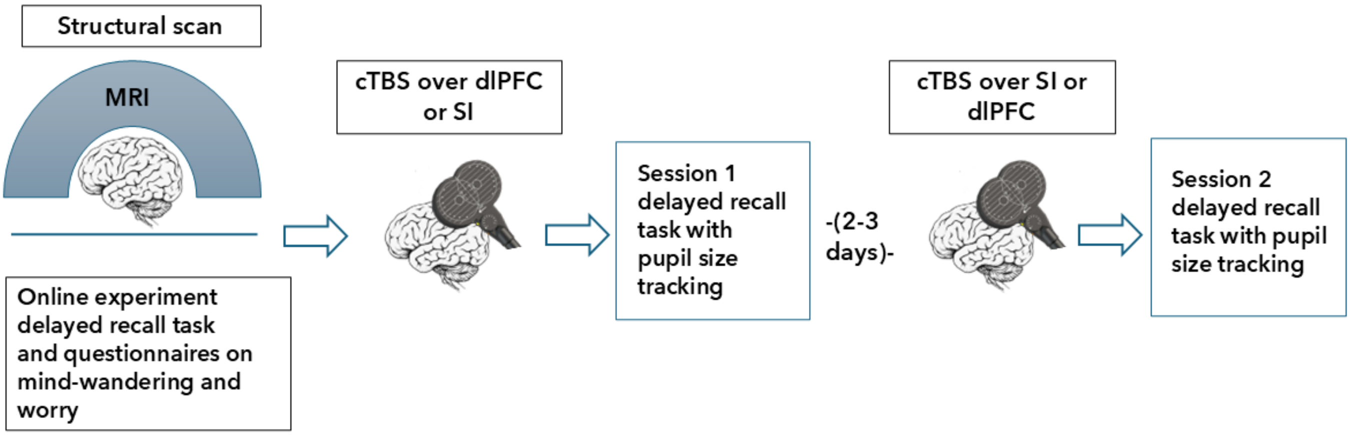


Figure 3. Main experiment scheme. Prior to the first session in the laboratory, a structural scan was acquired for neuronavigated scalp localisation of stimulation targets. Each participant underwent a short online experiment similar to the delayed recall task in the main experiment (instead of the static noise image, a fixation cross for presented in control trials); after the task, participants filled in questionnaires on spontaneous and deliberate mind wandering (Carriere et al., 2013, Vannucci et al., 2020) and trait worry (Penn State Worry Questionnaire, PSWQ, Meyer et al., 1990). For the first TMS session, participants underwent the procedure for defining their resting motor -threshold. At each TMS session, participants received cTBS over dlPFC or SI (order counter-balanced) followed immediately by the delayed recall task while their pupil size was tracked during the experiment.

The laboratory experiment consisted of an MRI scan and two TMS sessions. Participants performed the Delayed Recall Task in the laboratory after motor threshold determination and TMS target localisation (i.e., neuronavigation) at each TMS session. During the experiment, the head of a participant was maintained in a fixed position using a chinrest. The participants' eye movements and pupil size were recorded with the eye tracker.

Custom-written MATLAB script (The MathWorks Inc., 2021) using Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) were used to present the task, collect participants' responses, control the eye-tracker and collect the eye-movement data.

Behavioural data analysis

Memory recall precision was defined as the angular difference between the selected and target stimuli. To assess the distraction mitigation ability, we compared the mean memory recall precision in trials with dynamic versus static noise distractors using paired t-tests within sessions with TMS applied over dlPFC or S1. One participant was excluded from the analysis as an outlier using the Interquartile Range approach: statistical dispersion, representing the range within which the central 50% of the data in a dataset lies was calculated as the difference between the first quartile (Q1) and the third quartile (Q3), where Q1 marks the 25th percentile and Q3 marks the 75th percentile of the data.

Analysis of Mind-Wandering and Trait Worry questionnaire scores

To investigate the relationship between distractibility and the tendency to mind-wander or worry, mean memory precision was correlated with the related questionnaire scores. We used Spearman correlations with bootstrapping (with 10000 resamples) and tested these relationships separately for dlPFC TMS and control TMS sessions. All statistical analyses were performed using R Statistical Software (v4.3.2; R Core Team 2023).

Pupillometry

The pupil size was recorded throughout the experiment. All pupil data is reported in the units of measure by the EyeLink systems (arbitrary units). All pupil size analysis was performed in R Statistical Software (v4.3.2; R Core Team 2023). Eye movement data was not analysed as part of this study.

Original .edf files were converted to .ascii format and analysed using a custom-written script in R using the ‘eyelinker’ package. All missing data points were removed from the data. The average percentage of data removed was 7.29%, with the highest percentage of removed eye data being 15.49%.

To evaluate the fluctuation of attention throughout the experiment, we calculated the pre-trial variability (at fixation) reflecting the tonic (baseline) pupil response as in previous work (Unsworth & Robison, 2015). For each participant, the standard deviation of the mean pupil size across trials was divided by the mean pupil size at pre-trial fixation. Then, a linear regression ('lm' function) was used to test the relationship between the pre-trial pupil variability and mean memory recall precision. Stimulation site and distractor type were included in the model as regressors to test any interactions between pupil variability, TMS and distractors on the mean precision of object memory recall.

To test the effect of stimulation on pupil dynamics as a response to distractor presentation, we analysed the pupil size during the distractor phase. For each phase and trial, raw pupil size was re-sampled every 50 ms resulting in 10 bins for the pre-trial fixation phase and 100 bins for the distractor phase. The bins that contained missing data were calculated for each trial. The trials with more than 6 bins (300ms) with missing data were excluded from the analysis. To fill in the missing data points in the trials with less than 6 bins with missing data, the spline interpolation ('spline' function) was applied ensuring a good fit for non-linear data. The baseline was calculated as the mean of 10 pre-trial (fixation) phase bins. Distractor bins were baseline-corrected by subtracting the baseline value from each bin. If the percentage of rejected trials was more than 30% of the total (from both sessions), then the participant was excluded from the analysis. After this procedure, 10 out of 20 participants were included in the analysis of the distractor phase. Importantly, the session order in which TMS was applied over dlPFC remained counterbalanced; 5 participants received dlPFC stimulation in the first session and 5 in the second session.

Then, the median values were calculated for the binned pupil sizes for each distractor processing stage. The *initial dilation* stage usually lasts ~220 ms (Nieuwenhuis et al., 2011) and, thus, included the first 5 bins. For the *subsequent constriction* stage, we calculated the median pupil size from bins 5 to 20 as this stage usually lasts ~700-1000 ms and represents the depth of stimulus processing (Galeano-Keiner et al., 2023; Koevoet, et al., 2023). The *recovery stage* followed the pupil constriction and represented the increase of the pupil size back to its initial size before the stimulus presentation; this stage was defined by visually inspecting the plots; generally, the faster pupil rise after constriction denotes more effort exerted. Visual inspection of the pupil size plots revealed that this stage lasted from

~1000 to 2500 ms in all conditions, thus bins 20 to 50 were included in the calculation of *recovery stage* median. Finally, for the *maintenance stage* representing active memory retention (Beatty, 1982; Robison & Unsworth, 2019;), we included bins 50 to 100 corresponding to the 2500-5000 ms of distractor phase since the pupil response stayed stable after 2500 ms of the distractor onset.

Since averaged binned pupil size values did not follow the normal distribution (Shapiro-Wilk normality test, W = 0.95157, $p<0.01$), we used a non-parametric Wilcoxon signed rank test with continuity correction to compare medians of pupil response time courses during the distractor phase between stimulation sessions for each distractor type separately. The effect size (r) of the Wilcoxon test was calculated by dividing the Z-value (obtained by the Wilcox test) by the square root of the total number of observations (N).

To test if any stage of distractor processing reflected in the pupil dynamics correlated with subsequent performance (i.e. memory recall error on a trial), we ran Spearman's correlations for the median values of each processing stage (initial dilation, subsequent constriction, recovery and maintenance) with the absolute error on a trial-by-trial basis in each of the distractor types and stimulation sessions.

Results

Effect of dlPFC-TMS on memory precision

To evaluate the effect of TMS applied over dlPFC under distraction, we compared the mean differences in object memory precision (i.e. the inverse of the absolute circular error) for dynamic versus static distractor conditions. As Figure 4A illustrates, dynamic distractors did not lead to a significant decrease in memory precision (i.e. larger error) in comparison to the static noise distractors in the control TMS session ($t(25) =0.758$, $p=0.455$). In contrast, dlPFC-TMS resulted in significantly lower memory precision ($t(25) = 2.345$, $p = 0.027$, Cohen's $d=0.18$) in the presence of dynamic distractors compared to static distractors as shown in Figure 4B.

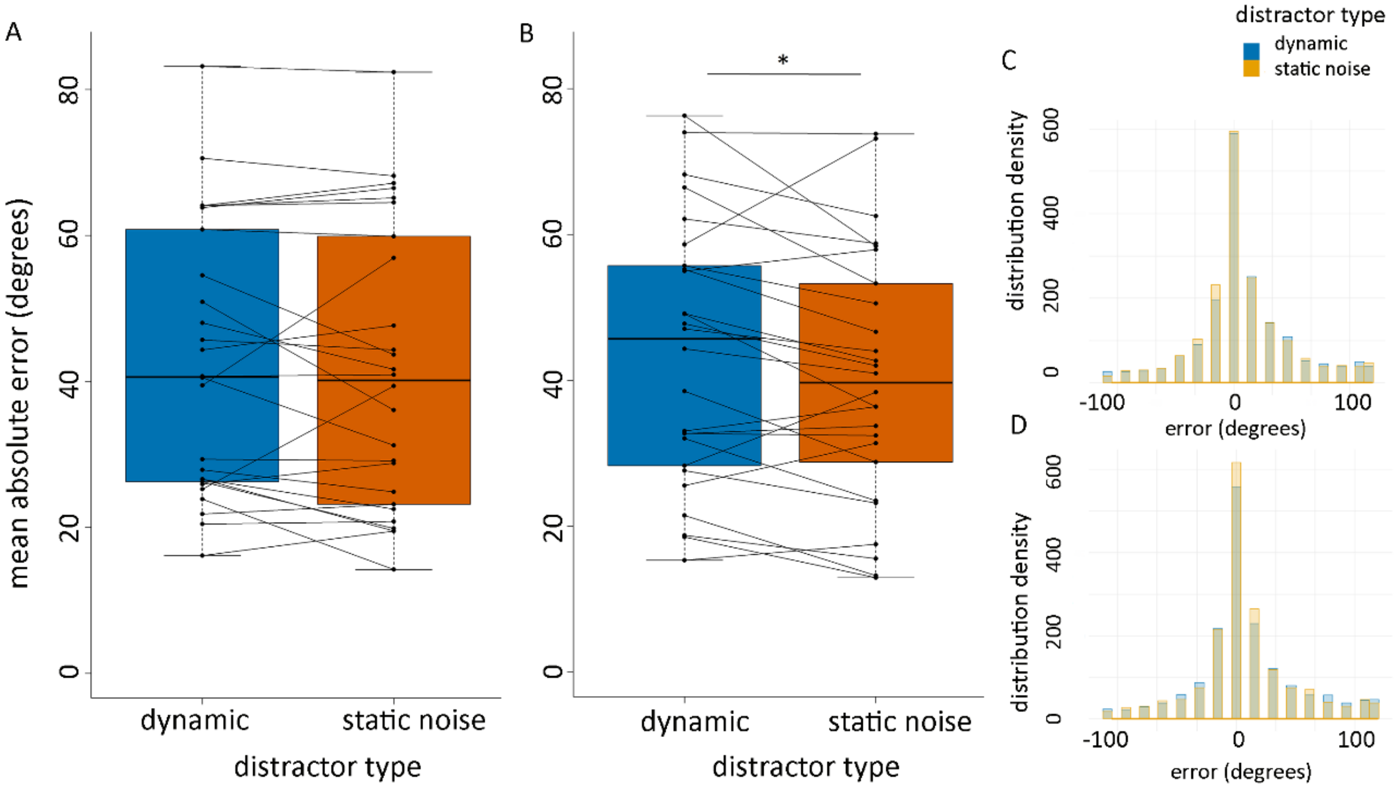


Figure 4. Effect of dlPFC stimulation on distractor mitigation ability (N=26). Distraction mitigation ability was defined as the difference between the mean absolute error in trials with dynamic (M=42.49, SD=18.54) and static distractors (M=41.45, SD=19.41). In the control stimulation session (A), no difference was found - dynamic distractors did not impact memory recall precision for objects. However, the stimulation over dlPFC (B) led to a significantly larger error in the trials with dynamic distractors (M=43.6, SD=17.94) compared to static distractors (M=40.31, SD=17.37) (*paired t-test, p<0.05). Distributions of the errors in two distractor conditions in the control stimulation session (C) and session with dlPFC TMS (D), show that stimulation over dlPFC resulted in slightly wider distribution in the dynamic distraction condition due to fewer zero errors and a larger number of large errors.

Relationship between mind-wandering and worry, and memory recall precision

In the SI TMS session, spontaneous mind wandering positively correlated with memory recall precision (r(24) = -0.53, p= 0.007, 95% CI [-0.89, -0.29]), whereas in dlPFC TMS session, this correlation disappeared (r(24)= 0.19, p=0.3672, 95% CI [-0.65, 0.23]).

There were no significant correlations between deliberate mind-wandering or worry and memory precision (see Table 1 for all correlation results and Figure 5 for all regression plots).

Averaged memory recall across both distractor conditions didn't significantly correlate with any questionnaire scores. However, the relationship between the trait worry and overall memory recall precision in the absence of stimulation over dlPFC demonstrated a slight trend (r(24) =.37, p=0.067, 95% CI [0.0656, 0.8261]).

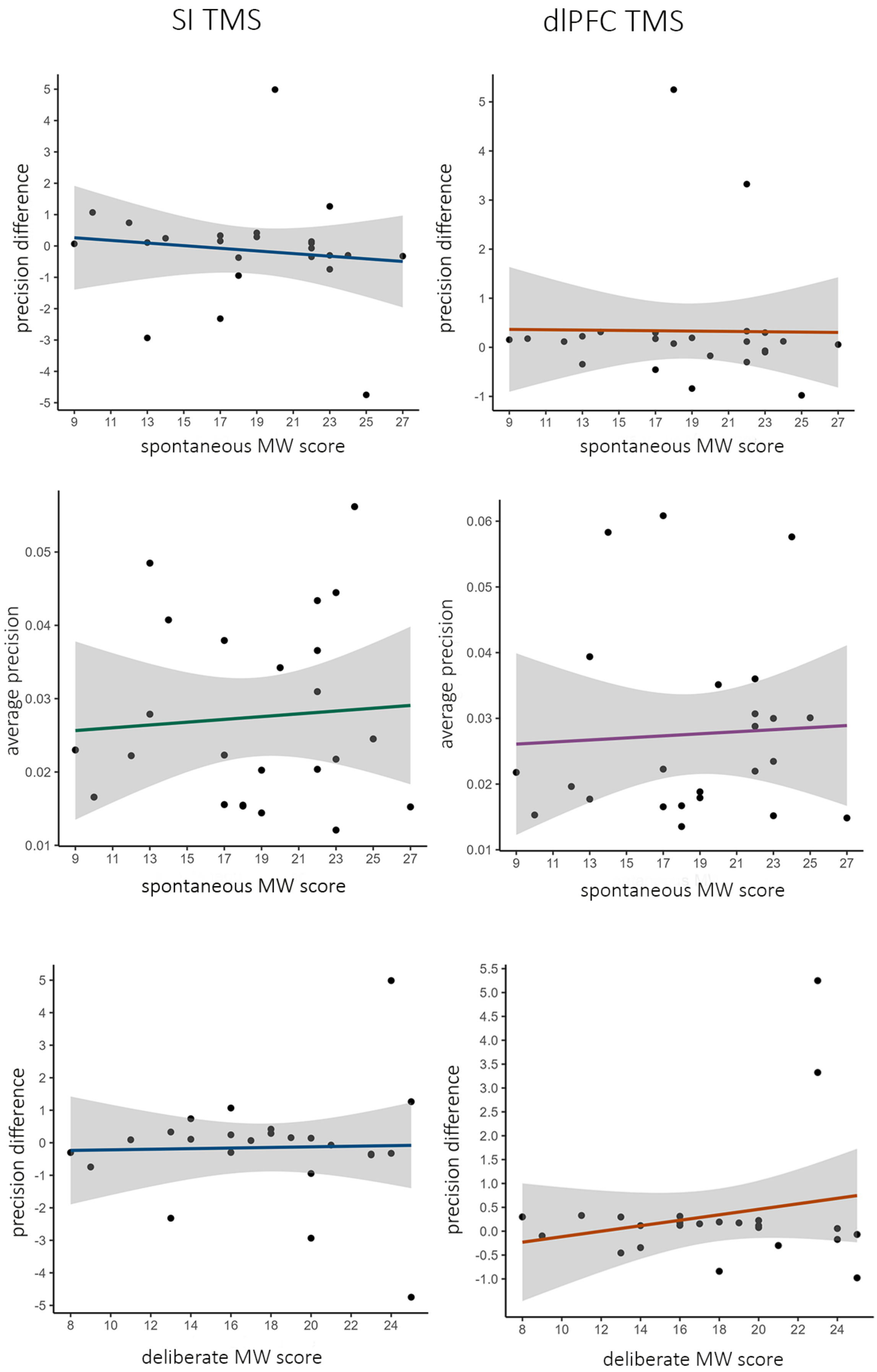
Table 1. Bootstrapped (10000 samples) Spearman correlations of scores of questionnaires and recall precision difference (dynamic minus static noise precision) or average precision across both distractor conditions in both TMS sessions. MW denotes Mind Wandering, PSWQ denotes Penn State Worry Questionnaire score.

Correlation pair	**dlPFC TMS**			**SI TMS**		
	r	p	95% CI	r	p	95% CI
Spontaneous MW and precision difference	-0.19	0.3672	[-0.65, 0.23]	-0.53	0.007	[-0.89, -0.29]
Spontaneous MW and average precision	-0.093	0.66	[-0.5283, 0.3475]	-0.0065	0.9758	[-0.4167, 0.4488]
Deliberate MW and precision difference	-0.058	0.78	[-0.5470, 0.3844]	-0.22	0.29	[-0.6498, 0.1580]
Deliberate MW and average precision	-0.13	0.52	[-0.1955, 0.6179]	-0.18	0.39	[-0.2626, 0.5683]
PSWQ and precision difference	-0.23	0.27	[-0.6506, 0.1064]	-0.09	0.67	[-0.2891, 0.5260]
PSWQ and average precision	-0.21	0.31	[-0.1660, 0.6636]	-0.37	0.067	[0.0656, 0.8261]

Tonic pupil response predicts memory recall precision irrespective of distractor type or stimulation

Since fluctuations in the tonic pupil response reflect the stability of cognitive control, we investigated whether variability in baseline pupil size (measured during pre-trial fixation) could predict subsequent memory recall errors, which are inversely related to precision. The linear regression analysis (Adjusted R^2 =0.09, F(3, 76) = 3.737, p=0.014) revealed a positive relationship between increased pupil variability during the pre-trial fixation and mean absolute error (shown in Figure 5): pupil variability can significantly predict the subsequent mean absolute recall error (B = 16.0137, SE = 4.8736, t(76) = 3.286, p = 0.00154) irrespective of distractor type (B= -0.3, SE = 0.29, t(76) = -1.048, p = 0.29 or stimulation site (B = -0.1754, SE = 0.29, t(76) = -0.6, p = 0.54)).

In addition, in a separate regression analysis, we tested whether tendencies to mind wander and worry influence the relationship between tonic pupil activity and mean absolute error. The result revealed that worry ($B = 0.03$, $SE = 0.03$, $t(12) = 1.13$, $p = 0.27$) as well as spontaneous ($B = 0.002$, $SE = 0.07$, $t(12) = 0.02$, $p = 0.98$) and deliberate ($B = 0.02$, $SE = 0.06$, $t(12) = 0.37$, $p = 0.72$) mind-wandering had no effect on the positive relationship between tonic pupil size and mean absolute error.


SI TMS
dlPFC TMS
precision difference
spontaneous MW score
5
4
3
2
1
0
-1
-2
-3
-4
-5
9
11
13
15
17
19
21
23
25
27
average precision
0.01
0.02
0.03
0.04
0.05
0.06
deliberate MW score
8
10
12
14
16
18
20
22
24
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0

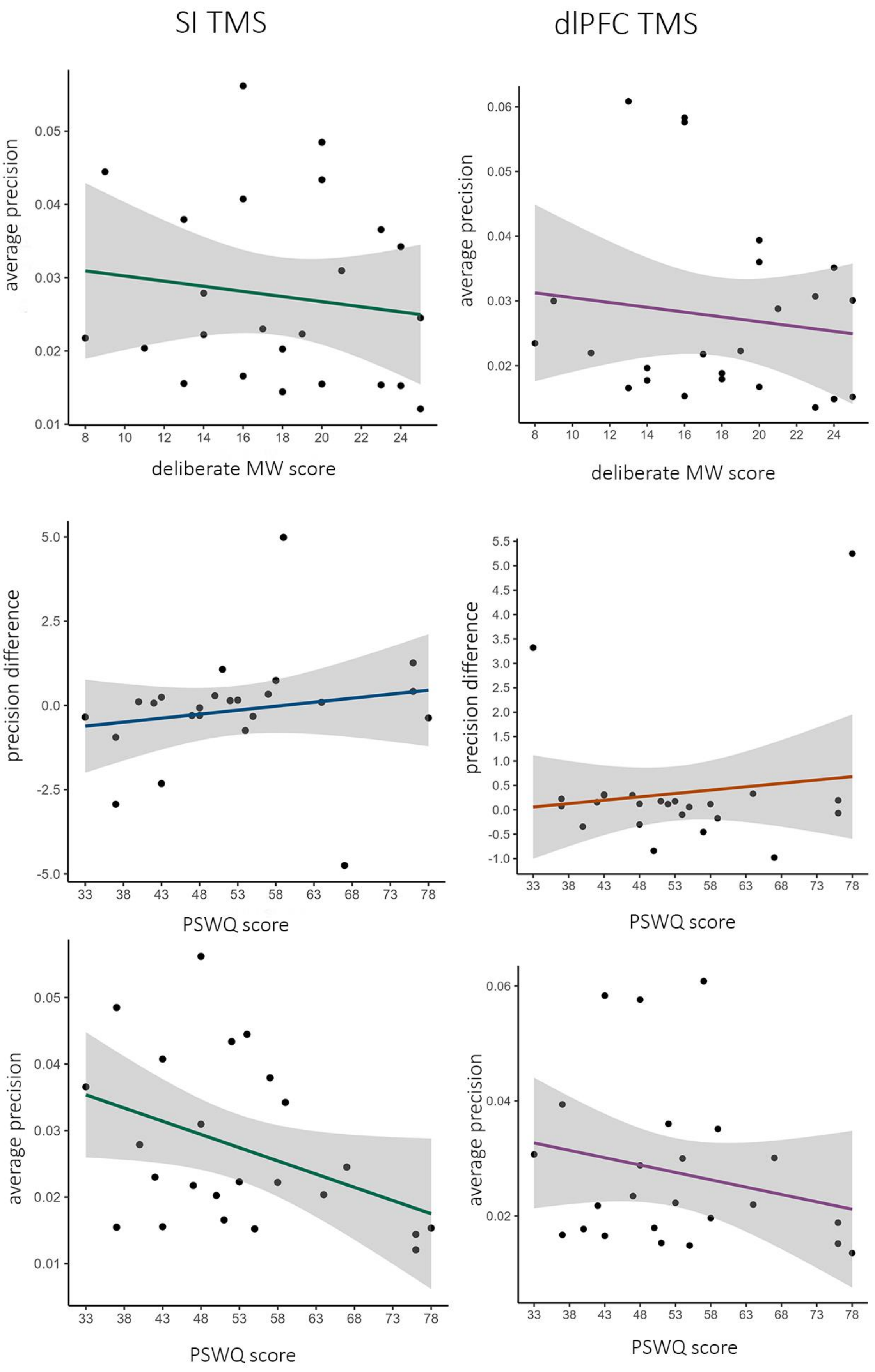

SI TMS
dlPFC TMS
average precision
0.01
0.02
0.03
0.04
0.05
0.06
8
10
12
14
16
18
20
22
24
deliberate MW score
precision difference
-5.0
-2.5
0.0
2.5
5.0
-1.0
-0.5
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.5
33
38
43
48
53
58
63
68
73
78
PSWQ score

Figure 5. Questionnaire score correlations with memory recall precision (bootstrapped, 10000 resamples) in two TMS sessions. MW stands for mind-wandering, PSWQ stands for Penn State Worry Questionnaire. X-axis denotes the score of each questionnaire. Y-axis denotes *precision difference* in the mean difference between trials with dynamic and static noise distractors or *average precision* across both distractor conditions.

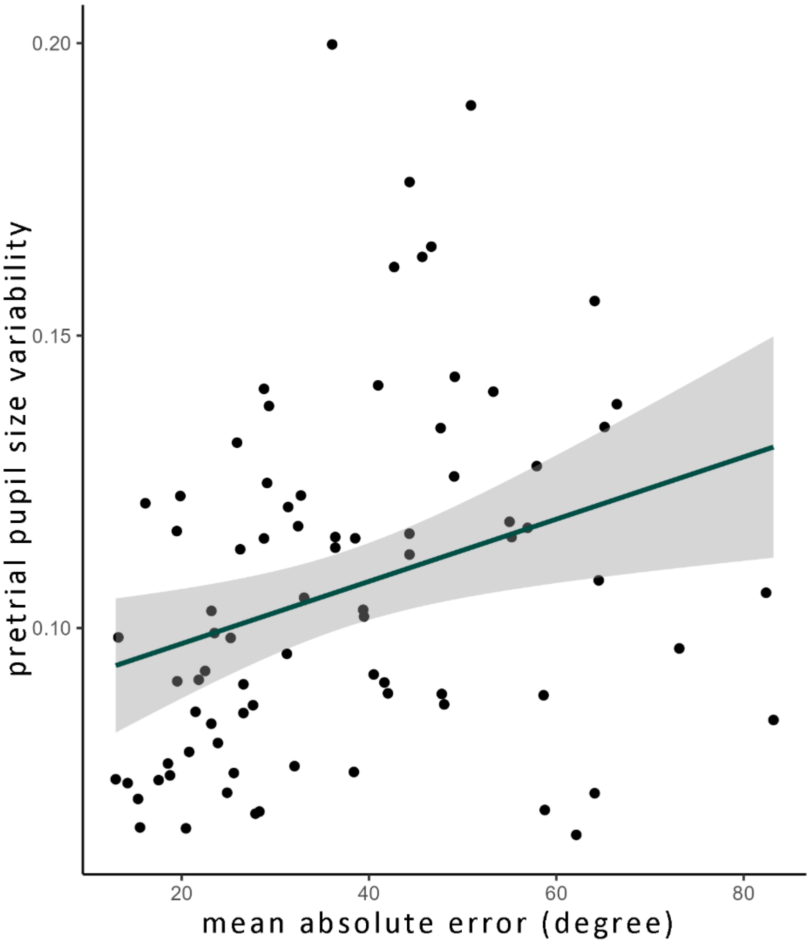


Figure 6. Tonic pupil response predicts mean absolute errors. Scatter plot of the mean absolute error in relation to the pretrial pupil variability (fluctuation of the tonic pupil response measured at fixation) and regression line demonstrate the strong positive relationship – pupil size variability predicts mean absolute recall error in the delayed memory task highlighting the role of fluctuations in cognitive control in WM recall.

Effect of dlPFC stimulation on pupil response to distractors

To evaluate the effect of stimulation on the pupil response to distractors, pupil size change during the distraction phase was averaged across the trials for each participant, stimulation site and bin (see Figure 6A for the pupil size time course); each bin represents the baseline-corrected pupil size change over a 50 ms interval.

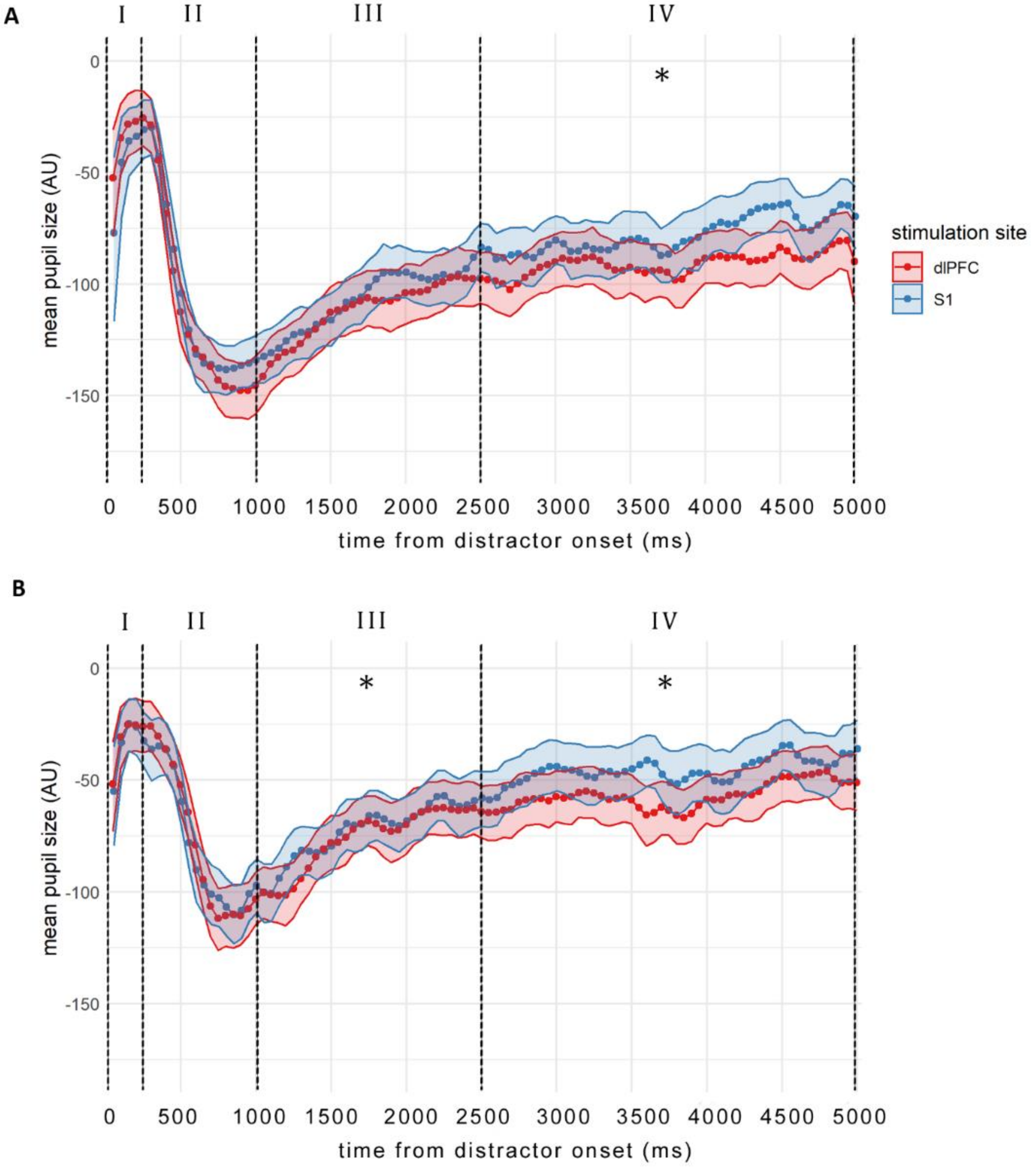


Figure 7. Effect of stimulation over pupil response dynamics during distractor phase. Pupil response was baseline-corrected and sampled to 50 ms bins, shaded plot areas correspond to the confidence intervals calculated using a bootstrapping approach (1000 samples) for mean value calculation. Roman numerals denote the processing stages: initial dilation (I), subsequent constriction (II), recovery (III) and maintenance (IV). Pupil response during static distractor presentation (A) was decreased by dlPFC TMS, but only in the

maintenance phase. Whereas pupil response during dynamic distractor presentation(B) was decreased by dlPFC TMS, both in recovery and maintenance phases. *p<0.05

There was a significant overall effect of stimulation on the pupil response to static distractors (V = 2179, p <0.001, r=0.75). Pupil constriction in trials with static distractors was higher in the session when the stimulation was applied over dlPFC (median: -94.819), compared to S1 stimulation (median: -84.927). dlPFC TMS also impacted pupil response to the dynamic distractors (V = 178877, p <0.001, r= 0.24): pupil constriction during distractor presentation was higher (median: -70.54338), than in SI TMS session (median: -58.46).

To examine the pupil dynamics as a response to distraction more closely, we divided the distractor phase into several processing stages: initial dilation, subsequent constriction, recovery and maintenance. We compared the pupil size between stimulation sessions within each processing stage and distractor type (see Table 1 for all the related median pupil size values). Stimulation over dlPFC did not impact the initial dilation phase of both static (V = 39343, p-value = 0.9167, r=0.005) and dynamic (V = 36999, p-value = 0.3538, r = 0.04) distractor processing. Subsequent pupil constriction as a reflection of the processing depth of a (distracting) stimulus also was not impacted by the stimulation; this was true for trials with static (V = 39343, p-value = .9167, r=.005) as well as dynamic (V = 36599, p-value = .2698, r=.05) distractors. However, stimulation did affect (V = 33001, p-value = .0071, r=.13) the pupil size recovery stage, but only during the processing of dynamic distractors; specifically, the pupil size was reduced by dlPFC TMS. Whereas during the static distractor processing the pupil response during the recovery was the same(V = 35170, p-value = .08312, r=.08) in both sessions. Finally, dlPFC TMS significantly reduced pupil size during the maintenance stage in both distractor conditions. Namely, when dynamic distractors were presented, pupil size was smaller (V = 31971, p-value = .00167, r=.15) when TMS was applied over dlPFC than over the SI. Similarly, the maintenance during static distractor presentation was reflected in lower pupil size (V = 31255, p-value = .000546, r=0.17) in dlPFC TMS session compared to the SI TMS session.

Table 2. Median pupil size values in each of the distractor conditions, stimulation sessions and distractor-processing stages.

	dynamic		**static**	
	dlPFC TMS	SI TMS	dlPFC TMS	SI TMS
initial dilation	32.47	28.79	-31.56	-26.22
subsequent constriction	-85.02	-78.86	-114.95	-105.12
recovery	-67.77	-46.64	-116.07	-105.86
maintenance	-89.28	-73.43	-89.28	-73.43

Relationships between tendencies to mind-wander and worry

We also analysed how the tendency to worry, and deliberate or spontaneous mind-wandering correlate with each other. We also tested whether these two types of mind-wandering measured distinct sides of mind-wandering.

The analysis revealed no significant relationship between spontaneous and deliberate mind-wandering (r(24) =.24, p=0.2422, 95% CI [-0.1601, 0.6089]), confirming that these two questionnaires capture distinct types of mind-wandering as displayed in Figure 8 (left). Similarly, the tendency to worry did not correlate with spontaneous mind-wandering (r(24) =.30, p=0.13, 95% CI 0.0362, 0.5878]; middle), nor with deliberate mind-wandering (r(24) =.23, p=0.26, 95% CI [-0.1160, 0.6918]; right).

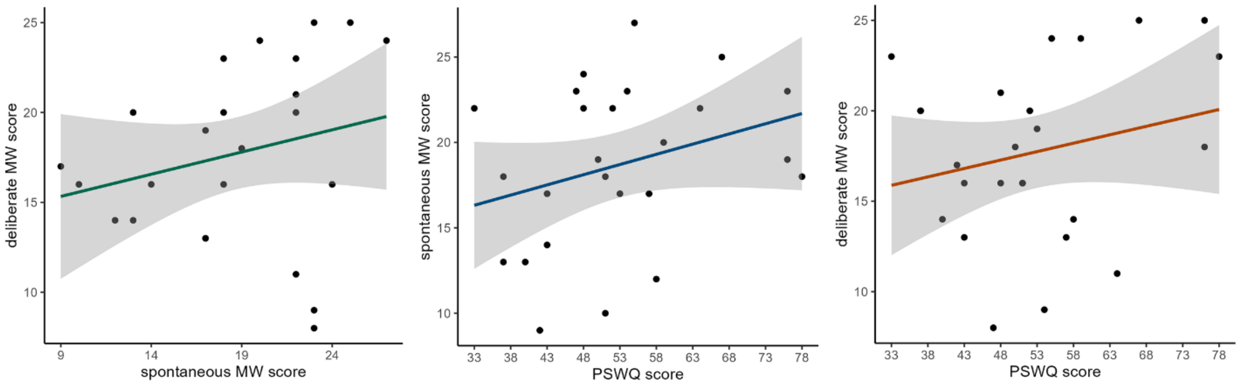


Figure 8. *Correlation plots of spontaneous and deliberate mind-wandering tendencies (left panel) and their relationship with worry (PSWQ) in the middle and right panels, respectively.* No relationship was found between any of the tendencies analysed. MW stands for Mind-Wandering, PSWQ stands for Penn State Worry Questionnaire.

Discussion

Effective mitigation of irrelevant information in WM is crucial for optimal daily performance. While numerous studies suggest the dorsolateral prefrontal cortex (dlPFC) plays a key role in managing distractions (Clapp et al., 2010; Dolcos et al., 2007; Yoon et al., 2006), few have used stimulation techniques to investigate this causally (Feredoes et al., 2011). This study examines the causal role of dlPFC in mitigating visual distractions during WM maintenance. We evaluated memory precision for objects when dynamic colourful images or static noise were displayed during the delay period, and applied stimulation over dlPFC or a control region (SI) in two counter-balanced sessions.

Memory precision for realistic objects did not differ between trials with dynamic versus static noise distractors in the control TMS session; when TMS was applied over dlPFC, the precision of memory recall decreased (i.e. higher mean error) in the presence of dynamic distractors versus trials with static noise. These results show that salient stimuli used as dynamic distractors in the present study could be successfully mitigated and did not impact memory precision when the control site (SI) was stimulated, but this mechanism was disrupted by the TMS applied over dlPFC, demonstrating a causal role of this region in distractor mitigation.

This study provides new evidence for the importance of dlPFC in mitigating visual distractions in WM. Building on the findings of Feredoes et al. (2011), we utilised continuous Theta Burst Stimulation (cTBS) instead of single-pulse TMS and assessed memory recall precision for objects. Importantly, rather than comparing an unfilled delay with delays involving distracting images of faces or houses, we introduced visual information in both distraction conditions by varying their visual characteristics: dynamic colourful images (aimed to be more engaging and consume cognitive resources more) versus static noise images (aimed to be less engaging). This approach was designed to ensure that visual processing was engaged in both types of trials while involving the cognitive effort to a different extent.

This study provides additional evidence for the distractor mitigation mechanism proposed by Feredoes et al. (2011), suggesting that the dlPFC plays a crucial role in maintaining relevant information by amplifying the neural signals associated with memory targets. This neural enhancement helps to protect memorandum from the potential interference caused by distracters. This was evidenced by increased activity in the brain regions associated with the current memory targets when TMS was applied to dlPFC during the presence of distracters. Another study used a concurrent TMS-fMRI-MVPA approach to provide further insights into this mechanism (Jackson et al., 2021). Specifically, the researchers investigated TMS of dlPFC modulated representations of relevant information in both the Multiple Demand (MD) network—which is involved in a variety of tasks requiring cognitive control (Duncan, 2010; Fedorenko et al., 2013; Assem et al., 2020)—and the visual cortices. This was examined when participants performed a selective attention task that required responding to a relevant item feature in the presence of an irrelevant one. Decoding analysis revealed the reduced coding of relevant information in MD regions and visual cortices following dlPFC TMS, indicating that this region enhances relevant information in those regions (Jackson et al., 2021). In contrast, irrelevant information coding was unaffected by dlPFC TMS, suggesting that the primary role of this region is to enhance relevant information rather than suppress irrelevant information. Furthermore, this finding supports the broader mechanism of dlPFC-driven cognitive control over memoranda by maintaining and strengthening goal-relevant representations in the posterior cortex. Further evidence for this mechanism comes from monkey neurophysiology studies (Suzuki and Gottlieb, 2013; Jacob and Nieder, 2014; Parthasarathy et al., 2019) demonstrating that cognitive control may be implemented by morphing memoranda code within dlPFC and preserving low-dimensional memory information for guiding behavior (Parthasarathy et al., 2019); while posterior regions maintain a stable memory code (Suzuki and Gottlieb, 2013; Jacob and Nieder, 2014) which may potentially serve as a source of detailed memory information that dlPFC may flexibly recruit to optimise goal-directed behavior under varying task demands. Moreover, Suzuki and Gottlieb, (2013) suggested that lower dlPFC activation during distractor presentation demonstrated inhibition of distracting stimuli processing. In contrast, Jacob and Nieder (2014) showed that dlPFC processed distracting information and did not suppress it. In contrast, information related to the memorised item was disrupted by the distractor but reinstated after the distractor phase, suggesting the target enhancement as the main mechanism for memory protection rather than distractor suppression. Although further human studies are needed to resolve these conflicting findings, Jackson et al. (2021) proposed a representational account of the mechanism underlying the enhancement of relevant information (e.g., memoranda). Therefore, the causal result in the present study may be explained in light of the following findings: dlPFC could enhance the memorandum in the presence of dynamic distractors, but when the stimulation disrupted the activity in this region, this enhancement was disrupted and led to lower object recall precision. Importantly, the present finding expands the previous evidence by demonstrating the causal role of dlPFC in mitigating visual distractions in WM for objects using a continuous recall to provide a more fine-grained view of the distraction impact. Future research should investigate the role of dlPFC from a network-level perspective, as this region likely orchestrates cognitive control in concert with other brain regions. One potential set of regions, as highlighted by the findings of Jackson et al. (2021), is the MD network. Future research could use similar concurrent methodologies or other connectivity-based analyses to further clarify the integrative role of dlPFC within broader cognitive control circuits.

To investigate the role of cognitive control further, we examined how tendencies towards mind-wandering or worry, characterised by reduced cognitive control, relate to the ability to mitigate distractors in WM. First, the results revealed the role of volitional control over mind wandering; individuals who mind-wandered deliberately were not more susceptible to distraction. Lower volitional control over mind wandering demonstrated in spontaneous mind wandering, revealed a different pattern. Opposite to our expectations, the memory recall precision of those individuals was less impacted by the external distraction. Because participants were less distracted by the dynamic stimuli presented during the delay period, they might have been more distracted by the static noise instead. Many participants reported after the experiment that they enjoyed viewing the abstract, vibrant shapes, which helped them stay focused on the task and likely enhanced their cognitive control over memory maintenance. Conversely, individuals who deliberately engaged in mind-wandering could have found the noise distractors boring and turned to internally generated thoughts, but due to their efficient cognitive control, this did not interfere with their task performance.

An alternative interpretation can be drawn from load theory (Lavie, 1995), which posits that participants perform better on tasks with increased perceptual load because their cognitive efforts are fully engaged by the task, leaving no spare resources for processing distractors (Lavie & Tsal, 1994). Spontaneous mind-wandering is associated with lower WM capacity (Robison & Unsworth, 2018); this type of mind-wandering can lead to lapses in focus, reducing the cognitive resources available for task-related perceptual processing. Therefore, there is a possibility that individuals prone to spontaneous mind-wandering allocated most of their cognitive resources toward memory encoding, leaving fewer processing efforts for dynamic distractors. At the same time, they still could have retained enough capacity to process static noise images, which required minimal effort and could potentially interfere with memory maintenance. Yet, the present study did not assess the online allocation of cognitive resources to memoranda or distractors to support this interpretation.

However, this negative relationship between spontaneous mind-wandering and the ability to mitigate external distractions disappeared when TMS was applied over dlPFC. Mind-wandering is associated with increased activation of the Default Mode Network (DMN), which includes brain regions such as the medial prefrontal cortex, posterior cingulate cortex, and precuneus (Weissman et al., 2006; Christoff et al., 2009). Research suggests that frontoparietal regions, including dlPFC, regulate the DMN by inhibiting its activity during task-focused behaviour (Smallwood et al., 2012). When cognitive control is no longer critical, this inhibition weakens, allowing the DMN to become more active and facilitate mind-wandering. Therefore, TMS applied over dlPFC in the present study could have decreased its inhibitory influence on the DMN, leading to increased activity within this network and, consequently, more episodes of spontaneous thoughts. This potentially led to an equalization of the distractor effect: the visual features of distracting stimuli became irrelevant as internally generated thoughts fully consumed cognitive resources, diminishing the impact of external distractions. However, the present experiment does not provide evidence for this mechanism. Future studies could apply dlPFC TMS in combination with neuroimaging and real-time assessments of mind-wandering (e.g., probes appearing at random intervals during the experiment) in a working memory task with distractors. This approach would help determine whether dlPFC stimulation increases DMN activity and spontaneous mind-wandering episodes, and importantly, whether these episodes coincide with greater susceptibility to distractors in the task. Moreover, to gain a deeper understanding of how cognitive control mechanisms manage both internal and external distractions, future studies should employ a network-based approach to explore the relationship between the two types of mind-wandering examined in this work—spontaneous and deliberate—and the ability to mitigate distractors in WM tasks.

In the present study, we also examined whether higher levels of trait worry affect distractor mitigation and found no significant relationship, except for a slight positive trend between overall memory recall precision and questionnaire scores. This finding can be better understood through the lens of Attentional Control Theory (ACT; Eysenck et al., 2007), which suggests that anxiety (which is accompanied by excessive worry) primarily impacts processing efficiency—how cognitive resources are utilized—rather than processing effectiveness, or the actual performance outcomes. According to this theory, individuals with higher levels of anxiety must exert more cognitive effort to attain the same level of performance as those with lower anxiety levels (Berggren & Derakshan, 2013). This compensatory mechanism helps explain inconsistencies observed in previous studies examining trait anxiety and WM, as reviewed by Berggren and Derakshan (2013). Specifically, the cognitive load, i.e. amount of mental effort and resources required to perform a particular task or process information, influences how susceptible anxious individuals are to distraction. In tasks with low to medium cognitive loads, anxious individuals may perform comparably to non-anxious individuals because their cognitive resources are sufficient to manage both the task demands and their anxiety. However, in tasks with high cognitive loads, their performance may decline due to the depletion of cognitive resources needed to cope with both the task and the anxiety-induced processing inefficiencies (Berggren & Derakshan, 2013). Therefore, potentially, the lack of a significant connection between worry and distractor mitigation in the present work could be due to the low cognitive load of the task. Participants were required to memorise only one image, which may not have been demanding enough to exhaust their cognitive resources. As a result, even individuals with higher levels of trait worry did not exhibit a significant reduction in memory recall precision because the task did not strain their processing capacity. Additionally, these compensatory mechanisms could explain why dlPFC TMS did not interact with the tendency to worry in the present study, despite the involvement of this region in regulating anxiety (Basten et al., 2011; Sagliano et al., 2019). However, future studies are needed to test these compensatory mechanisms neurally to confirm these interpretations.

Nevertheless, the observed trends indicate that distractors could potentially affect the memory performance of anxious individuals by increasing the cognitive load of the task. The presence of distractors may require additional cognitive effort to maintain focus, thereby revealing the processing inefficiencies associated with anxiety. Future studies could vary cognitive load in WM to better understand the threshold at which anxiety impairs performance, explore the impact of different types of distractors (emotional vs neural) and how these factors interact to affect cognitive effort as well as investigate the neural compensatory mechanisms employed by individuals prone to excessive worry. In addition, since both mind wandering and worry are characterised by increased internal distraction, i.e. internally generated (worrying) thoughts, future research needs to address what cognitive control mechanisms are involved in resolving internal distraction and how these mechanisms interact with those involved in external distraction resolution.

Moreover, we examined the relationship between worry and spontaneous and deliberate mind-wandering. None of the tendencies measured correlated. This result was surprising since anxiety often results in increased mind-wandering;

namely, Figueiredo et al. (2020) examined mind-wandering and anxiety symptoms in individuals with Attention-Deficit Hyperactivity Disorder (ADHD) and demonstrated that mind-wandering was associated with anxiety levels, independently of an ADHD diagnosis.

Furthermore, dlPFC plays an important role in both regulating anxiety (Basten et al., 2011, 2012) and mind-wandering (Christoff et al., 2009; Axelrod et al., 2015). However, the present study lacks sufficient measures to examine the interaction between worry and mind-wandering during WM protection. The study design did not include real-time assessments of active worry or mind-wandering episodes resulting from dlPFC stimulation, nor did it capture their subsequent effects on WM performance and memory protection. Therefore, future studies may pair neural stimulation together with Experience Sampling Methods (ESM), asking individuals to report their thoughts at random intervals, providing real-time measurement of active worrying (Sari et al., 2017) or state of mind-wandering (Unsworth & Robison, 2015) during the WM task performance. One prediction would be that following dlPFC stimulation, participants would experience more active worry during the experiment and, consequently, engage more in mind-wandering, which could potentially lead to more distractibility in a WM task. It would be of particular interest to test if the level of cognitive control may be reflected in the type of mind wandering increased; specifically, if individuals have low cognitive control, they may engage in spontaneous mind wandering not being able to focus on the task, but if their cognitive control is higher, they may potentially engage in deliberate mind wandering using it as a coping strategy (Baird et al., 2011). Although speculative, these possibilities may be addressed by future studies designed to measure the real-time effects of dlPFC disruption on the complex interaction between worry and mind-wandering.

Another approach we used to elucidate the role of cognitive control in distractor mitigation was pupillometry. As highlighted in numerous studies reviewed by Van der Wel & Van Steenbergen (2018), pupil dilation serves as an indicator of cognitive effort exerted during tasks, making pupillometry a valuable tool for measuring mental effort. The locus coeruleus-norepinephrine (LC-NE) system plays a crucial role in regulating pupil dilation by modulating arousal and it operates in two firing modes: tonic and phasic (Aston-Jones & Cohen, 2005; Usher et al., 1999). Tonic pupil activity reflects overall task engagement, where larger fluctuations are associated with weaker cognitive control in WM as shown by Unsworth & Robison (2015) and Robison & Unsworth (2018). We used the same approach to measure the fluctuations in tonic pupil response and found that these fluctuations predicted average recall precision for objects irrespective of the distractor type. This finding indicates that maintaining cognitive control to sustain effort throughout the experiment is essential for task performance and handling both types of distractions. Furthermore, variations in tonic pupil response predicted memory recall precision, even when TMS was applied over dlPFC, suggesting that the pupil fluctuations captured by this analysis reflect a different control which does not require dlPFC. Indeed, activity in dlPFC is regulated by the LC-NE system, which signals dlPFC when increased effort is needed to meet the demands of the task at hand (Grueschow et al., 2022; Grueschow et al., 2020). Future studies should explore the interaction between the LC-NE system and dlPFC in mitigating distractions in WM to better understand the contribution of each type of control to this process.

Whereas phasic pupil activity tracks the immediate cognitive effort in response to specific stimuli or events (Naber et al., 2013). In WM, this type of pupil response at memory encoding predicts subsequent recall (Naber et al., 2013) and reflects the active maintenance of memorandum (Robison & Unsworth, 2019; Zokaei et al., 2019). In the present study, we asked if dlPFC TMS would impact the allocation of cognitive efforts exerted for distractor processing and memory maintenance by comparing phasic pupil response during distractor presentation across two sessions and processing stages, i.e. initial dilation, subsequent constriction, recovery and maintenance. We found that dlPFC stimulation reduced pupil dilation during the later stage of the distractor phase (2500 ms after distractor onset), which likely reflected active memory maintenance (Beatty, 1982; Robison & Unsworth, 2019). Importantly, this effect was observed irrespective of the distractor type, indicating that dlPFC stimulation modulated cognitive effort allocation in a generalisable manner rather than being specific to certain distractor characteristics. In contrast, dlPFC TMS did not affect the phasic pupil response during the earlier stages of the distractor phase (<2500 ms), suggesting that initial distractor processing remained unaffected. These findings collectively indicate that stimulation over dlPFC did not disrupt the processing of distracting information but rather impaired the maintenance of the memory target, as reflected in reduced pupil size during the later phase of the distractor period. Specifically, this reduction in pupil size suggests that dlPFC TMS led to a decreased allocation of cognitive resources for sustaining memoranda. This interpretation aligns with previous TMS-fMRI-MVPA evidence (Jackson et al., 2021), which demonstrated that disrupting dlPFC-driven cognitive control primarily affected the representation of relevant information (memorandum) rather than distractor inhibition. Thus, our findings support the notion that dlPFC plays a crucial role in sustaining cognitive resources for memory retention rather than filtering out irrelevant information. Moreover, the reduced overall phasic response resulting from dlPFC stimulation may indicate an anticipatory enhancement of the memorandum, as suggested by earlier neuroimaging findings from Sakai et al., (2002) and the monkey neurophysiology study by Suzuki and Gottlieb (2013). Specifically, the dlPFC may enhance memory representations in posterior brain areas to prepare for anticipated distractions, protecting them from distortion (Sakai et al., 2002; Suzuki & Gottlieb, 2013). In the present study, since participants could anticipate a distracting input on every trial, the dlPFC may have been actively amplifying memory maintenance throughout each trial. Consequently, when dlPFC activity was disrupted by TMS, it led to a lower overall phasic response—indicating reduced cognitive effort deployed for memory maintenance.

Altogether, these results indicate that pupillometry can be a useful tool for measuring different facets of cognitive control in WM and distraction. Specifically, the stability of the cognitive control can be revealed by tonic pupil activity, and the allocation of cognitive resources to the different phases of the task (i.e., stimulus/distractor processing, memory maintenance) reflected by the phasic pupil response. Although tonic pupil activity predicted object memory precision in the present study, further investigation is needed to validate this measure of cognitive control using different paradigms, such as those involving continuous recall. In addition, phasic pupil responses reflected the effect of dlPFC TMS on memory maintenance, suggesting a reduction in cognitive effort when dlPFC activity was perturbed. Therefore, to capture the real-time effects of stimulation on cognitive processes, future research may benefit from combining neural stimulation with pupillometry to study distractor mitigation in WM. For instance, to measure the immediate effects of TMS stimulation on cognitive control, future studies may use online TMS protocols (when TMS pulses are applied within each trial at a certain phase) while tracking phasic pupil responses to the task.

In conclusion, this study expands limited evidence of the causal role of dlPFC in mitigating visual distractors in WM and uses naturalistic objects as memory targets with continuous recall to provide a more detailed view of the distraction impact on the memorandum. A deeper understanding of the mechanisms by which dlPFC implements cognitive control will improve the development of targeted therapeutic interventions – neurostimulation or/and cognitive training – for neurological conditions characterised by deficits in WM and distractibility, such as attention deficit hyperactivity disorder (ADHD), schizophrenia, and for the forms of dementia (e.g. Alzheimer's disease) where the function of prefrontal control mechanisms is impaired. Assessing individual differences in cognitive control showed that the relationship between mind wandering, worry, and distractor mitigation is complex. Worry appears to mask deficits in control through compensatory mechanisms, while spontaneous mind wandering can be more distracting than external stimuli, paradoxically helping some individuals stay focused on tasks. Therefore, future research must account for individual differences in cognitive control, as these variations may manifest distinctly during WM tasks and uncover important underlying mechanisms that contribute to cognitive function. Additionally, pupillometry can be a useful tool for future studies exploring cognitive control and allocation of cognitive resources during distraction in WM.

References

Allen, R. J., Baddeley, A. D., & Hitch, G. J. (2017). Executive and perceptual distraction in visual working memory. Journal of Experimental Psychology: Human Perception and Performance, 43(9), 1677.

Alnaes, D., Sneve, M. H., Espeseth, T., Endestad, T., van de Pavert, S. H.P., & Laeng, B. (2014). Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the locus coeruleus. Journal of Vision, 14, 1 20.

Assem, M., Glasser, M. F., Van Essen, D. C., & Duncan, J. (2020). A domain general cognitive core defined in multimodally parcellated human cortex. Cerebral Cortex, 30(8), 4361 4380.

Aston Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus–norepinephrine function: Adaptive gain and optimal performance. Annual Review of Neuroscience, 28, 403 450.

Axelrod, V., Rees, G., Lavidor, M., & Bar, M. (2015). Increasing propensity to mind-wander with transcranial direct current stimulation. Proceedings of the National Academy of Sciences, 112(11), 3314-3319.

Baldo, J. V., & Shimamura, A. P. (2000). Spatial and color working memory in patients with lateral prefrontal cortex lesions. Psychobiology, 28(2), 156 167.

Basten, U., Stelzel, C., & Fiebach, C. J. (2011). Trait anxiety modulates the neural efficiency of inhibitory control. Journal of Cognitive Neuroscience, 23(10), 3132 3145.

Basten, U., Stelzel, C., & Fiebach, C. J. (2012). Trait anxiety and the neural efficiency of manipulation in working memory. Cognitive, Affective, & Behavioral Neuroscience, 12, 571-588.

Beatty, J. (1982). Task evoked pupillary responses, processing load, and the structure of processing resources. Psychological Bulletin, 91(2), 276 292.

Berggren, N., & Derakshan, N. (2013). Attentional control deficits in trait anxiety: Why you see them and why you don't. Biological Psychology, 92(3), 440 446.

Berry, A. S., Zanto, T. P., Rutman, A. M., Clapp, W. C., & Gazzaley, A. (2009). Practice related improvement in working memory is modulated by changes in processing external interference. Journal of Neurophysiology, 102(3), 1779 1789.

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

Bruning, A. L., Mallya, M. M., & Lewis Peacock, J. A. (2023). Rumination burdens the updating of working memory. Attention, Perception, & Psychophysics, 85, 1452–1460.

Carriere, J. S. A., Seli, P., & Smilek, D. (2013). Wandering in both mind and body: Individual differences in mind wandering and inattention predict fidgeting. Canadian Journal of Experimental Psychology, 67(1), 19 31.

Chao, L. L., & Knight, R. T. (1995). Human prefrontal lesions increase distractibility to irrelevant sensory inputs. Neuroreport, 6(12), 1605 1610.

Chao, L. L., & Knight, R. T. (1998). Contribution of human prefrontal cortex to delay performance. Journal of Cognitive Neuroscience, 10(2), 167 177.

Chen, L., Thomas, E. H., Kaewpijit, P., Miljevic, A., Hughes, R., Hahn, L., & Fitzgerald, P. B. (2021). Accelerated theta burst stimulation for the treatment of depression: a randomised controlled trial. Brain Stimulation, 14(5), 1095 1105.

Christoff, K., Gordon, A. M., Smallwood, J., Smith, R., & Schooler, J. W. (2009). Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. Proceedings of the National Academy of Sciences, 106(21), 8719 8724.

Clapp, W. C., Rubens, M. T., & Gazzaley, A. (2010). Mechanisms of working memory disruption by external interference. Cerebral Cortex, 20(4), 859 872.

Crosswell, A. D., Coccia, M., & Epel, E. S. (2020). Mind wandering and stress: When you don’t like the present moment. Emotion, 20(3), 403.

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. Behavior Research Methods, 47(1), 1 12.

Dolcos, F., Miller, B., Kragel, P., Jha, A., & McCarthy, G. (2007). Regional brain differences in the effect of distraction during the delay interval of a working memory task. Brain Research, 1152, 171 181.

Duncan, J. (2010). The multiple demand (MD) system of the primate brain: mental programs for intelligent behaviour. Trends in Cognitive Sciences, 14(4), 172 179.

Esslinger, C., Schuler, N., Sauer, C., Gass, D., Mier, D., Braun, U., ... & Meyer Lindenberg, A. (2014). Induction and quantification of prefrontal cortical network plasticity using 5 Hz rTMS and fMRI. Human Brain Mapping, 35, 140–151.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. Emotion, 7, 336–353.

Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. Proceedings of the National Academy of Sciences, 110(41), 16616 16621.

Feredoes, E., Heinen, K., Weiskopf, N., Ruff, C., & Driver, J. (2011). Causal evidence for frontal involvement in memory target maintenance by posterior brain areas during distracter interference of visual working memory. Proceedings of the National Academy of Sciences, 108(42), 17510 17515.

Fitzgerald, P. B., Brown, T. L., & Daskalakis, Z. J. (2002). The application of transcranial magnetic stimulation in psychiatry and neurosciences research. Acta Psychiatrica Scandinavica, 105(5), 324 340.

Galeano Keiner, E. M., Pakzad, S., Brod, G., & Bunge, S. A. (2023). Examining the role of attentional allocation in working memory precision with pupillometry in children and adults. Journal of Experimental Child Psychology, 231, 105655.

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

Grueschow, M., Kleim, B., & Ruff, C. C. (2020). Role of the locus coeruleus arousal system in cognitive control. Journal of Neuroendocrinology, 32(12), e12890.

Grueschow, M., Kleim, B., & Ruff, C. C. (2022). Functional coupling of the locus coeruleus is linked to successful cognitive control. Brain Sciences, 12(3), 305.

Gustavson, D. E., & Miyake, A. (2016). Trait worry is associated with difficulties in working memory updating. Cognition and Emotion, 30(7), 1289 1303.

Holmes, N. P., Tamè, L., Beeching, P., Medford, M., Rakova, M., Stuart, A., & Zeni, S. (2019). Locating primary somatosensory cortex in human brain stimulation studies: experimental evidence. Journal of Neurophysiology, 121(1), 336 344.

Huang, Y. Z., Edwards, M. J., Rounis, E., Bhatia, K. P., & Rothwell, J. C. (2005). Theta burst stimulation of the human motor cortex. Neuron, 45(2), 201 206.

Jackson, J. B., Feredoes, E., Rich, A. N., Lindner, M., & Woolgar, A. (2021). Concurrent neuroimaging and neurostimulation reveals a causal role for dlPFC in coding of task relevant information. Communications Biology, 4(1), 588.

Jacob, S. N., & Nieder, A. (2014). Complementary roles for primate frontal and parietal cortex in guarding working memory from distractor stimuli. Neuron, 83(1), 226-237.

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox 3? Perception 36 ECVP Abstract Supplement.

Krimsky, M., Forster, D. E., Llabre, M. M., & Jha, A. P. (2017). The influence of time on task on mind wandering and visual working memory. Cognition, 169, 84–90.

Koevoet, D., Naber, M., Strauch, C., Somai, R. S., & Van der Stigchel, S. (2023). Differential aspects of attention predict the depth of visual working memory encoding: Evidence from pupillometry. Journal of Vision, 23(6), 9.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. Journal of Experimental Psychology: Human Perception and Performance, 21(3), 451.

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. Perception & Psychophysics, 56, 183 197.

Lorenc, E. S., Mallett, R., & Lewis Peacock, J. A. (2021). Distraction in visual working memory: Resistance is not futile. Trends in Cognitive Sciences, 25(3), 228 239.

Lorenc, E. S., Sreenivasan, K. K., Nee, D. E., Vandenbroucke, A. R., & D'Esposito, M. (2018). Flexible coding of visual working memory representations during distraction. Journal of Neuroscience, 38(23), 5267 5276.

Lowe, C. J., Manocchio, F., Safati, A. B., & Hall, P. A. (2018). The effects of theta burst stimulation (TBS) targeting the prefrontal cortex on executive functioning: A systematic review and meta analysis. Neuropsychologia, 111, 344 359.

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. Nature Neuroscience, 17(3), 347 356.Mallett, R., & Lewis Peacock, J. A. (2019). Working memory prioritization impacts neural recovery from distraction. Cortex, 121, 225 238.

Meyer, T., Miller, M., Metzger, R., & Borkovec, T. (1990). Development and validation of the Penn State Worry Questionnaire. Behavior Research and Therapy, 28, 487–495.

Mylius, V., Ayache, S. S., Ahdab, R., Farhat, W. H., Zouari, H. G., Belke, M., ... & Lefaucheur, J. P. (2013). Definition of DLPFC and M1 according to anatomical landmarks for navigated brain stimulation: inter rater reliability, accuracy, and influence of gender and age. Neuroimage, 78, 224 232.

Naber, M., & Nakayama, K. (2013). Pupil responses to high level image content. Journal of Vision, 13(6), 7.

Nemes, V. A., Parry, N. R., Whitaker, D., & McKeefry, D. J. (2012). The retention and disruption of color information in human short term visual memory. Journal of Vision, 12(1), 26 26.

Ngetich, R., Zhou, J., Zhang, J., Jin, Z., & Li, L. (2020). Assessing the effects of continuous theta burst stimulation over the dorsolateral prefrontal cortex on human cognition: a systematic review. Frontiers in Integrative Neuroscience, 14, 35.

Nieuwenhuis, S., De Geus, E. J., & Aston Jones, G. (2011). The anatomical and functional relationship between the P3 and autonomic components of the orienting response. Psychophysiology, 48(2), 162–175.

Oliveri, M., Turriziani, P., Carlesimo, G. A., Koch, G., Tomaiuolo, F., Panella, M., & Caltagirone, C. (2001). Parieto frontal interactions in visual object and visual spatial working memory: Evidence from transcranial magnetic stimulation. Cerebral Cortex, 11, 606–618.

Parthasarathy, A., Tang, C., Herikstad, R., Cheong, L. F., Yen, S. C., & Libedinsky, C. (2019). Time-invariant working memory representations in the presence of code-morphing in the lateral prefrontal cortex. Nature communications, 10(1), 4995.

Poh, J. H., Chong, P. L. H., & Chee, M. W. L. (2016). Sleepless night, restless mind: Effects of sleep deprivation on mind wandering. Journal of Experimental Psychology: General, 145, 1312–1318.

Postle, B. R., Ferrarelli, F., Hamidi, M., Feredoes, E., Massimini, M., Peterson, M., Alexander, A., & Tononi, G. (2006). Repetitive transcranial magnetic stimulation dissociates working memory manipulation from retention functions in the prefrontal, but not posterior parietal, cortex. Journal of Cognitive Neuroscience, 18, 1712–1722.

Rademaker, R. L., Bloem, I. M., De Weerd, P., & Sack, A. T. (2015). The impact of interference on short term memory for visual orientation. Journal of Experimental Psychology: Human Perception and Performance, 41(6), 1650.

Rademaker, R. L., Chunharas, C., & Serences, J. T. (2019). Coexisting representations of sensory and mnemonic information in human visual cortex. Nature Neuroscience, 22(8), 1336 1344.

Robison, M. K., & Unsworth, N. (2018). Cognitive and contextual correlates of spontaneous and deliberate mind wandering. Journal of Experimental Psychology: Learning, Memory, and Cognition, 44(1), 85.

Robison, M. K., Gath, K. I., & Unsworth, N. (2017). The neurotic wandering mind: An individual differences investigation of neuroticism, mind wandering, and executive control. Quarterly Journal of Experimental Psychology, 70(4), 649 663.

Rossini, D., Lucca, A., Zanardi, R., Magri, L., & Smeraldi, E. (2005). Transcranial magnetic stimulation in treatment resistant depressed patients: A double blind, placebo controlled trial. Psychiatry Research, 137(1 2), 1 10.

Sagliano, L., Atripaldi, D., De Vita, D., D'Olimpio, F., & Trojano, L. (2019). Non invasive brain stimulation in generalized anxiety disorder: A systematic review. Progress in Neuro Psychopharmacology and Biological Psychiatry, 93, 31 38.

Sari, B. A., Koster, E. H., & Derakshan, N. (2017). The effects of active worrying on working memory capacity. Cognition and Emotion, 31(5), 995 1003.

Suzuki, M., & Gottlieb, J. (2013). Distinct neural mechanisms of distractor suppression in the frontal and parietal lobe. Nature neuroscience, 16(1), 98-104.

Schicktanz, N., Fastenrath, M., Milnik, A., Spalek, K., Auschra, B., Nyffeler, T., ... & Schwegler, K. (2015). Continuous theta burst stimulation over the left dorsolateral prefrontal cortex decreases medium load working memory performance in healthy humans. PloS One, 10(3), e0120640.

Seli, P., Carriere, J. S., & Smilek, D. (2015). Not all mind wandering is created equal: Dissociating deliberate from spontaneous mind wandering. Psychological Research, 79, 750 758.

Seli, P., Schacter, D. L., Risko, E. F., & Smilek, D. (2019). Increasing participant motivation reduces rates of intentional and unintentional mind wandering. Psychological Research, 83(5), 1057 1069.

Soemer, A. (2019). Task unrelated thoughts and forgetting in working memory. Journal of Memory and Language, 106, 118 134.

Stojanoski, B., Emrich, S. M., & Cusack, R. (2019). Representation of semantic information in ventral areas during encoding is associated with improved visual short term memory. BioRxiv, 2019 12.

Stojanoski, B., & Cusack, R. (2014). Time to wave good bye to phase scrambling: Creating controlled scrambled images using diffeomorphic transformations. Journal of Vision, 14(12), 6 6.

Teng, C., & Kravitz, D. J. (2019). Visual working memory directly alters perception. Nature Human Behaviour, 3(8), 827 836.

Unsworth, N., & Robison, M. K. (2015). Individual differences in the allocation of attention to items in working memory: Evidence from pupillometry. Psychonomic Bulletin & Review, 22, 757 765.

Unsworth, N., & Robison, M. K. (2016). The influence of lapses of attention on working memory capacity. Memory & Cognition, 44, 188 196.

Unsworth, N., & Robison, M. K. (2018). Tracking working memory maintenance with pupillometry. Attention, Perception, & Psychophysics, 80, 461 484.

Usher, M., Cohen, J. D., Servan Schreiber, D., Rajkowski, J., & Aston Jones, G. (1999). The role of locus coeruleus in the regulation of cognitive performance. Science, 283, 549 554.

Van der Wel, P., & Van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. Psychonomic Bulletin & Review, 25, 2005 2015.

Vannucci, M., Chiorri, C., Nocentini, A., & Menesini, E. (2020). Distinguishing spontaneous from deliberate mind wandering in adolescents: The role of attentional control and depressive symptoms. British Journal of Developmental Psychology, 38(3), 434 441.

Vékony, T., Németh, V. L., Holczer, A., Kocsis, K., Kincses, Z. T., Vécsei, L., & Must, A. (2018). Continuous theta burst stimulation over the dorsolateral prefrontal cortex inhibits improvement on a working memory task. Scientific Reports, 8(1), 14835.

Veldsman, M., Mitchell, D. J., & Cusack, R. (2017). The neural basis of precise visual short term memory for complex recognizable objects. NeuroImage, 159, 131 145.

Voytek, B., & Knight, R. T. (2010). Prefrontal cortex and basal ganglia contributions to visual working memory. Proceedings of the National Academy of Sciences, 107(42), 18167 18172.

Weissman, D. H., Roberts, K. C., Visscher, K. M., & Woldorff, M. G. (2006). The neural bases of momentary lapses in attention. Nature Neuroscience, 9(7), 971 978.

Wischnewski, M., & Schutter, D. J. (2015). Efficacy and time course of theta burst stimulation in healthy humans. Brain Stimulation, 8(4), 685 692.

Yoon, J. H., Curtis, C. E., & D'Esposito, M. (2006). Differential effects of distraction during working memory on delay period activity in the prefrontal cortex and the visual association cortex. Neuroimage, 29(4), 1117 1126.

Zokaei, N., Board, A. G., Manohar, S. G., & Nobre, A. C. (2019). Modulation of the pupillary response by the content of visual working memory. Proceedings of the National Academy of Sciences, 116(45), 22802 22810.

Chapter 4. The Role of Multiple-Demand Network in Protecting Visual Working Memory Representations from Interruption

Abstract

The ability to protect working memory (WM) from the disruptive effect of irrelevant information is key for successful performance in everyday life. Yet, it is unclear what neural mechanisms implement this ability. Multiple studies show that effortful cognitive tasks activate the frontoparietal regions collectively called the Multiple-Demand (MD) network. In the present study, we asked if the MD network protects memory representations when a demanding secondary task appears during memory maintenance. In the dual-task paradigm, participants had to memorise an image of a tool or an animal over a short delay period (*memory task*). During this period, an unrelated easy versus hard digit manipulation task (*digit task*) was presented. To evaluate how shifting attention to the digit task impacts memorised information, we varied the difficulty of the digit task. To investigate the role of the MD network in protecting memory representations from digit task interruption, we used functional magnetic resonance imaging (fMRI) and multivariate pattern analysis (MVPA) to evaluate memory decoding in MD regions and the early visual cortex. As expected, the harder version of the digit task resulted in decreased memory performance (lower accuracy and slower reaction times). Neural decoding in the MD network showed increased decoding accuracy of memorandum when the digit task was harder. Analysis of the memory decoding in separate MD regions revealed that the intraparietal sulcus (IPS) subregion showed the strongest effect. Memory representations in the early visual cortex remained unaffected by the digit task demands. The results point to the central role of the IPS in protecting memory representations in the face of increasing task demands, critically, by enhancing the neural decoding of memorandum. We propose a mechanism through which the MD network and IPS in particular, protects WM from interruptions.

Introduction

Everyday life tasks rely on our ability to hold information in mind referred to as working memory (WM) (Baddeley, 1992). However, irrelevant information such as distractors (irrelevant stimuli in our environment) or interrupters (events or tasks that temporarily halt or disrupt ongoing task) may interfere with this ability. Therefore, it is critical to protect WM from interference.

Recent studies investigated the impact of distractors (i.e. irrelevant stimuli displayed during memory maintenance) on WM using neural decoding of functional MRI data (Bettencourt & Xu, 2016; Lorenc et al., 2018; Rademaker et al., 2019; Hallenbeck et al., 2021; Vu-Cheung et al., 2021; Kiyonaga et al., 2017). This approach provides unique insights into the fate of memory representations in the face of distraction. Specifically, the interfering effect of distractors is reflected in the drop of memory decoding in the presence of distractors; potentially, because the same neural populations encode both memorandum and distractors leading to noisier memory representations (Hallenbeck et al., 2021).

Using this approach, several mechanisms for WM protection have been proposed. One line of research highlights the importance of the parietal cortex in memory maintenance and specifically its role in the protection of memoranda from visual distractors (Bettencourt & Xu, 2016; Lorenc et al., 2018; Rademaker et al., 2019). Bettencourt & Xu (2016) decoded memory representations in the occipital cortex (V1, V2, V3, V4, V3A, V3B) and superior parts of the intraparietal cortex (sIPS) in the absence of distraction. However, visual distraction (i.e. flickering images of faces or gazebos) impacted orientation decoding in the occipital cortices, but not in the sIPS suggesting that this region stores memory representations protecting them from upcoming sensory information.

Although Lorenc et al., (2018) also highlight the role of the sIPS in mitigating visual distractors, the authors emphasize the adaptive nature of representations in this region by showing that reliable decoding of orientations only occurred when distracting stimuli (i.e., orientations) were present during the delay period, and not when the delay period was unfilled. Researchers attribute this to the flexible coding of memory representations in the sIPS based on task demands and expectations:

when distractors are likely, the brain adaptively encodes the memorandum not only in the occipital cortices but also in the sIPS (Lorenc et al., 2018).

The role of the IPS in WM protection from distracting information was further investigated by Rademaker et al. (2019); the study found a change in the representational format of memorised orientations in the face of visual distractors (orientations, faces, or gazebos) during the delay period. Specifically, the findings demonstrated that memory representations shifted from sensory to mnemonic format in IPS to resist interference, while those in the visual cortex remained in a more distortion-prone sensory format.

In contrast, another line of research provides evidence for memory representations within the visual cortex being sufficiently resilient to distractors by demonstrating unchanged decoding in the face of distraction (Hallenbeck et al., 2021; Vu-Cheung et al., 2021). Visual polar angle was reliably decoded in the sensory cortex as well as in parietal and frontal regions after a distracting stimulus (black and white moving dots) was shown during the delay period, suggesting the robust nature of memory representations stored across the cortex (Hallenbeck et al., 2021). Similarly, Vu-Cheung et al. (2021) found no difference in the decoding accuracy of the spatial location of the memorised coloured dot in retinotopic areas with and without an interrupter (radial checkerboard) presented in the delay period.

Despite using different behavioural and decoding methods to assess the impact of distractors on WM and holding contrasting views on the role of sensory areas in memory protection, these studies consistently emphasise the importance of memory storage in the posterior cortex. This suggests that the posterior cortex plays a crucial role in protecting memory, potentially without relying on top-down cognitive control (Hallenbeck et al., 2021).

Cognitive control may be especially important when dealing with interrupters that require active cognitive effort and cannot be passively ignored. Yet, few decoding studies addressed the impact of interruptions on memory representations (Kiyonaga et al., 2017). Researchers asked participants to memorise either one or two items (faces or houses) for a later memory probe. During memory maintenance, participants performed a visual search task, acting as an interrupter, where they had to identify a vertical target among distractors (tools or bodies) that were either easy or hard to differentiate from the target. The study found that decoding accuracy in sensory regions was influenced by the difficulty of the visual search, suggesting that cognitive resources are shared between maintaining memory representations and processing external stimuli (Kiyonaga et al., 2017). Univariate fMRI analysis revealed an interaction effect in the prefrontal cortex (PFC) when both WM load and visual search difficulty were manipulated. Specifically, the PFC showed increased activity when both the WM load was high and the visual search task was challenging (Kiyonaga et al., 2017). This indicates that the PFC was particularly engaged when there were high demands on managing cognitive control over both the memory task and the interruption, suggesting that this adaptable mechanism may be key to resolving interference from interruptions in WM. However, the PFC does not operate in isolation; it likely works in concert with other neural systems to implement cognitive control (Menon & D'Esposito, 2022).

One of the candidates for a network implementing such a control over memorandum could be a set of regions in the parietal and frontal cortex, referred to as the Multiple Demand (MD) network. It is proposed to play a crucial role in managing attention by dividing complex tasks into focused sub-tasks, allowing for efficient problem-solving and goal-directed behaviour (Duncan, 2010). Namely, the regions identified as parts of this network are anterior inferior frontal sulcus (aIFS), posterior inferior frontal sulcus (pIFS), premotor cortex (PM), inferior frontal junction (IFJ), anterior insula/frontal operculum (AI/FO), intraparietal sulcus (IPS), and bilateral pre-supplementary motor area/anterior cingulate cortex (pre-SMA/ACC). MD regions engage simultaneously during various demanding tasks, such as those involving WM, selective attention, and problem-solving (Fedorenko et al., 2013; Assem et al., 2020; Woolgar et al., 2011; Woolgar et al., 2016). MD network might achieve this by dynamically adapting neural activity to the current task demands, selectively coding relevant information while ignoring irrelevant details, as evidenced by recent studies using a decoding approach (Jackson & Woolgar, 2018; Jackson et al., 2017). In these experiments participants saw the same abstract novel objects and had to attend to one of the object features in the presence of another; neural decoding revealed that the MD network coded only relevant information, whereas early visual areas coded both relevant and irrelevant features (Jackson & Woolgar, 2018; Jackson et al., 2017). This representational mechanism might protect the relevant feature from interference by the irrelevant feature. Could a similar mechanism operate in the WM to protect it from external irrelevant information?

Since there is evidence for some of the MD regions to be involved in WM protection, it is plausible that the MD network may orchestrate this ability. For example, IPS is part of the MD network, and the decoding studies discussed above suggest this region protects WM by maintaining distractor-resistant memoranda (Bettencourt & Xu, 2016; Lorenc et al., 2018; Rademaker et al., 2019). The role of the prefrontal MD regions in resolving distraction has been demonstrated by studies using univariate analysis (Dolcos et al., 2007) paired with a functional connectivity approach (Yoon et al., 2006; Clapp et al., 2010). Specifically, Dolcos et al. (2007) asked participants to memorise faces while being presented with confusable (faces) or non-confusable distracters (scrambled faces) during the delay period; researchers found that sustained activity within the dorsolateral prefrontal cortex (dlPFC) reflected memory maintenance and preparation, especially in the presence of high-confusion distracters; whereas inferior frontal gyrus (IFG) showed increased activation linked to resolving the impact of confusable distractors. Yoon et al. (2006) also asked participants to hold images of faces in mind while displaying a congruent (face) or incongruent (scene) distracting image during the delay period and measured the functional connectivity between the dlPFC and visual association cortex (VAC); the results revealed that congruent distractors disrupted connectivity between dlPFC and VAC, suggesting that dlPFC elicits top-down control to maintain memorandum in sensory areas. Using a similar connectivity approach, Clapp et al. (2010) further investigated the different effects of distraction and interruption on the connectivity between dlPFC and VAC. Specifically, while participants held an image of a face in mind, a distracting face that could be ignored (distraction) was presented during the delay period and did not impact the connectivity between these regions;

however, when participants had to respond to a face image presented during the delay period (interruption), the connectivity between dlPFC and VAC was disrupted and then re-instantiated after the interruption (Clapp et al., 2010). This finding indicates that while interruptions temporarily disrupt the maintenance of memoranda, the dlPFC can re-engage with sensory regions to restore WM performance. This role of the dlPFC is further supported by the causal evidence provided by Feredoes et al., (2011). In their experiment researchers applied transcranial magnetic stimulation (TMS) over the dlPFC to modulate activity in this region while simultaneously acquiring functional MRI; this approach provides a unique insight into the effect of a region perturbation on the connected regions and their activity. The study found that sensory activity related to the memorandum was enhanced but only when stimulation was applied over the dlPFC and in the presence of visual distractors (Feredoes et al., 2011). Together, using different techniques, prior research demonstrates the involvement of MD regions such as IPS, IFG and dlPFC (represented by two MD subregions – anterior and posterior inferior frontal gyrus, IFG) in resolving interference in WM. However, it remains unclear whether these regions operate in a coordinated manner to mediate this process.

To test if the whole MD system is involved in WM protection, we used neural decoding to investigate the fate of memory representations when an interrupter (i.e. secondary task) appeared during memory maintenance in the MD regions combined, as well as separate. The interrupter aimed to increase the task demand and the degree of control required to protect memorandum from distortion. First, we asked participants to memorise an image of an animal or tool and hold it in mind for the delay period (referred to as *memory task*). Then, during the memory maintenance phase, an interrupter task, referred to as *digit task*, appeared on the screen. This task involved sorting a list of four digits in ascending order and then identifying the second digit from the sorted list. To manipulate the degree of attentional modulation over a memorandum and reduce the role of perceptual interference, we kept the visual display constant but varied the difficulty of the digit task: the list could be hard (active hard) or easy to sort (active easy), or could be ignored (passive).

We hypothesised that if the MD network elicits cognitive control to protect memoranda during the digit task, it will preserve the memory representations in the face of the interruption since the memory task remained relevant throughout the trial and had to be kept in mind. Alternatively, if the MD system exerts control only over the task at hand and adjusts its responses to the immediately relevant information, it will switch to coding the digit task leading to a decrease in memory decoding with the higher difficulty of the digit task resulting in lower memory decoding.

Methods

Procedure

Participants

Twenty-eight participants (18 females; mean age=26.1, SD= 7.9) took part in the study. All participants had normal or corrected-to-normal vision and no history of neurological or psychiatric disorders. Participants gave written informed consent and were reimbursed for their time. The experiment was approved by the University of Reading Research Ethics Committee.

Data from two participants was excluded from the analysis due to low performance in the memory task (<60% recall accuracy). Since excessive movements may compromise the integrity of the fMRI analysis (Power et al., 2012; Van Dijk et al., 2012), after applying a 2 mm threshold for translation or rotation of the movement parameters, another four datasets were excluded after the realignment preprocessing step. Specifically, two participants had excessive movements in 3 runs, and another two – in 4 runs. In addition, after analysis of decoding accuracies, one subject appeared as an outlier revealed by the Interquartile Range (IQR) method (i.e. the decoding accuracy in the combined MD ROI was 82%). The final sample resulted in twenty-one individual data sets.

Memory Task Stimuli

The stimuli set for the memory task consisted of black-and-white morphed images of animals and tools. All images were generated using the diffeomorphic transformation algorithm (Cusack & Stojanoski, 2014), with an equal amount of transformation between each image. Each stimulus set for the training and for the main experiment consisted of one pair of animals and one pair of tools; three morphed exemplars were generated for each image. Figure 1, panel A depicts all the stimuli used as memory targets in the main experiment. The aim of this stimulus set was to make the memory task more difficult, because participants had to memorise not just an object (i.e. identity), but also its fine details to choose the correct item at recall.

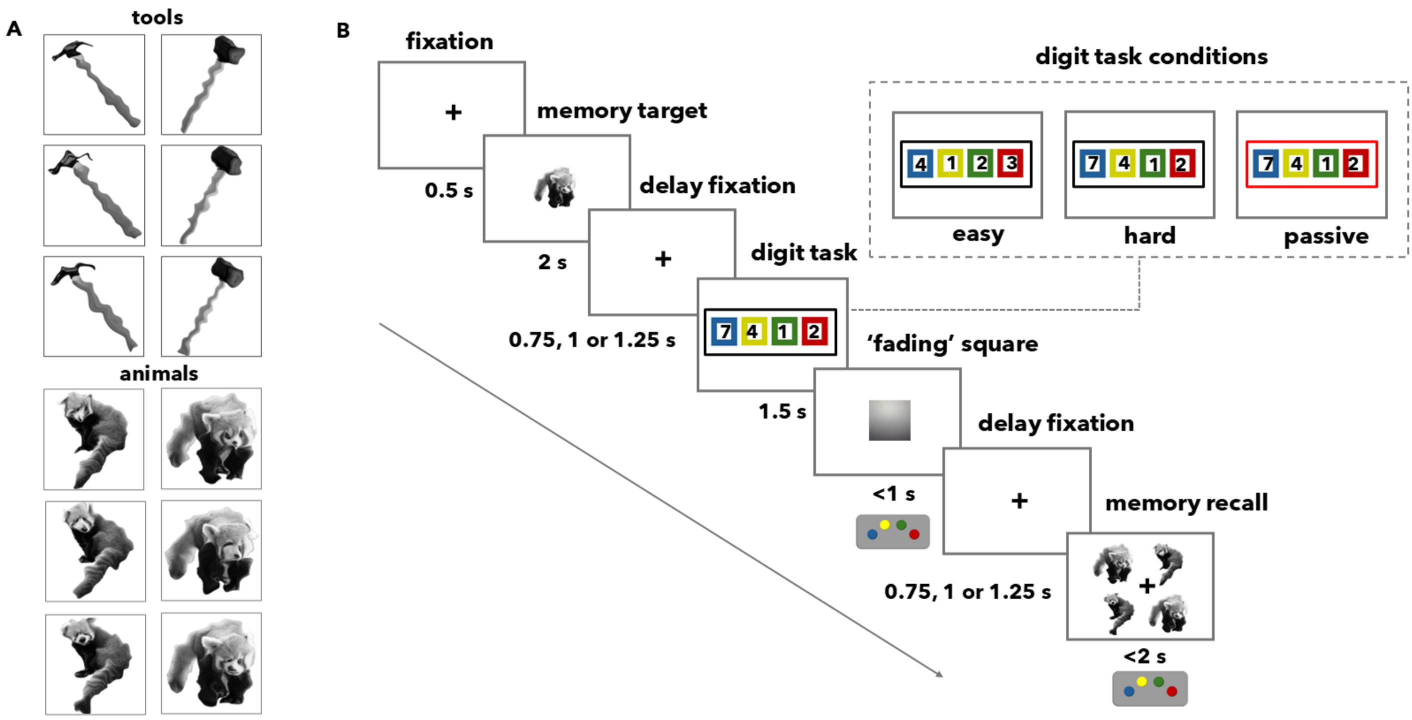


Figure 1. Stimulus set for the memory task (A) and Dual Task Paradigm (B). (A) The memory task stimulus set included two categories–tools and animals–each with 3 exemplars to increase task difficulty. (B) The dual-task paradigm involved a memory task and a digit task, with the latter serving as an interrupter during memory maintenance. Participants saw an image (tool or animal) to remember, followed by a fixation cross. Next, four coloured squares with digits appeared, and participants had to sort the digits and press the button for the second item in the list before a 'fading square' disappeared. Task difficulty was manipulated by frame colour: a black frame indicated sorting, while a red frame signalled a passive condition where participants ignored the digits and pressed a button when the 'fading square' appeared. After a delay, participants recalled the initial image by selecting the memorized image.

Dual task

As displayed in Figure 1, panel B, on each trial, participants were asked to perform a memory task and a digit task (as an interrupter). A memory task trial started with a fixation cross (0.7° ×0.7°) displayed for 0.5s, followed by the memory target: an image (7° × 7°) of a morphed object (tool or animal) which was to be recalled after a delay period.

After the memory target, a fixation cross (0.7° ×0.7°) was displayed on the screen for 0.75, 1, or 1.25 seconds. Then four squares (3°× 3° each, 1.5° width apart from each other) were presented centrally, each filled with one of the following colours, in this order, from left to right: blue, yellow, green, and red. The order of the colours was always the same and corresponded to the colours of the response button box. Within each square, a black digit was placed against a white background. The squares were presented for 1.5 seconds and were followed by a gray "fading" square (3°× 3°), gradually changing colour from gray to white to denote the time passing, displayed up to 1 second. Participants were instructed to mentally arrange the digits in ascending order and respond with the second digit in the ordered sequence when the "fading" square appeared.

To respond to the digit task, participants used a button box with four coloured buttons (blue, yellow, green, and red), which matched the colours of squares displayed on the screen with digits. Participants were required to press the button that matched the colour of the square associated with the digit they needed to respond to (the second item in a sorted list). This is referred to as the digit task.

The digit task included three variations: two active types and one passive type. Two active digit tasks represented two levels of difficulty - easy and hard. In the easy digit task, the digits were randomly sampled from the interval of 1 to 4. Whereas in the hard digit task, the sample interval was from 1 to 9. In this scenario, arranging a sequence of digits where each digit differs by more than one from the next would be more challenging within the restricted timeframe (hard digit task), compared to organizing a sequence of digits where each one differs by exactly one from the next (easy digit task). The passive digit task served as a control and required only pressing the red button when the fading square appeared; participants were asked not to order the digits. In the passive task, the digits were specifically sampled the same way as for the hard digit task, so participants would not order the digits unintentionally.

The type of the digit task was cued by the frame (thickness: 0.25°) around the squares with digits; the black frame informed about the active digit task – both hard and easy, and the red frame appeared for the passive digit task. Then, followed by another fixation cross displayed on the screen (0.7° ×0.7°) for 0.75, 1, or 1.25 seconds, participants were prompted to recall the image they had memorised. To test memory recall accuracy for the memorandum, four images were presented at the memory probe stage: one target image, one that was a different version of the same image, and two versions of a visually similar paired image (see Figure 1B). In a 4 AFC paradigm, participants selected the memorised image (7° ×7°) using the same button box, with each button mapped to the location of an image. After selecting an image, a 0.3° thick black frame appeared around it for 50 ms, followed by a blank screen for an inter-trial interval (jittered duration from 2 to 10 seconds).

Each participant underwent 4-5 training runs (9 trials each, lasting ~2.5 minutes) to familiarize themselves with button assignments for memory recall and digit task responses, and a further 9 practice trials in the scanner before data collection. The number of digit task types was balanced in each training run.

The main experiment consisted of 8 runs, each run had 24 trials and lasted ~6.5 minutes. The number of each type of digit task was balanced in each run (8 trials per type). The category of the memory target was also balanced in each run (4 tools and 4 animals were randomly chosen for each digit task type). The main experiment resulted in 192 trials in total, with 64 trials per digit task type.

MR Data Acquisition

Participants were scanned on a 3T MRI scanner (Prisma Siemens) with a 32-channel head coil. During the scanning session, functional MRI scans were acquired with the following sequence parameters: 48 transversal slices collected in interleaved order with repetition time (TR) = 1825 ms, echo time (TE) = 30 ms, multiband acceleration factor =2, voxel size = 2.5 × 2.5× 2.5 mm3, field-of-view = 192 × 192 mm, matrix size = 64 × 64, slice thickness = 3.5 mm, 0 mm slice gap, flip angle = 67°) covering the entire cerebrum except cerebellum. For each participant we also acquired a T1-weighted MR image (MP-RAGE; 176 sagittal slices, TR = 2300 ms, TE = 2.29 ms, voxel size = 1 × 1 × 1 mm3, field-of-view = 240 × 240 mm, matrix size = 240 × 240, slice thickness = 0.9 mm, 0 mm slice gap, TI = 900 ms, flip angle = 8°).

Preprocessing

MRI data were pre-processed using SPM12 (Welcome Department of Imaging Neuroscience, www.fil.ion.ucl.ac.uk/spm) in MATLAB 2021a (MathWorks, Inc., 2019). Functional MRI data were converted from DICOM to NIFTII format, spatially realigned to the mean functional scan and slice timing corrected. EPIs from the main experiment were smoothed (4 mm FWHM Gaussian kernel) to improve the signal-to-noise ratio, as applied in previous similar work (Jackson et al., 2018; Jackson et al., 2017). Structural scans were co-registered to the mean EPI and spatially normalised, using the segmentation and normalisation routines of SPM12, to derive the (inverse) normalisation parameters for individual ROI definition.

ROIs definition

Similar to a recent study investigating relevant over irrelevant representations in the MD network (Jackson et al., 2024), we used the parcellated map of 13 frontal and parietal MD ROIs provided by Fedorenko et al. (2013; available online at imaging.mrc-cbu.cam.ac.uk/imaging/MDsystem). This definition of the MD network closely aligns with the previous definition by Duncan & Owen (2000), derived from meta-analytic data, and used in previous studies applying MVPA to MD regions (Jackson et al., 2017; Jackson & Woolgar, 2018; Jackson et al., 2021).

The MD ROIs included the left and right anterior inferior frontal sulcus (aIFS; center of mass (COM) MNI coordinates = ±35 47 19, volume = 5.0 cm^3), left and right posterior inferior frontal sulcus (pIFS; COM ±40 32 27, 5.7 cm^3), left and right premotor cortex (PM; COM ±28 −2 56, 9.0 cm^3), left and right inferior frontal junction (IFJ; COM ±44 4 32, 10.1 cm^3), left and right anterior insula/frontal operculum (AI/FO; COM ±34 19 2, 7.9 cm^3), left and right intraparietal sulcus (IPS; COM ±29 −56 46, 34.0 cm^3), and bilateral presupplementary motor area/anterior cingulate cortex (pre-SMA/ACC; COM 0 15 46, 18.6 cm^3). We defined the early visual cortex (BA17: COM −1 −79 6, 31 cm3) from the Brodmann template provided with MRICroN (Rorden, 2007). All templates were provided in the MNI space and were then transformed to the native space by applying the inverse of the normalisation parameters for each participant using SPM12's 'segment' and 'normalise' routine. For the MD network all 13 ROIs were combined into a single mask using 'ImCalc' in SPM12 (Wellcome Department of Imaging Neuroscience, www.fil.ion.ucl.ac.uk/spm).

General Linear Model (GLM)

We applied the General Linear Model (GLM) in SPM12 to estimate beta values for subsequent MVPA. For each run, we estimated separately the responses to each category (tool, animal) and digit task condition – easy, hard and passive. Two additional regressors representing "baseline" activity (tool, animal) during the inter-trial interval were also added to the model. This produced beta values for each run corresponding to 8 conditions. Trials were modelled as events as the delay period (i.e. starting after memory target onset until memory recall onset) and convolved with the canonical hemodynamic response of SPM12. In addition, baseline response (inter-trial interval), motion parameters, and grand means of each run were also modelled as separate nuisance regressors resulting in fifteen regressors in total.

Multivariate Pattern Analysis (MVPA)

To examine the fate of the memory representations in the presence of the (interrupter) digit task, we used MVPA to decode the memory object category (tools versus animals). We implemented MVPA using the CosmoMVPA (Oosterhof et al., 2016; https://www.cosmomvpa.org/).

For each participant, within each ROI (early visual cortex, combined MD ROI, within each MD sub-region) and for each digit task condition (easy, hard and passive) we trained a linear support vector machine (lSVM) to decode the category of memoranda.

Classification accuracies were calculated using a leave-one-run-out cross-validation technique, in which patterns generated by beta values from N-1 runs formed the training set, and the pattern from the remaining run was used as the test set; this procedure was iterated for each of the eight runs.

To test the significance of the decoding accuracies, we conducted a t-test against chance-level decoding (50%) for each ROI in each digit task condition. To evaluate the impact of the digit task on the memory decoding in each ROI, we performed an ANOVA with the digit task condition as a main factor, separately for each region; we also included an error term accounting for the within-subject variability. Pairwise comparisons were conducted using the Bonferroni correction to control for Type I errors. Statistical analysis of the decoding results was performed in R Statistical Software (v4.3.2; R Core Team 2023)

Behavioral Data Analysis

Behavioral responses were aggregated using a custom-written MATLAB script in Matlab 2021a (MathWorks, Inc., 2019) and analysed using R Statistical Software (v4.3.2; R Core Team 2023). To test the effect of the digit task on memory recall, we conducted paired t-tests to compare memory task performance (accuracy and reaction time) between each digit task combination: passive vs. hard, passive vs. easy, and easy vs. hard.

Although we ensured that each participant could perform the digit task with at least 60% accuracy prior to scanning, we also evaluated their performance during the experiment in the scanner. To ensure that participants attended to the digit task during memory maintenance, we calculated accuracy in each digit task condition and compared it against the chance level (25% for active easy and hard; 50% for passive).

We compared reaction times in the digit task using paired t-tests to confirm the task demand manipulation, aiming to create a gradient of interrupter task difficulty. The active hard task was intended to be the most difficult, the active easy task less difficult, and the passive task the simplest to perform.

Results

Performance on the digit task reflected successful task demand manipulation

First, we assessed the mean accuracy (i.e. per cent correct) in each digit task condition. Mean accuracy as displayed in Figure 2 (panel A) was the highest for the easy condition (M=94.18%, SD=10.59%), followed by the passive condition (M=90.8%, SD=20.35%), and further dropped for the hard digit task condition (M=88.05%, SD=16.42%). Such a high accuracy rate shows that participants were performing the task and diverting their focus to the interrupter as intended by the dual-task paradigm. Note that two participants showed low (below chance) performance, but we did not exclude their data since they performed well on the memory task and the digit task at 100 per cent accuracy during the training phase. Some responses were not recorded due to being outside of the time window, or a response not being provided on a trial.

The reaction times, as depicted in Figure 2 (panel B), reveal a gradient of the digit task difficulty. The hard digit task (M=0.38 sec, SD=0.17 sec) resulted in significantly slower responses in comparison with easy (M=0.34 sec, SD=0.15 sec; t(21)=4.35, p<.001, Cohen's d= .92) or passive condition (M=0.29 sec, SD=0.13 sec; t(21)=4.38, p<.001, Cohen's d= .93). Responses to the easy digit task were slower (t(21)=2.52, p=.01, Cohen's d= .53) than in the trials with passive condition. This result demonstrates that the manipulation of the digit task difficulty successfully worked in the present study.

Notably, two participants had low response times (RTs) as shown in Figure 2 (panel B), but their overall digit task accuracy was above chance (13% and 15%, with the chance level at 12.5%). After removing these participants, the pattern of statistical results remained unchanged. Therefore, we decided to keep the data from these participants in the dataset.

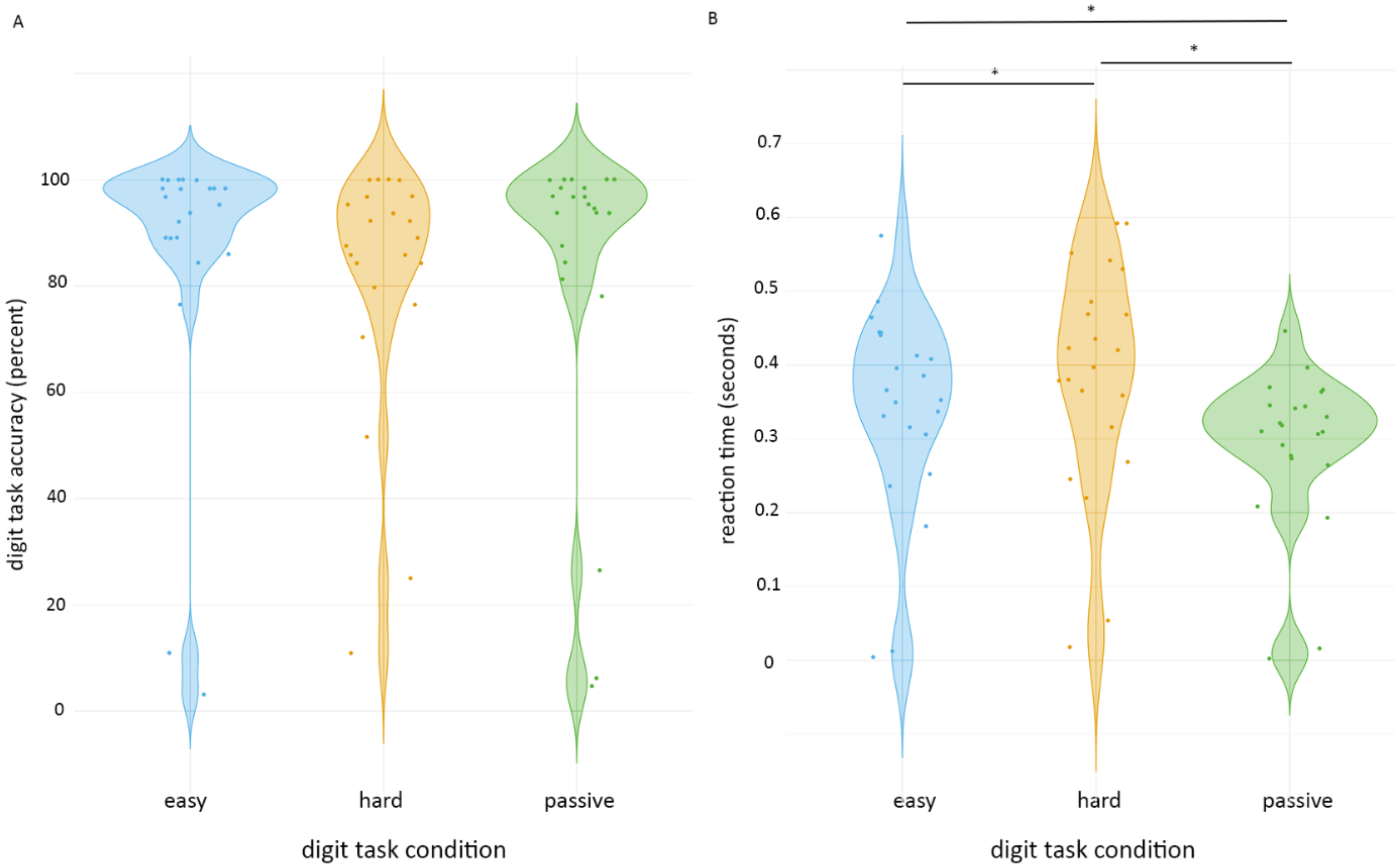


Figure 2. Accuracy and reaction times in the digit task conditions reflected successful task demand manipulation. First, participants were performing the digit task with high accuracy (A), except for several outliers, suggesting that the digit task was attended. The reaction times (B) of participants' responses reflected the gradient of difficulty intended by the task manipulation: responses were slower when the difficulty of the digit task was high (i.e. in the hard condition), then faster when the task was easier, and the fastest in the passive (control) condition.

Effect of Digit Task Difficulty on Memory Recall Accuracy and Reaction Times

Although there was no main effect of condition type on recall accuracy ($F(2, 63) = 1.32$, $p = .275$, $\eta^2=.04$), pairwise comparisons revealed that the difficulty of the digit task impacted the accuracy of recall in the memory task (see Figure 3, panel A). Specifically, memory accuracy dropped ($t(21) = -2.26$, $p=.034$, Cohen's $d=.48$) when the easy active digit task (M=80.82%, SD=10.28%) was performed during the memory delay period compared to the passive condition (M=85.11%, SD=8.99%). Similarly, compared to the passive condition, the hard active digit task (M=81.1%, SD=10.85%) significantly lowered (t(21) = -2.79, p=.01, Cohen's d=.59) the accuracy of memory recall. The difficulty of the active digit task did not significantly impact recall accuracy (t(21) =.16, p=.86).

As illustrated in Figure 3, panel B, participants' response to the memory task was slower (t(21)= 2.16, p= .042, Cohen's d=.46) when the hard digit task (M= 1.35, SD= .12) appeared at the delay period in comparison with the passive condition (M= 1.31, SD= .12). There were no differences in response times between trials with the easy (M= 1.33, SD= .09)) and hard or easy and passive digit tasks.

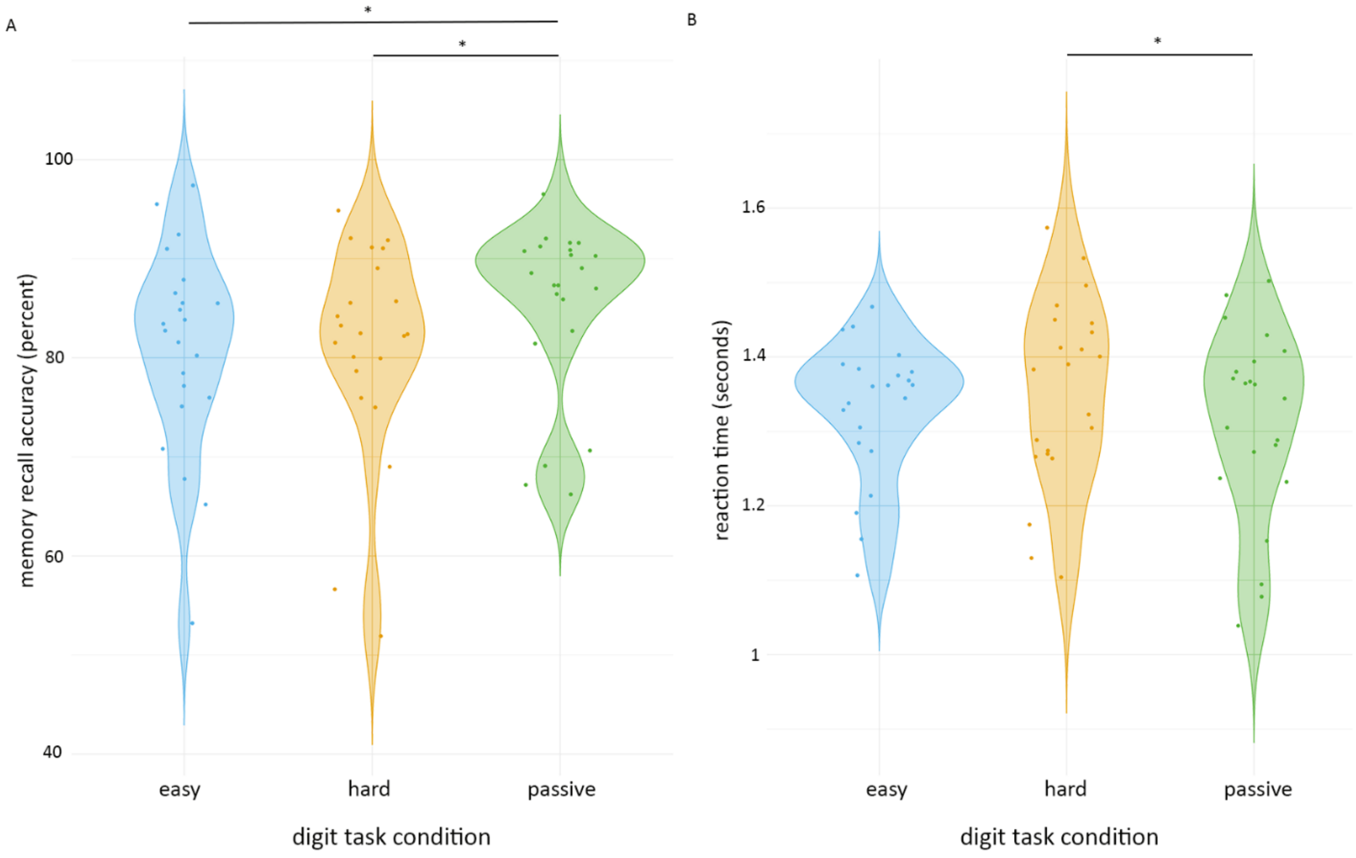


Figure 3. Effects of digit task difficulty on WM recall accuracy and response times. (A) Increasing difficulty of the digit task impacted the memory recall accuracy in the memory task. Specifically, both the hard and easy conditions resulted in lower recall accuracy in comparison with the passive digit task condition. The hard condition also resulted in slower responses (B) than in the passive condition of the digit task. *Denotes paired t-tests, p<0.05.

Decoding within MD Network

To address our first question of whether diverting attention to the digit task modulated the memory representations in the MD network, we compared the decoding of memory targets in the combined MD ROI in all digit task conditions. First, the decoding accuracy was statistically greater than chance in the active digit conditions – easy (M=55.1%, SD=8.09%, t(21)= 2.95, p= .007) and hard (M=59.36%, SD=9.59%, t(21)= 4.57, p=.0001); but not in the passive condition (M=52.35%, SD=6.32%, t(21)= 1.74, p=.09).

There was a statistically significant effect of the digit task difficulty on the decoding of memoranda in the MD network (F(2, 42) = 4.02, p = .025, η^2=.17). Pairwise comparisons revealed the significant difference between hard and passive conditions (t(42) = 3.22, p=.007); no significant differences were found in memory decoding in easy versus hard (t(42) = -1.958, p= .17) and easy versus passive (t(42) = 1.26, p= .63) conditions.

Eight MD regions showed significant memory decoding in one or more digit task conditions. Memory representations could be decoded (ts>2.3, ps<0.05) in the hard digit task condition in right and left IPS, right IFG, right and left pIFS, left and right PM and left AI/FO (See Figure 4 for the decoding results from all MD sub-regions). In the easy digit task condition, only right pIFS, right and left IPS significantly coded memoranda (ts>2.4, ps<0.05), whereas the decoding of memory representations in the passive condition was above chance in right IFJ and right IPS (ts >2, ps<0.05). Memory decoding was statistically non-significant (ts<1.9, ps>0.05) in all digit task conditions in five MD regions: ACC, right AI/FO, left and right aIFS, left IFG.

Interestingly, the impact of digit task difficulty was most evident in left IPS (Figures 4 & 5). Namely, there was the main effect of digit task condition on memory decoding (F(2, 42) = 14.39, p< .001, η^2=.42); performing the hard digit task led to an increase in the decoding accuracy in comparison to passive (t(42) = 5.248, p< .001) or to easy digit task condition (t(42)= -3.587, p=.0027). There was no significant difference in memory decoding between easy and passive digit task conditions in the left IPS (t(42)= 1.66, p=.31).

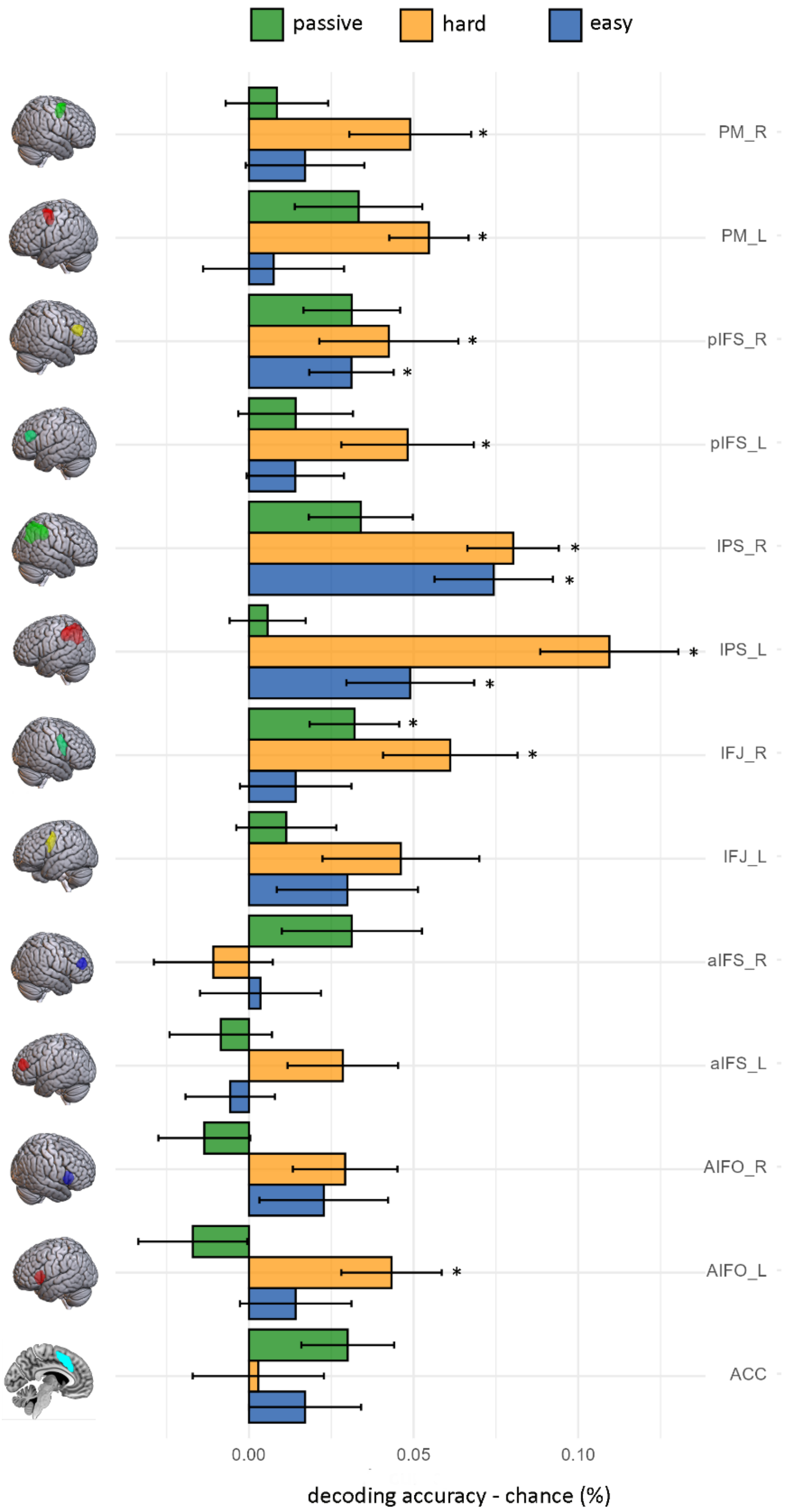


Figure 4. Decoding accuracy in each MD ROI in each digit task condition. Memory representations could be decoded (ps<0.05) in the hard digit task condition in right and left IPS, right IFJ, right and left pIFS, left and right PM and left AI/FO. In the easy digit task

condition only right pIFS, right and left IPS coded memoranda(ps<0.05). Decoding of memory representations in the passive condition was above chance only in right IFJ and right IPS(p<0.05). No significant memory decoding was found in the ACC, right AI/FO, left and right aIFS, left IFG (ps>0.05). ACC: bilateral presupplementary motor area/anterior cingulate cortex; AI/FO_L and AI/FO_R: left and right anterior insula/frontal operculum; aIFS_L and aIFS_R: left and right anterior inferior frontal sulcus; IFJ_L and IFJ_R: left and right inferior frontal junction; IPS_L and IPS_R: left and right intraparietal sulcus; pIFS_L and pIFS_R: left and right posterior inferior frontal sulcus; PM_L and PM_R: left and right premotor cortex. *Denotes t-test against chance-level decoding (50%), p<0.05

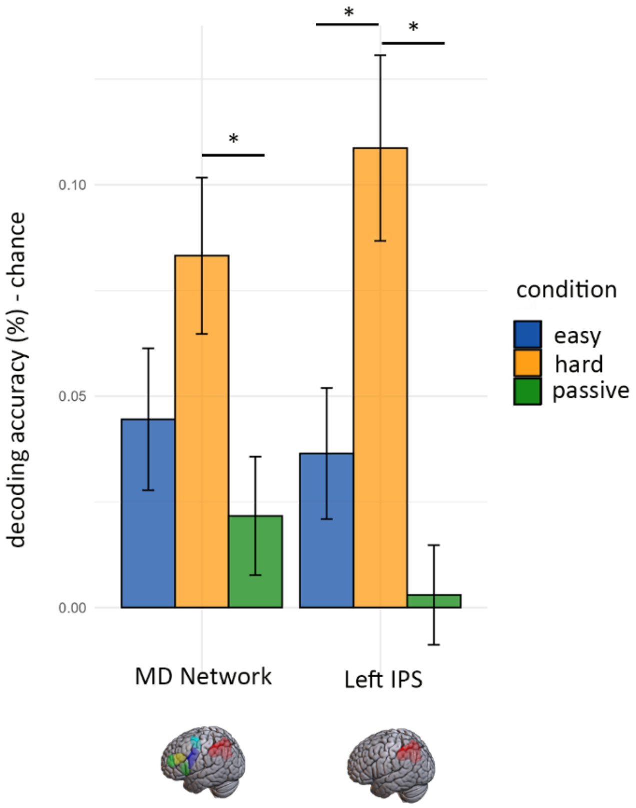


Figure 5. Effect of task demands on memory decoding in the MD Network and left IPS. The decoding accuracy of the memorandum in the combined MD ROI was enhanced with task difficulty (statistically significant main effect of the digit task condition); this effect was most pronounced in the decoding in the left IPS subregion: the accuracy of memory decoding increased with increasing task demands (i.e. hard digit task condition resulted in enhancement of the decoding accuracy in the left IPS in comparisons with easy or passive digit task conditions). *Denotes paired t-test, p<.05

Decoding in Early Visual Cortex

As depicted in Figure 6, memoranda could be reliability decoded from the early visual cortex irrespective of the digit task difficulty (i.e. in all three conditions, easy: M= 57.45%, SD= 9.22%, t(21)=3.79, p=.001; hard: M=57.73, SD=7.98, t(21)=4.54, p=.0001; passive: M=60.29%, SD=7.38%, t(21)= 6.53, p< .0001). There was no significant main effect of the digit task condition on memory decoding in early visual areas (F(2, 42) = .935, p = .4). This finding demonstrates that demands to the secondary task performed during the delay period did not change the representations in the early visual cortex.

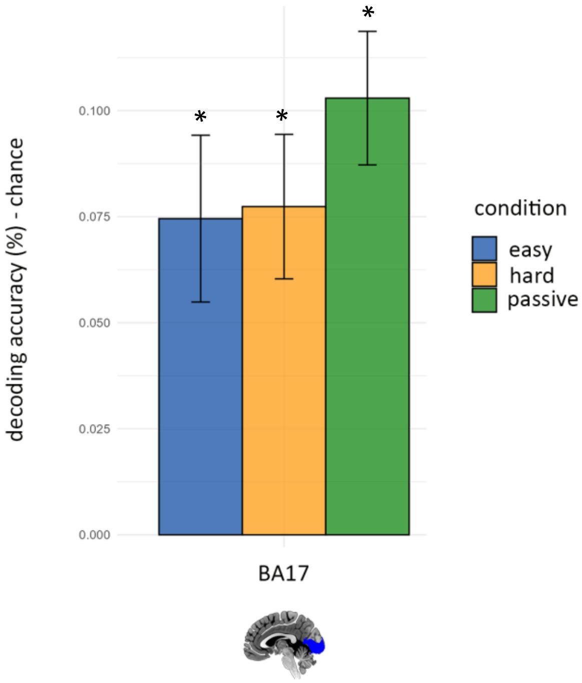


Figure 6. Effect of increasing task demands on memory category decoding in early visual cortex. The representations of memories could be reliably decoded independently of the task demands of the interrupter presented during memory maintenance (i.e. in all digit task conditions). There was no significant effect of the digit task difficulty on decoding within this region. *Denotes t-test against chance (50%), p<.05

Did MD regions code for the digit task?

In addition to the main analysis, we tested whether observed decoding accuracy was related to the BOLD activity evoked by the digit task. We correlated mean decoding accuracy and mean BOLD activation (resulting from the GLM analysis) for that purpose. Specifically, the values were calculated to represent the digit task difficulty by subtracting values in easy condition from the hard condition values and corrected by the passive condition serving as a baseline, i.e. no digit task.

We found no relationship between averaged decoding accuracy and BOLD activity across subjects and MD ROIs (r(20)=−0.22, p=.34, 95% CI [−0.58,0.23]), as illustrated in Figure 7. Furthermore, decoding and BOLD activity did not correlate in any of MD regions, as shown in Table 1 and Figure 8.

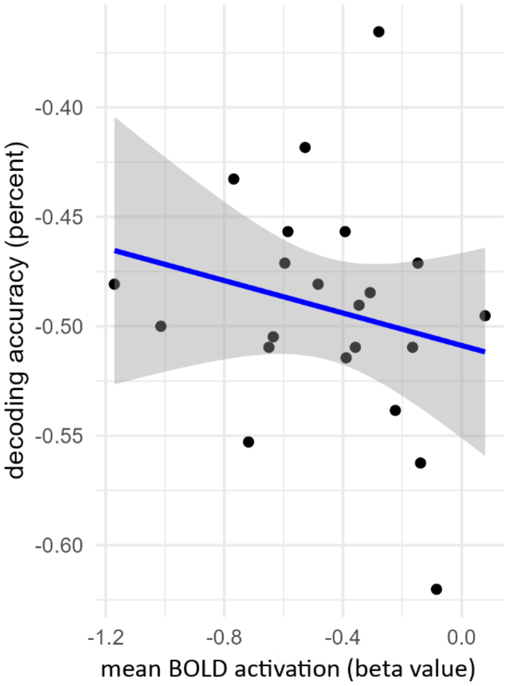
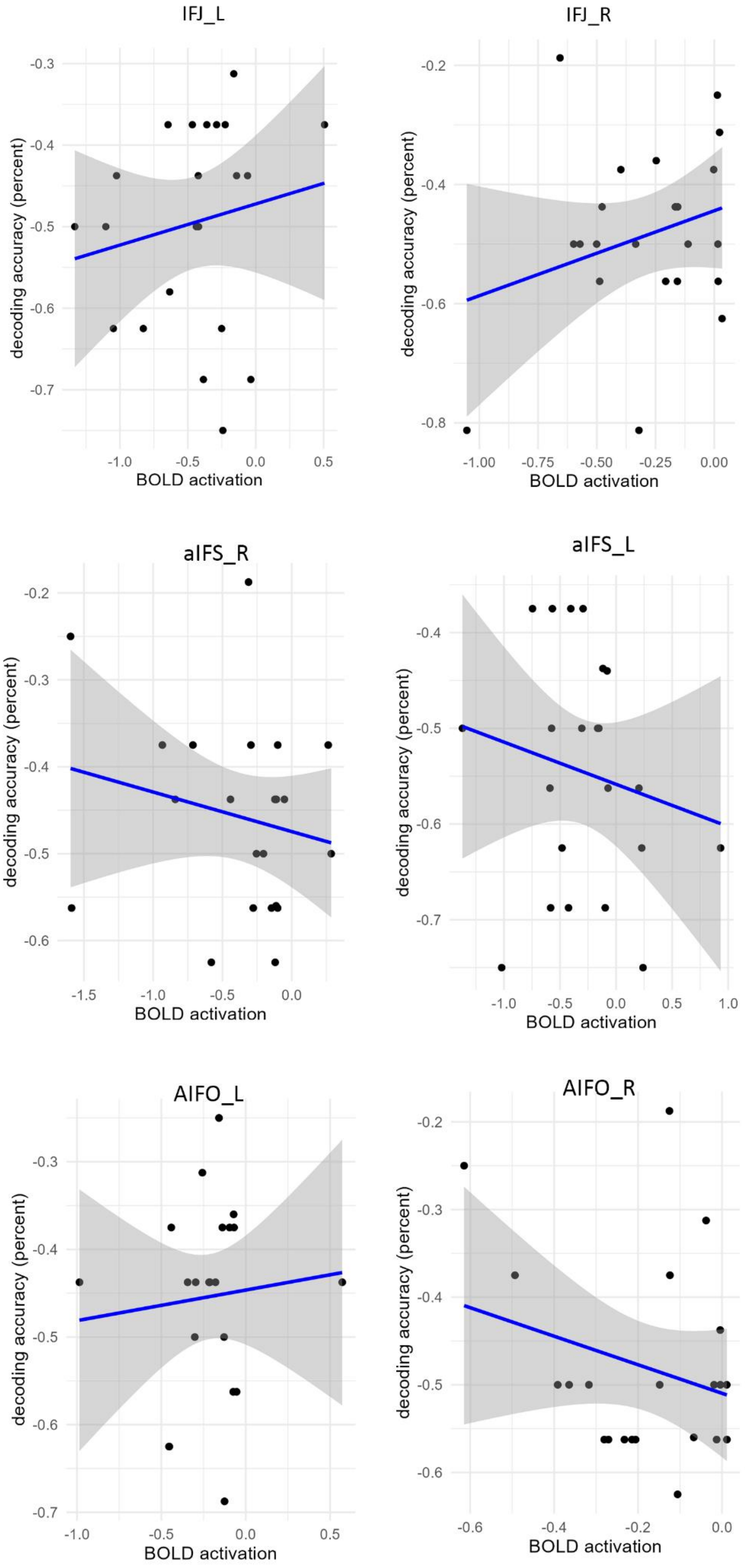
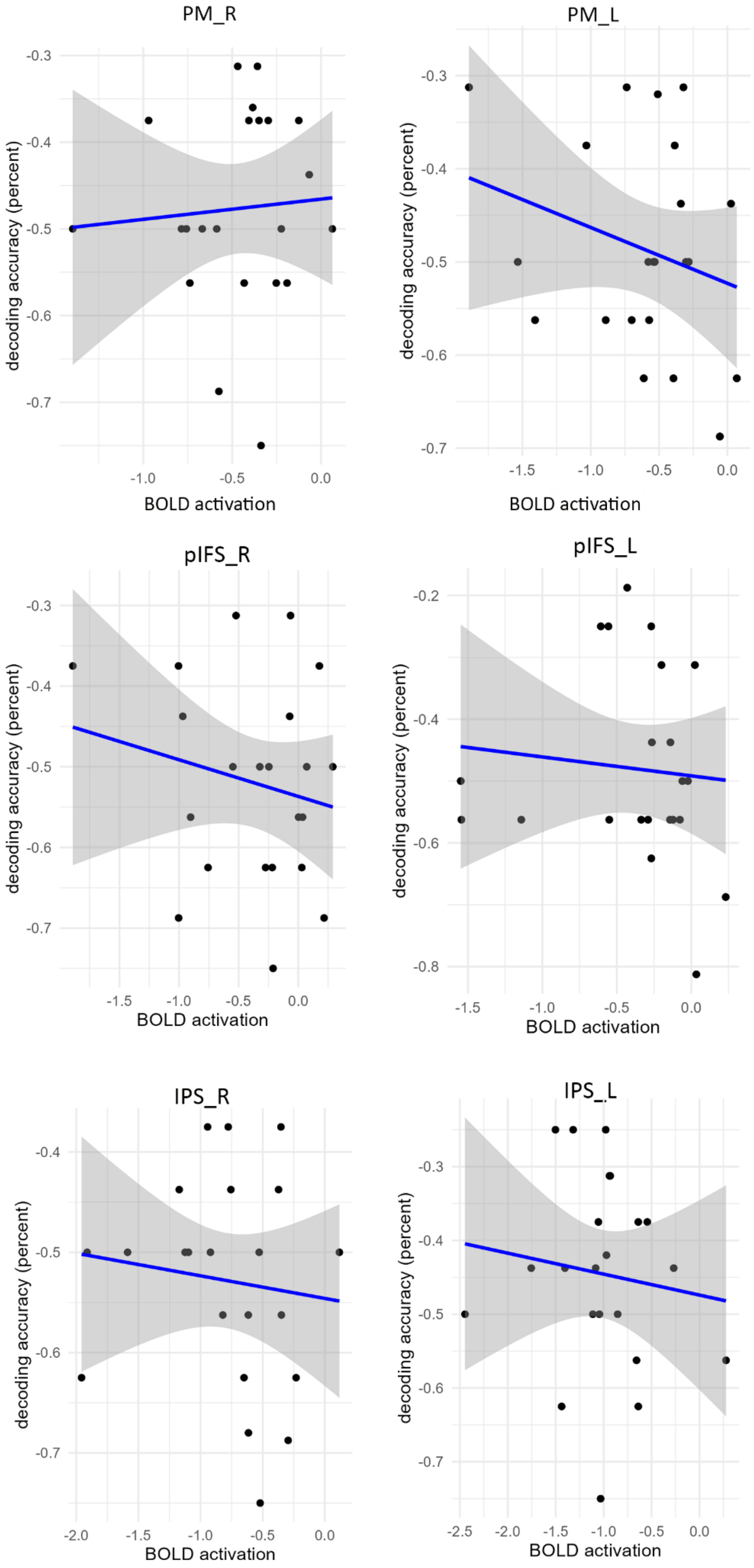


Figure 7. Correlation plot of the BOLD activation and decoding accuracy within MD regions averaged across subjects. The values represent the subtraction of easy and passive condition values from the hard condition value. Analysis did not reveal any relationship between the decoding result and BOLD activation.

Table 1. Results of Spearman correlation between decoding accuracy and BOLD activation in each MD region. ACC: bilateral presupplementary motor area/anterior cingulate cortex; AI/FO_L and AI/FO_R: left and right anterior insula/frontal operculum; aIFS_L and aIFS_R: left and right anterior inferior frontal sulcus; IFJ_L and IFJ_R: left and right inferior frontal junction; IPS_L and IPS_R: left and right intraparietal sulcus; pIFS_L and pIFS_R: left and right posterior inferior frontal sulcus; PM_L and PM_R: left and right premotor cortex.

MD region	r	p
PM_R	0.045	0.841
PM_L	-0.139	0.536
pIFS_R	-0.143	0.524
pIFS_L	-0.261	0.240
IPS_R	-0.241	0.278
IPS_L	-0.096	0.668
IFJ_R	0.069	0.759
IFJ_L	-0.171	0.445
aIFS_R	-0.143	0.524
aIFS_L	-0.158	0.48
AIFO_R	-0.119	0.597
AIFO_L	0.131	0.56
ACC	-0.04	0.842


IFJ_L
-0.3
-0.4
-0.5
-0.6
-0.7
decoding accuracy (percent)
-1.0
-0.5
0.0
0.5
BOLD activation
IFJ_R
-0.2
-0.4
-0.6
-0.8
decoding accuracy (percent)
-1.00
-0.75
-0.50
-0.25
0.00
BOLD activation
aIFS_R
-0.2
-0.3
-0.4
-0.5
-0.6
decoding accuracy (percent)
-1.5
-1.0
-0.5
0.0
BOLD activation
aIFS_L
-0.4
-0.5
-0.6
-0.7
decoding accuracy (percent)
-1.0
-0.5
0.0
0.5
1.0
BOLD activation
AIFO_L
-0.3
-0.4
-0.5
-0.6
-0.7
decoding accuracy (percent)
-1.0
-0.5
0.0
0.5
BOLD activation
AIFO_R
-0.2
-0.3
-0.4
-0.5
-0.6
decoding accuracy (percent)
-0.6
-0.4
-0.2
0.0
BOLD activation


PM_R
decoding accuracy (percent)
-0.3
-0.4
-0.5
-0.6
-0.7
-1.0
-0.5
0.0
BOLD activation
PM_L
decoding accuracy (percent)
-0.3
-0.4
-0.5
-0.6
-0.7
-1.5
-1.0
-0.5
0.0
BOLD activation
pIFS_R
decoding accuracy (percent)
-0.3
-0.4
-0.5
-0.6
-0.7
-1.5
-1.0
-0.5
0.0
BOLD activation
pIFS_L
decoding accuracy (percent)
-0.2
-0.4
-0.6
-0.8
-1.5
-1.0
-0.5
0.0
BOLD activation
IPS_R
decoding accuracy (percent)
-0.4
-0.5
-0.6
-0.7
-2.0
-1.5
-1.0
-0.5
0.0
BOLD activation
IPS_L
decoding accuracy (percent)
-0.3
-0.4
-0.5
-0.6
-0.7
-2.5
-2.0
-1.5
-1.0
-0.5
0.0
BOLD activation

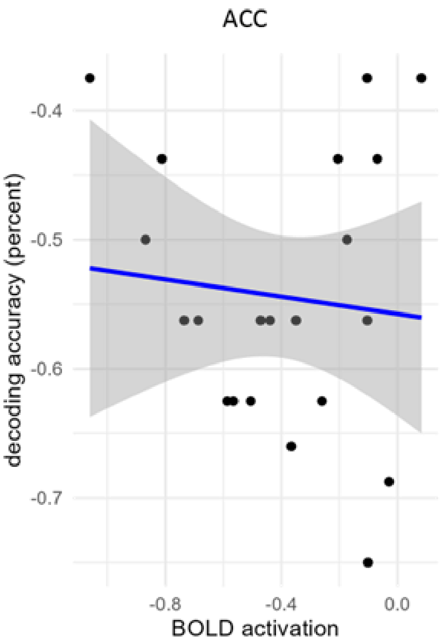


Figure 8. Correlation plots demonstrating the relationship between averaged BOLD activation (beta values) and decoding accuracy in MD regions. No significant correlation was found in any of the MD ROIs. ACC: bilateral presupplementary motor area/anterior cingulate cortex; AI/FO_L and AI/FO_R: left and right anterior insula/frontal operculum; aIFS_L and aIFS_R: left and right anterior inferior frontal sulcus; IFJ_L and IFJ_R: left and right inferior frontal junction; IPS_L and IPS_R: left and right intraparietal sulcus; pIFS_L and pIFS_R: left and right posterior inferior frontal sulcus; PM_L and PM_R: left and right premotor cortex.

Discussion

This study sought to understand how the MD network is involved in protecting WM contents from interruption. We used MVPA to examine how memory representations in MD regions and the early visual cortex change when a digit task is presented during the delay period, aimed to interfere with memory maintenance. We found that digit task difficulty modulated the decoding of the memorandum in the MD network. This effect was most pronounced in the IPS subregion.

In line with previous evidence of the MD network supporting demanding cognitive tasks, including WM (Fedorenko et al., 2013) and coding adaptively relevant information in the presence of irrelevant information (Jackson, 2018), we found memoranda decoding in the MD regions. Importantly, this decoding was enhanced by the digit task difficulty. This finding leads to several important implications. First, it demonstrates the protective mechanism against interference in WM. In light of this finding, prior research demonstrating the involvement of prefrontal regions (Feredoes et al., 2011; Dolcos et al., 2007; Clapp et al., 2009; Yoon et al., 2006) and parietal cortex (Bettencourt & Xu, 2016; Lorenc et al., 2018; Rademaker et al., 2019) in memory protection can be reconciled by proposing their collective role as subregions of the MD network. In addition, while Hallenbeck et al. (2021) proposed that the primary function of these regions is to maintain a stable storage of memoranda, ensuring their reinstatement in sensory areas after the distraction, the present findings suggest a broader role. Beyond simple memory maintenance, these regions are also involved in the flexible modulation of memory, specifically enhancing its fidelity, in response to the increasing demands of the task. There is a clear difference between these two mechanisms. The mechanism proposed by Hallenbeck et al. (2021) suggests that memory is encoded consistently across multiple cortical areas, regardless of distraction, and is protected through redundancy. If memory becomes distorted in one area, a copy of the memory can be transferred from another area and reinstated, ensuring its stability. However, our results demonstrate that MD regions do not merely maintain a stable representation in the face of interruption; rather, they enhance the representation in response to the difficulty of the interrupter task, demonstrating their adaptive, task-dependent role in modulating memory strength to meet cognitive demands. Moreover, the code enhancement observed was not related to the digit task itself, as revealed by additional correlation analysis of decoding accuracy and mean BOLD activation within MD ROIs. Second, more broadly, present results deepen our understanding of how the MD system elicits cognitive control over cognitive tasks, specifically, by demonstrating that it enhances the codes of the relevant information.

Comparing memory protection mechanisms proposed by previous decoding evidence (Bettencourt & Xu, 2016; Lorenc et al., 2018; Rademaker et al., 2019; Hallenbeck et al., 2021; Vu-Cheung et al., 2021; Kiyonaga et al., 2017) is difficult because of distinct approaches to studying this process. Namely, memory targets vary between the previous decoding studies, e.g. Bettencourt & Xu (2016), Lorenc et al. (2018) and Rademaker et al. (2019) decoded orientations, Hallenbeck et al.(2021) decoded visual polar angle, Vu-Cheung et al. (2021) decoded coloured dot location, and Kiyonaga et al. (2017) decoded object category (faces vs houses); therefore, since the types of information maintained in memory differ, distinct neural substrates may be involved in its protection. Indeed, one possibility is that simple features such as orientations rely more on the sensory regions such as the primary visual cortex as demonstrated by neural decoding evidence (Serences et al., 2009), while more complex, abstract information like object categories may require the involvement of higher-order areas such as the prefrontal cortex as recorded during delay period in monkeys (Freedman et al., 2001); thus, the decoding in present study within MD region could be explained by the usage of naturalistic objects as memory targets; hence, future research is needed to investigate the role of the MD network in protecting memory representations of simple features as well.

Notably, while interrupters produce a larger detrimental impact on WM performance than distractors and cause more severe disruption in brain connectivity (Clapp et al., 2010), most of the decoding studies have focused on the effects of distractors on WM representations (Bettencourt & Xu, 2016; Lorenc et al., 2018; Rademaker et al., 2019; Hallenbeck et al., 2021; Vu-Cheung et al., 2021). Only a limited number of studies using neural decoding have addressed the impact of interruptions, including the present study and the study by Kiyonaga et al. (2017), which used a visual search task to disrupt the memoranda but tested memory decoding only in sensory regions. Thus, the present findings contribute significantly to the study of interruptions in WM and provide evidence for the mechanism of memory protection driven by cognitive control within the MD network. Furthermore, the enhancement of memory representations we found in our study can explain the improvement in task performance in the presence of interrupters (Zickerick et al., 2020; Kiyonaga et al., 2017): The brain may quickly adapt to the difficulty of the interrupter task, often overcompensating by deploying more cognitive resources than required for the primary task. Specifically, Zickerick et al. (2020) used EEG to demonstrate that participants exhibited increased amplitudes of the fronto-central N2 component following an interruption - a marker reflecting cognitive control and conflict monitoring. In addition, an increase in fronto-central slow waves was observed after interference, particularly in conditions with WM load, indicating that interference may trigger a mechanism that enhances memory rehearsal and maintenance (Zickerick et al., 2020). Using a decoding approach to neuroimaging data, Kiyonaga et al. (2017) reported a surprising finding: despite the greater task demand, decoding accuracy in the hard interruption condition was higher at the later time points in the WM delay than in trials with easy interruption. Researchers proposed that the observed decoding pattern likely reflected a compensatory reinstatement of WM representations; initially suppressed memory traces were reactivated once attentional resources became available after completing the interrupting task, enabling their retrieval and improved decoding at later time points (Kiyonaga et al., 2017). The present finding is in accord with these results since it also showed an increase in memory decoding accuracy following the higher difficulty of the interrupting task. However, the current experiment did not provide any temporal resolution of the decoding dynamics; therefore, it is unclear whether the representations were enhanced after the interruption as in previous studies (Zickerick et al., 2020; Kiyonaga et al., 2017) or during the interrupter task performance. Thus, future research could explore the dynamics of the compensatory mechanism proposed by these findings; namely, studies could this question by examining how the MD network adjusts its coding strategies in response to varying task demands, especially under conditions of increased cognitive load or interruptions while also considering individual differences in cognitive flexibility and resource allocation.

In addition, to provide a deeper understanding of the role of the MD system in protecting memory representations, further research should address the question of communication among these regions and its dynamics. For example, comparing the effective connectivity (Friston et al., 2003) among MD regions under high versus low interrupter task difficulty could reveal how the strength and directionality of connections shift as task demands increase, requiring greater memory protection. Furthermore, since the present decoding analysis was based on the activity within the whole duration of the delay period, the decoding results do not disentangle the activity during and after the interruption. Thus, it remains unclear whether the enhancement of memory codes within MD network happened during or after the interrupter task. The MD regions might simultaneously encode the interrupter task while amplifying the memorandum to shield it from distortion. Alternatively, the MD system could transition from representing the memories to focusing on processing the interrupter, and subsequently reinstating the enhanced memory representation once the interrupter is processed. To clarify these possibilities and overcome the low temporal resolution of fMRI, future studies may analyse the representational flow between MD regions and sensory cortices, as demonstrated by Goddard et al.

(2016) using MEG and MVPA: researchers showed that object identity was coded in the occipital cortex as early as 80 ms and in the frontal regions by 265 ms. This approach will help reveal how MD regions and sensory areas code the memorandum during encoding, delay, and recall, as well as during and after interference.

Moreover, the present findings provide further support for the adaptive coding hypothesis, which proposes that MD neurons adjust their representations of relevant information (Duncan, 2010). Evidence for this is based on extensive evidence, including decoding (Jackson et al., 2017; Jackson & Woolgar, 2018; Jackson et al., 2024). The current study extends this hypothesis by demonstrating that MD regions adjust their representation of a memory task, particularly enhancing it when faced with interruptions. Although we did not investigate the coding of the digit task, it is likely that the MD system was engaged in its processing, as has been observed in various demanding cognitive tasks (Fedorenko et al., 2013; Woolgar et al., 2011; Woolgar et al., 2016).

The adaptive flexibility of the MD system demonstrated in the present study suggests a potential mechanism for task-switching, wherein both tasks are simultaneously coded within the same network. Supporting this idea, a recent decoding study found that MD regions encode modality-specific information for both visual and auditory tasks within the same voxels (Jackson et al., 2024). These findings underscore the ability of the MD system to allocate resources across different tasks while maintaining distinct, modality-specific codes, enabling adaptive responses to varying cognitive demands. As a result, the flexibility in MD neuron coding may form the basis for task-switching. Future research is needed to explore how multiple task representations coexist within the MD regions to better understand this mechanism.

Consistent with previous findings we decoded memory representations in the IPS in the face of irrelevant information (i.e. digit task) presented at the delay period (Bettencourt & Xu, 2016; Lorenc et al., 2018; Rademaker et al., 2019). However, unlike similar decoding in the presence or absence of distractors in these studies (Bettencourt & Xu, 2016; Lorenc et al., 2018), we observed an increase in memory decoding accuracy with the increasing cognitive demands of the digit task. Importantly, in contrast with previous studies (Bettencourt & Xu, 2016; Lorenc et al., 2018) which displayed a visual stimulus, i.e. orientation in Lorenc et al., (2018) and faces or gazebos in Bettencourt & Xu, (2016), as a distractor, the increase in decoding in the IPS could not be attributed to visual processing of distractor stimulus since the visual display of the digit task was nearly identical in all digit task conditions; the frame around the digits was black or red to indicate the task rule to apply, but this cannot account for differences in IPS decoding accuracy. This finding suggests that the IPS protects memory from irrelevant information by not only maintaining the memorandum in the face of irrelevant information but also enhancing its representation, possibly rendering it more resistant to the negative impact of distraction under increased task demands.

Multiple studies of visual representations suggest the multifaceted role of the parietal cortex in integrating visual, cognitive, and motor functions to support adaptive visual processing and interactions with the environment (see Xu, 2018, for a review). For example, Bray et al. (2015) used an fMRI connectivity approach to demonstrate that IPS is activated in both spatial and arithmetic manipulations but establishes distinct connections with other regions for different tasks, suggesting a domain-general role in tasks that require the manipulation of currently relevant information (Bray et al., 2015). In another study, the left IPS showed increased activation with a larger WM load both in visual and auditory modalities, pointing to its domain-general function in maintaining memoranda (Cowan et al., 2011). Considering these views, the enhancement of memory representations with increased task demands found in the present study could reflect increased cognitive control over memoranda in a goal-directed, adaptive manner, i.e., the higher cognitive load, the more relevant representations are boosted to keep them in an active state.

Another intriguing possibility is that the modulation of memory representations in the IPS found in this study could be explained by increased activation related to the processing of the interrupting task in this region. Although MVPA is based on the spatial patterns of neural activity, the overall magnitude of the activation can potentially sharpen this pattern. Albers et al., (2018) investigated the relationship between overall blood-oxygen-level-dependent (BOLD) signal and the decoding accuracy of stimulus information in the early visual cortex using MVPA. The ability to decode stimulus information relied on subtle but reliable differences in BOLD signal magnitude between voxels tuned to preferred and non-preferred orientations; importantly, the magnitude of these BOLD signal differences correlated with the accuracy of the MVPA decoding (Albers et al., 2018). Thus, it is possible that the memory representations maintained in IPS in our study were potentially enhanced by the overall increase in BOLD signal magnitude related to digit task processing.

Indeed, left IPS is involved in various tasks involving numerosity (Cappelletti et al., 2007; Bugden et al., 2012) which would also be required to perform the digit sorting task in the present study. For instance, Cappelletti et al. (2007) applied repetitive TMS to the left intraparietal sulcus (IPS) while participants compared numerical quantities presented in both symbolic form (Arabic numerals) and non-symbolic form (dot arrays) to a reference number. The disruption in performance on both tasks following left IPS stimulation suggests that common neural mechanisms underlie these processes. This finding highlights the causal role of the left IPS, rather than the right IPS, in numerosity processing (Cappelletti et al., 2007), which is involved in our digit task as well. Therefore, there is a possibility that numerosity-related processing, as reflected in increased BOLD signal magnitude, could enhance the accuracy of decoding in the left IPS. Hence, more studies using the decoding approach are needed to compare the effects of different types of interruptions, and perceptual distractors as well, to clarify the role of the (left) IPS in memory protection. In addition, not all IPS regions are involved in WM equally: as demonstrated by Bettencourt & Xu (2016), decoding in superior IPS correlates with WM performance, whereas other parietal regions such as inferior parietal lobule (IPL) and superior parietal lobule (SPL) do not show the same consistency in decoding. Since here we used an ROI-based approach to decoding as in prior studies of information coding within the MD network (Jackson et al., 2018; Jackson et al., 2024), the parcellations within the IPS were not considered and, thus, need to be addressed in the future studies to reveal their specific functions in the WM protection and communication with other MD regions.

Finally, memory representations in the early visual cortex remained reliably decodable despite interference from the digit task, contrasting with prior decoding evidence that demonstrated lower memory decoding accuracy in the presence of visual distractors (Bettencourt & Xu, 2015; Lorenc et al., 2018; Rademaker et al., 2019) or interrupters (Kiyonaga et al., 2017). Although this result is in line with previous decoding studies showing the same decoding accuracy in early visual areas with and without distractors (Hallenbeck et al., 2021; Vu-Cheung et al., 2021), it is possible that the modulation of memory codes we found in the MD network may have impacted the decoding in sensory areas. For example, this proposal may be supported by the TMS-fMRI-MVPA study investigating the causal role of the dlPFC in modulating relevant information in the visual cortex and MD network (Jackson et al., 2021). Researchers applied TMS to the dlPFC during a selective attention task during fMRI scanning; decoding analysis revealed the reduced coding of relevant information in MD regions and visual cortices following dlPFC TMS, suggesting that this region enhances relevant information. In contrast, irrelevant information coding was unaffected by dlPFC TMS, suggesting that the primary role of this region is to boost relevant information rather than suppress irrelevant (Zanto et al., 2011; Zanto & Gazzaley, 2009). Therefore, the causal evidence for the dlPFC's role in mitigating distractors, as discussed in the previous section, may be attributed to its ability to strengthen memory representations, thereby reducing the impact of external distractions. However, future studies using a concurrent TMS-fMRI approach are required to examine the role of the dlPFC in modulating MD code for memory representations both within MD regions and areas connected to them in this process such as early visual areas. In addition, future studies need to address causally the specific role of the (left) IPS in maintaining and/or modulating memory presentations within MD regions and sensory cortices.

In addition, it could be argued that early visual areas were not critical for maintaining object representations in the present study since object processing involves multiple areas from early visual areas to higher-level regions (Hebart et al., 2018). However, participants were instructed to memorise objects in detail, as the recall phase involved pairs of similar objects difficult to resemble, thus requiring them to retain fine details in memory, which likely involved early visual areas. However, the visual display of digit tasks during the delay period remained the same across conditions and, possibly, was too distinct to interfere with memoranda in the early visual cortex.

Importantly, understanding how the brain protects WM from interruptions is critical for designing better clinical interventions for neurological conditions in which this ability is impaired, such as attention deficit hyperactivity disorder (ADHD), schizophrenia, Alzheimer's disease or other forms of dementia where WM is impaired. Specifically, techniques such as cognitive training or neuromodulation, e.g., TMS, could be optimised to enhance the functioning of MD regions and improve cognitive control and WM.

To conclude, the present study investigated the role of the MD network in exerting cognitive control over memoranda to protect them from interference caused by an interrupting task. Our findings suggest that the MD network plays a critical role in maintaining and enhancing memory representations under increased cognitive demands. This highlights the adaptive flexibility of the MD system in protecting relevant information, even in the face of challenging interruptions. These findings not only advance our understanding of the neural underpinnings of cognitive control and memory maintenance but also highlight the essential role of the MD network in optimising cognitive performance in the face of interference. Moreover, they potentially point to representational mechanisms that implement task-switching, suggesting that the MD network can dynamically adapt to handle MDs by flexibly coding and protecting relevant information across tasks. Future research should further explore the mechanisms by which the MD network modulates memory representations and investigate potential applications for enhancing cognitive resilience in clinical populations.

References

Albers, A. M., Meindertsma, T., Toni, I., & de Lange, F. P. (2018). Decoupling of BOLD amplitude and pattern classification of orientation-selective activity in human visual cortex. NeuroImage, 180, 31-40.

Assem, M., Glasser, M. F., Van Essen, D. C., & Duncan, J. (2020). A domain-general cognitive core defined in multimodally parcellated human cortex. Cerebral Cortex, 30 (8), 4361-4380.

Baddeley, A. (1992). WM: The interface between memory and cognition. Journal of cognitive neuroscience, 4 (3), 281-288.

Bae, G. Y., & Luck, S. J. (2019). What happens to an individual visual WM representation when it is interrupted?. British Journal of Psychology, 110(2), 268-287.

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.

Bugden, S., Price, G. R., McLean, D. A., & Ansari, D. (2012). The role of the left intraparietal sulcus in the relationship between symbolic number processing and children's arithmetic competence. Developmental cognitive neuroscience, 2(4), 448-457.

Bray, S., Almas, R., Arnold, A. E., Iaria, G., & MacQueen, G. (2015). Intraparietal sulcus activity and functional connectivity supporting spatial WM manipulation. Cerebral Cortex, 25(5), 1252-1264.

Cappelletti, M., Barth, H., Fregni, F., Spelke, E. S., & Pascual-Leone, A. (2007). rTMS over the intraparietal sulcus disrupts numerosity processing. Experimental brain research, 179, 631-642.

Clapp, W. C., Rubens, M. T., & Gazzaley, A. (2010). Mechanisms of WM disruption by external interference. Cerebral cortex, 20 (4), 859-872.

Cowan, N., Li, D., Moffitt, A., Becker, T. M., Martin, E. A., Saults, J. S., & Christ, S. E. (2011). A neural region of abstract WM. Journal of cognitive neuroscience, 23(10), 2852-2863.

Dolcos, F., Miller, B., Kragel, P., Jha, A., & McCarthy, G. (2007). Regional brain differences in the effect of distraction during the delay interval of a WM task. Brain Research, 1152, 171-181.

Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. Trends in Cognitive Sciences, 14(4), 172-179.

Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. Proceedings of the National Academy of Sciences, 110(41), 16616-16621.

Feredoes, E., Heinen, K., Weiskopf, N., Ruff, C., & Driver, J. (2011). Causal evidence for frontal involvement in memory target maintenance by posterior brain areas during distracter interference of visual WM. Proceedings of the National Academy of Sciences, 108 (42), 17510-17515.

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

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. Neuroimage, 19(4), 1273-1302.

Goddard, E., Carlson, T. A., Dermody, N., & Woolgar, A. (2016). Representational dynamics of object recognition: Feedforward and feedback information flows. Neuroimage, 128, 385-397.

Hallenbeck, G. E., Sprague, T. C., Rahmati, M., Sreenivasan, K. K., & Curtis, C. E. (2021). WM representations in visual cortex mediate distraction effects. Nature Communications, 12(1), 4714.

Jackson, J. B., Feredoes, E., Rich, A. N., Lindner, M., & Woolgar, A. (2021). Concurrent neuroimaging and neurostimulation reveals a causal role for dlPFC in coding of task-relevant information. Communications Biology, 4(1), 588.

Jackson, J. B., & Woolgar, A. (2018). Adaptive coding in the human brain: Distinct object features are encoded by overlapping voxels in frontoparietal cortex. Cortex, 108, 25-34.

Jackson, J., Rich, A. N., Williams, M. A., & Woolgar, A. (2017). Feature-selective attention in frontoparietal cortex: multivoxel codes adjust to prioritize task-relevant information. Journal of Cognitive Neuroscience, 29(2), 310-321.

Kiyonaga, A., Dowd, E. W., & Egner, T. (2017). Neural representation of WM content is modulated by visual attentional demand. Journal of Cognitive Neuroscience, 29(12), 2011-2024.

Lorenc, E. S., Sreenivasan, K. K., Nee, D. E., Vandenbroucke, A. R., & D'Esposito, M. (2018). Flexible coding of visual WM representations during distraction. Journal of Neuroscience, 38(23), 5267-5276.

MathWorks, Inc. (2019). MATLAB Release 2019a. Natick, MA: The MathWorks, Inc.

Menon, V., & D’Esposito, M. (2022). The role of PFC networks in cognitive control and executive function. Neuropsychopharmacology, 47(1), 90-103.

Oosterhof, N. N., Connolly, A. C., & Haxby, J. V. (2016). CoSMoMVPA: multi-modal multivariate pattern analysis of neuroimaging data in Matlab/GNU Octave. Frontiers in Neuroinformatics, 10, 27.

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage, 59(3), 2142-2154.

Rademaker, R. L., Chunharas, C., & Serences, J. T. (2019). Coexisting representations of sensory and mnemonic information in human visual cortex. Nature Neuroscience, 22(8), 1336-1344.

Rorden, C. (2007). Mricron [computer software].

Ruff, C. C., Bestmann, S., Blankenburg, F., Bjoertomt, O., Josephs, O., Weiskopf, N., ... & Driver, J. (2008). Distinct causal influences of parietal versus frontal areas on human visual cortex: evidence from concurrent TMS–fMRI. Cerebral Cortex, 18(4), 817-827.

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.

Stojanoski, B., & Cusack, R. (2014). Time to wave good-bye to phase scrambling: Creating controlled scrambled images using diffeomorphic transformations. Journal of Vision, 14(12), 6-6.

Van Dijk, K. R., Sabuncu, M. R., & Buckner, R. L. (2012). The influence of head motion on intrinsic functional connectivity MRI. Neuroimage, 59(1), 431-438.

Vu-Cheung, K., Ester, E. F., & Sprague, T. C. (2021). Spatial WM representations in human cortex are robust to a task-irrelevant interrupting stimulus. bioRxiv, 2021-09.

Woolgar, A., Jackson, J., & Duncan, J. (2016). Coding of visual, auditory, rule, and response information in the brain: 10 years of multivoxel pattern analysis. Journal of cognitive neuroscience, 28(10), 1433-1454.

Woolgar, A., Thompson, R., Bor, D., & Duncan, J. (2011). Multi-voxel coding of stimuli, rules, and responses in human frontoparietal cortex. Neuroimage, 56(2), 744-752.

Yoon, J. H., Curtis, C. E., & D'Esposito, M. (2006). Differential effects of distraction during WM on delay-period activity in the prefrontal cortex and the visual association cortex. Neuroimage, 29(4), 1117-1126.

Zickerick, B., Thönes, S., Kobald, S. O., Wascher, E., Schneider, D., & Küper, K. (2020). Differential effects of interruptions and distractions on working memory processes in an ERP study. *Frontiers in Human Neuroscience*, *14*, 84.

Chapter 5. General Discussion

Key Findings

In this thesis, I investigated neural mechanisms of resolving interference in working memory focusing on the role of cognitive control in this process. Specifically, I examined two types of interference: external distractors and interrupters.

To investigate a fine-grained measure of distractor impact on working memory for objects, I tested a continuous recall approach to measuring the memory precision of naturalistic objects (Veldsman et al., 2017; Stojanoski et al., 2019) in the presence of visual distractors during the delay period. The results revealed that object memory was robust to external distractors independent of their visual features, i.e. movement and vibrant colours.

To test the role of cognitive control in distractor mitigation, I examined how individual differences in tendencies to mind wander or worry, since they are associated with lower cognitive control in working memory (Robison and Unsworth, 2018; Gustavson & Miyake, 2015). Therefore, I investigated the link between these tendencies and the ability to manage external distractions in working memory. The results revealed that cognitive control over mind-wandering interacts with susceptibility to external distraction. Further, individuals who have less control over mind-wandering episodes benefit from more visually engaging distractors, perhaps by helping them to stay on task. In addition, trait worry was not related to a lower ability to mitigate external distraction. However, a trend in the data indicates that distractors may still consume cognitive resources in individuals prone to worry, potentially making it harder for them to maintain compensatory strategies that support task performance over time.

Pupil dilation serves as an indicator of cognitive effort (Van der Wel et al., 2018), with tonic pupil activity reflecting overall task engagement and larger fluctuations denoting weaker cognitive control (Robison & Unsworth, 2018), while phasic pupil activity reflects immediate cognitive effort (Naber et al., 2013). Therefore, I applied the pupillometry approach to show that phasic pupil response during distractor presentation reflects the allocation of cognitive resources and not the effort for visual processing per se. The phasic pupil response during the delay period reflected unimpaired increased active maintenance of the memorandum. Tonic pupil fluctuations may reflect individual differences in cognitive control during working memory tasks.

To expand the causal evidence of the dorsolateral prefrontal cortex (dlPFC) in resolving distraction in visual working memory (Feredoes et al., 2011), I applied a neural stimulation approach and demonstrated in healthy humans that dlPFC is causally involved in mitigating distracting stimuli, reinforcing its role in exerting cognitive control over memory in the presence of external distractions.

Finally, I investigated the role of the Multiple Demand (MD) network (Duncan, 2010) in protecting memory representations from interruption, based on research demonstrating its key role in supporting cognitive control across a wide range of tasks (Fedorenko et al., 2013; Assem et al., 2020). I applied multivariate pattern analysis (MVPA) to decode memory contents within MD regions and found that memory representations in MD network are enhanced by interrupter task difficulty, providing a representational account for the cognitive control mechanism of MD system to protect memoranda from interference.

In the following sections, I will delve deeper into these findings, discuss their connections and implications, explore practical applications, and propose future directions for research on interference in working memory.

Causal Evidence of dlPFC-Driven Cognitive Control in Mitigating External Distraction

Causal evidence demonstrates that dlPFC is crucial for working memory maintenance (Chao & Knight, 1998; Oliveri et al., 2001; Postle et al., 2006; Esslinger et al., 2014) and resolving visual distraction in working memory (Feredoes et al., 2011), proposing that this region may be key in eliciting cognitive control over memory maintenance (D'Esposito & Postle, 2015). However, neural decoding studies suggest that memory storage, rather than control mechanisms, is key to distractor mitigation. Specifically, these studies propose that resistant representations across multiple cortical regions (Hallenbeck et al., 2021; Vu-Cheung et al., 2021; Lorenc et al., 2018) or robust mnemonic formats (Rademaker et al., 2019) help maintain memory representations despite disruptions.

To clarify the role of dlPFC-driven cognitive control in working memory and reconcile previous findings from studies employing diverse methodologies to investigate this process, Chapter 3 investigated whether this region is causally involved in mitigating visual distractions by applying transcranial magnetic stimulation (TMS) to disrupt dlPFC activity. Consistent with previous causal evidence (Feredoes et al., 2011; Postle et al., 2006; Esslinger et al., 2014), the results confirmed that dlPFC is crucial for resolving external visual distractions in working memory. The mechanism of this process is suggested by the findings of Feredoes et al. (2011), where researchers used concurrent TMS-fMRI to perturb dlPFC activity during the memory delay period when visual distractors might be presented, while simultaneously acquiring fMRI data. Analysis revealed that TMS over dlPFC increased activity in visual areas representing the memory items, but only in the presence of distractors. This finding suggests that distractor mitigation might be achieved by dlPFC enhancing representations in sensory cortices, as indicated by increased connectivity between this region and the visual cortex following stimulation (Feredoes et al., 2011).

The results from Chapter 3 expand these findings by revealing the effect of distraction on object memory precision in the presence of dlPFC stimulation compared to the match-to-sample approach used in the study by Feredoes et al. (2011). Additionally, this study and the work by Feredoes et al. (2011) employed different TMS protocols designed to achieve distinct effects. In the present study, continuous Theta Burst Stimulation (cTBS) was used to induce long-lasting suppression of dlPFC, whereas Feredoes et al. (2011) applied three single TMS pulses during the distractor phase of each trial, aiming to momentarily perturb activity in dlPFC without causing prolonged disruption. The findings from this work demonstrated that the cTBS approach was effective in disrupting activity in dlPFC, further highlighting the versatility of TMS in studying distractor mitigation in working memory.

Notably, distraction had no effect in the absence of stimulation in the present study. This finding suggests that dlPFC effectively mitigated the interference caused by visual distractors. Therefore, the null behavioural findings may indicate the success of distractor mitigation, highlighting the necessity of using such techniques as TMS and neuroimaging to reveal how the brain is able to mitigate distractions and complete a working memory task successfully (i.e., without a detectable drop in accuracy).

The current results do not contradict the findings of neural decoding studies (Hallenbeck et al., 2021; Vu-Cheung et al., 2021; Lorenc et al., 2018; Rademaker et al., 2019) but instead highlight the importance of understanding how the interaction between dlPFC and the posterior cortex contributes to effective distraction mitigation. While neural decoding studies emphasise that memory storage across multiple cortical regions (Hallenbeck et al., 2021; Vu-Cheung et al., 2021) or the use of robust mnemonic formats helps resist distractions (Lorenc et al., 2018; Rademaker et al., 2019), the current findings continue to underscore the critical role of dlPFC in exerting cognitive control during distraction.

In a broader context, the involvement of dlPFC in mitigating visual distractions highlights its key role in cognitive control processes. According to the hierarchical control model of the prefrontal cortex proposed by Badre (2008), the posterior PFC (i.e., premotor cortex) mediates actions based on immediate sensory stimuli, while the anterior PFC manages higher-order representations, such as abstract rules, goals, and expectations that span longer time frames and guide behaviour beyond the immediate context. Dlpfc is located between these regions in the hierarchy and is responsible for episodic control—monitoring the temporal context of the current episode by integrating past information and future expectations to guide present behaviour (Badre, 2008). Irrelevant information can disrupt this process, interfering with episodic control and impacting the ability to focus on current goals. Therefore, mitigating distractions is essential for maintaining flexible yet stable behaviour, which is critical for decision-making, reasoning, and adapting to dynamic environments.

Future research needs to address the question of what the (distinct) representational roles of prefrontal, parietal and sensory cortices are in distractor mitigation in working memory, to reconcile previous findings using the decoding approach. In addition, the enhancement mechanism proposed by Feredoes et al. (2011) could be tested by future studies using neural decoding combined with a concurrent TMS-fMRI approach. By decoding memory representations across the cortex in the presence of distractions while perturbing dlPFC activity, researchers could examine how this region modulates memory representations in a broader range of cortical areas.

Furthermore, in Chapter 3, dlPFC TMS revealed how this region interacts with individual differences in mind wandering, worry, and phasic pupil responses during distractor presentation. This finding suggests that dlPFC engages with multiple brain networks involved in cognitive control, emphasizing the need to understand its function from a broader network perspective, which is explored in subsequent sections together with the suggestions for future studies addressing this question.

Role of the Multiple Demand Network in Protecting Working Memory Representations from Interruption

The causal finding discussed in the previous section highlights the necessity of dlPFC in eliciting cognitive control over memory maintenance in the face of distraction. However, a network of brain regions may be crucial to protect memory from interference. MD network is a likely candidate network since it is activated during a range of demanding cognitive tasks (Fedorenko et al., 2013). To examine the role of this network in resolving interference in working memory for objects, in Chapter 4, I introduced an interrupting task during the delay period since this kind of interference is more potent and ecologically valid (compared to visual distractors) and examined the fate of memory representations within MD network.

The decoding analysis in MD regions revealed that the difficulty of the interrupting task modulated memory representations within MD system, specifically, by enhancing them. This finding highlights the cognitive control mechanism that protects working memory against interference by adaptively maintaining and enhancing memory representations within MD network in the face of interference. In contrast to studies that interpreted decoding in frontoparietal regions as additional storage while considering visual cortices central for working memory maintenance (Hallenbeck et al., 2021; Vu-Cheung et al., 2021), the present finding demonstrated that the memory representations in MD regions serve not just a storage function, but flexible maintenance modulated by the task demands.

The enhancement of memory decoding seen in the present study can be explained in light of the study by Jackson et al. (2021). Researchers applied TMS to dlPFC during a selective attention task during fMRI scanning; decoding analysis revealed the reduced coding of relevant information in MD regions and visual cortices following dlPFC TMS, suggesting that this region enhances relevant information. In contrast, irrelevant information coding was unaffected by dlPFC TMS, suggesting that the primary role of this region is to boost relevant information rather than suppress irrelevant (Zanto et al., 2011; Zanto & Gazzaley, 2009). Therefore, the causal evidence for dlPFC's role in mitigating distractors, as discussed in the previous section, may be attributed to its ability to strengthen memory representations, thereby reducing the impact of external distractions. However, future studies using a concurrent TMS-fMRI approach need to investigate the role of dlPFC in modulating MD code for memory representations both within MD regions and areas connected to them in this process.

The present findings also provide further support for the adaptive coding hypothesis, which proposes that MD neurons adjust their representations of relevant information (Duncan, 2010). Evidence for this is based on extensive evidence, including decoding (Jackson et al., 2017; Jackson & Woolgar, 2018; Jackson et al., 2024). The present study extends this hypothesis by showing that MD regions adjust their representation of a memory task, enhancing it in the face of interruption. Although the present analysis did not provide the decoding of the interrupter task, MD system was likely engaged in processing the demanding task, as previously observed in multiple cognitive tasks (Fedorenko et al., 2013; Woolgar et al., 2011; Woolgar et al.,2016).

This adaptive flexibility of MD system suggests a potential mechanism for task switching, implemented by the simultaneous coding of both tasks within the same network. Supporting this idea, a recent decoding study found that MD regions code modality-specific information for visual and auditory tasks within the same voxels (Jackson et al., 2024). These findings highlight the ability of MD system to allocate resources across different tasks while preserving distinct, modality-specific codes, allowing for adaptive responses to varying cognitive demands. Therefore, the flexibility in MD neuron coding may serve as a foundation for multitasking. Future research is needed to explore how multiple task representations coexist within MD regions to address this question.

The decoding analysis also revealed that the modulatory effect of the interrupter task was most pronounced in the left intraparietal sulcus (IPS). This finding aligns with studies that demonstrated the key role of IPS in protecting memory representations against visual distractors (Bettencourt & Xu, 2015; Lorenc et al., 2018; Rademaker et al., 2019). Specifically, Bettencourt & Xu (2015) and Lorenc et al. (2018) proposed that this region maintains a stable memorandum in the face of visual distraction. Furthermore, Rademaker et al. (2019) examined the format of these codes and demonstrated that memory representations in IPS are maintained in a mnemonic format, making it more resilient to disruption from visual distractors. Importantly, these studies highlight the stable maintenance of memory codes as a key mechanism for protecting memory from distortion, demonstrating that decoding in the IPS remains unchanged regardless of the presence or absence of visual distraction. However, this mechanism contrasts with the current findings, which suggest that memory protection from interruption is achieved through enhancement of memory representations, reflected in increased decoding accuracy when the interrupter was more challenging, rather than solely maintaining distraction-resistant representations.

The discrepancy between IPS decoding results in present and previous studies could be potentially attributed to methodological differences. Specifically, in the present study, the interrupter during the delay period acted as interference that may require more cognitive effort for memory protection compared to visual distractors, and, thus, enhancement of memory representations. Previous studies also employed different decoding methods that could uncover distinct aspects of memory representations in the presence of interference, making direct comparisons challenging. Specifically, the Inverted Encoding Models (IEMs) used by Lorenc et al. (2018) and Rademaker et al. (2019) provide insights into specific features (e.g., stimulus orientation) encoded in neural activity, whereas MVPA applied in the current study and the study by Bettencourt & Xu (2015) offers information about the discriminability of neural patterns, but not necessarily about which specific features are being encoded. Using the IEM approach, Rademaker et al. (2019) demonstrated that the IPS maintains memoranda in a mnemonic format distinct from the format of sensory information, which might make it ~~more~~ resistant to distraction. However, the present decoding study did not examine differences in representational formats between the visual and parietal cortices, leaving it unclear whether the IPS maintained and enhanced memory codes for objects in a mnemonic format to protect them from interruption. The primary reason for choosing the MVPA approach over IEMs in this study is that IEMs are best suited for decoding simple features; in contrast, this study examined memory representations for realistic objects, which required the use of MVPA. An alternative approach for capturing the transformation of memory codes of realistic objects is representational similarity analysis (RSA; Kriegeskorte et al., 2008). Specifically, RSA compares patterns of brain activity and measures how similar or dissimilar they are, enabling researchers to assess the underlying structure of neural representations; unlike approaches that decode specific features like IEMs, RSA focuses on the relationships between these patterns, allowing it to capture complex changes in how the brain encodes information, such as shifts from sensory processing to memory-based representations. Namely, RSA measures the dissimilarity between neural representations, assessing how close or far away they are in the representational space, i.e. how similar or dissimilar they are. Therefore, future studies could apply the RSA approach to reveal how interference might change object memory representations. Specifically, this approach may help determine whether memory representations in sensory and parietal regions diverge further in representational space as a mechanism for memory protection, thereby reducing the likelihood of memory code disruption due to interference.

In addition, since the present study used an ROI-based approach to decoding, the parcellations within the IPS (i.e., IPS0, IPS1, IPS2, IPS3, IPS4) were not considered as in previous studies decoding memory representations of orientations in this region (Bettencourt & Xu, 2015; Lorenc et al., 2018). Therefore, future decoding research needs to pinpoint the differences in object memory representations between IPS subregions.

Despite the differences in analysis methods, it is more likely that the greater difficulty of the interrupter task in the present study, compared to visual distraction used by previous decoding research, demanded a greater degree of cognitive control over the memorandum, thereby possibly enhancing its representation further, in order to protect it from interference. Together, the present decoding results and previous research on IPS in working memory (see reviews by Curtis, 2006; Xu, 2018; Cowan, 2011) indicate that this region plays a key role in maintaining and protecting working memory. Whereas the causal findings on the necessity of dlPFC for modulating relevant information within MD system (Jackson et al., 2021) suggest that dlPFC contributes to the enhancement of decoding observed in the current analysis, suggesting a coordinated involvement of both regions in supporting working memory processes. Thus, to investigate the specific role of the IPS in resolving interference in working memory, future studies could explore the directionality of the information flow between dlPFC, IPS and sensory areas as well as beyond visual modality. This question can be addressed by decoding the signal recorded with Electro- or Magnetoencephalography (E/MEG), providing the temporal dynamics of this process. For instance, Goddard et al. (2016) used this approach to investigate how object recognition involves both feedforward (from occipital to frontal regions) and feedback (from frontal back to occipital) flows of information in the brain using neural decoding of MEG signal. Future studies of working memory may adopt this approach to reveal the flow of information during encoding, memory maintenance with and without distractors, and the recall phase, hence, revealing the specific roles of MD network and sensory regions, and their interaction in working memory. For instance, at encoding, memory representations could be transferred from visual cortex to MD regions and then reinstated back in visual cortex at the recall or when a distractor disrupts representations in this region.

Decoding memory representations in the early visual cortex revealed that unlike the coarser memory representations observed in previous studies of visual distractors (Bettencourt & Xu, 2015; Lorenc et al., 2018; Rademaker et al., 2019) and interrupters (Kiyonaga et al., 2017), the degree of interrupter difficulty in the present study did not impact memory representations in the early visual cortex. This finding can be explained by the causal evidence of dlPFC enhancing relevant information in the visual cortex (Jackson et al., 2021) to protect it in the face of interfering stimuli. It can be argued that this region may not be critical for the maintenance of object memorandum since object processing involves multiple areas from early visual areas to higher-level regions (Hebart et al., 2018). However, in the memory task of the present study, participants were asked to memorise object details, as the recall phase involved pairs of similar objects that were difficult to distinguish, requiring them to maintain fine details in memory that presumably involve early visual areas (Coggan et al., 2017). It is possible that we did not observe lower decoding in the early visual cortex since the visual display during the interrupter task was visually distinct from the memory targets; hence, the visual processing of the interrupter task likely involved different neural populations and did not lead to memoranda interference.

Additionally, as was noted in previous work (Lorenc et al., 2018), memory representations may be flexibly adjusted for the task at hand (Machizawa et al., 2012); specifically, when a distractor is expected to appear during the delay period, it appears that this can be mitigated in advance by storing the copies (Lorenc et al., 2018; Bettencourt & Xu, 2015) or changing the format (Rademaker et al., 2019) of the memorandum in the parietal cortex. However, this idea does not fit with the present decoding showing flexible modulation as a response to increasing task demands. If cognitive control modulated the representations in anticipation of interruption, this modulation would be equal irrespective of the interrupter task difficulty. Yet, the opposite was the case: the harder the interrupter task, the more cognitive control was required to protect the memorandum from disruption resulting in the enhancement of memory representations. This result highlights that MD system flexibly allocated representational resources for memorandum as a response to increasing (interrupter) task demands. Interestingly, this could explain the improvement in task performance in the presence of interrupters (Zickerick et al., 2020; Kiyonaga et al., 2017). MD network adapts quickly to the difficulty of the interrupter task, and, thus, overcompensates and deploys even more cognitive resources that are required by the primary task. Future research could further explore these adaptive mechanisms by examining how MD network adjusts its coding strategies in response to varying task demands, particularly under conditions of increased cognitive load or interruptions as well as how MD system adapts to the (high) constant task demands to elicit proactive strategies for stable task performance.

Broadly, the findings expand our understanding of how MD system implements adaptive coding in the brain, specifically by revealing the mechanisms of cognitive control that protect working memory from interruptions. Importantly, understanding the interaction dynamics between MD regions and other brain areas in this process could have profound implications for improving cognitive performance in more distracting environments and creating interventions for individuals with cognitive control deficits (see section Practical Implications of This Work for more details).

Interaction Between Internal and External Distraction: The Role of Individual Differences in Cognitive Control

Mind Wandering and Mitigation of External Distraction in Working Memory

Individuals who are prone to internal distractions such as internally generated thoughts (Smallwood & Schooler, 2015) demonstrate lower performance in working memory tasks (Robison and Unsworth, 2018). Therefore, across two similar experiments in Chapters 2 and 3, I investigated whether individual tendencies to mind-wander (spontaneous or deliberate), which are characterised by increased internal distraction, are related to susceptibility to external distraction in working memory.

First, results revealed that cognitive control over mind wandering is reflected in distinct relationships with external distraction. Namely, the tendency to mind wander deliberately (characterised by higher control over the mind-wandering state compared to individuals who mind wander spontaneously) did not affect the ability to mitigate external distractors in working memory in both experiments.

Second, Chapter 3 revealed that spontaneous mind wandering was linked to better resistance against visual distractors in working memory. Specifically, the more individuals were prone to spontaneous mind wandering, the less dynamic distractors affected their object memory recall. Since participants were less distracted by the dynamic stimuli presented during the delay period, they could be more distracted by the static noise instead. Indeed, many participants reported after the experiment that they liked viewing the abstract colourful shapes that helped them to stay on task, potentially through more effective cognitive control for memory maintenance. Of course, individuals who mind wandered deliberately also could be bored by noise distractors and engage with internally generated thoughts, but because of their efficient cognitive control, it did not interfere with the task performance.

A slightly opposite interpretation may be done in light of Load Theory (Lavie, 1995) which posits that participants perform better on tasks with increased perceptual load because their cognitive efforts are consumed by the task, and they therefore do not have resources for distractor processing (Lavie & Tsal, 1994). Indeed, spontaneous mind-wandering is related to lower working memory capacity (Robison & Unsworth, 2018); this type of mind-wandering can result in lapses in focus, reducing the cognitive resources available for task-related perceptual processing (Robison & Unsworth, 2018). Therefore, individuals prone to spontaneous mind wandering may allocate their cognitive resources towards memory encoding, leaving fewer processing efforts for dynamic distractors. However, they retain enough capacity to process static noise images, which require minimal effort and thus interfere with memory maintenance.

Interestingly, in the present study, the negative relationship between spontaneous mind wandering and the ability to mitigate external distractions disappeared when TMS was applied over dlPFC. Mind wandering is associated with increased activation of the Default Mode Network (DMN), a network of brain regions, including the medial prefrontal cortex, posterior cingulate cortex, and precuneus (Weissman et al., 2006; Christoff et al., 2009). Research suggests that frontoparietal regions, including dlPFC, regulate DMN by inhibiting its activity during task-focused behaviour (Fox et al., 2005; Smallwood et al., 2012). When control is no longer critical, the inhibition weakens allowing the DMN to become more active facilitating mind-wandering. Therefore, TMS applied over dlPFC in the present study could have led to a decrease in inhibition over the DNM evoking more activity within this network and, hence, more episodes of spontaneous thoughts. This potentially led to an equalisation of the distractor effect: the visual features of distracting stimuli became irrelevant, as internally generated thoughts fully consumed cognitive resources diminishing the impact of external distractions.

Furthermore, it is noteworthy to point out that dlPFC is involved in multiple functions, including mitigation of external distraction as discussed in the first section, and, therefore, perturbation of this region could lead to several modulations of performance in the working memory task in the present study. To illustrate this, a recent study used an excitatory intermittent theta burst stimulation (iTBS) TMS protocol to enhance activity in the right dlPFC and continuously monitored participants' state of mind wandering during a finger-tapping random sequence generation task (Aasen et al., 2024). Researchers found that stimulation caused both task performance and the number of mind-wandering episodes to increase suggesting that excitatory stimulation of dlPFC could enhance cognitive resources that support both mind-wandering and task efficiency (Aasen et al., 2024). Therefore, in the present findings, it is challenging to disentangle whether the effects of DLPFC perturbation are specifically related to increased mind wandering or to impaired ability to mitigate distractions.

In addition, in Chapter 2 I found the negative relationship between spontaneous mind wandering and overall reaction time. Individuals who reported being more prone to spontaneous mind wandering recalled memory targets faster than those who reported less mind wandering. Although this finding is counterintuitive, there may be several explanations. First, individuals whose mind wander spontaneously have lower cognitive control and, therefore, potentially, could rely on a more automatic, less deliberate retrieval process without spending too much cognitive effort at the recall and responding faster. Second, distractors appeared in two-thirds of the trials in that experiment and may have helped keep individuals, whose minds tend to wander spontaneously, engaged with the task. Indeed, improvements in the presence of distractors were found in the auditory domain (Nagaraj, 2021) and motor performance (Hemond et al., 2010).

It is important to note that all the interpretations discussed here are not supported by the data in the present study and need to be addressed in future studies. For instance, the studies might use neuroimaging to investigate functional connectivity between the sensory regions responsible for distractor processing and MD regions eliciting cognitive control to examine how volitional control in mind wandering affects this connectivity during memory recall. Namely, if individuals prone to spontaneous mind-wandering rely more on automatic recall, the connectivity between these regions would be weaker than in individuals who mind wander deliberately, and, hence, elicit more control over memory recall. Furthermore, since dlPFC plays an important role in the mitigation of external distraction (as discussed in the first section), there may be multiple mechanisms to achieve that, one of which may be by reducing mind wandering; thus, future studies should examine the interaction of the DMN and MD networks and its relation to distractor mitigation ability. Specifically, studies could explore whether the suppression of the DMN through enhanced MD network engagement helps maintain cognitive control and reduce mind wandering during external distraction in WM.

Finally, mind wandering could be considered an adaptive state of mind which can facilitate creativity by allowing the brain to unconsciously process information, leading to sudden insights or creative solutions (Baird et al., 2011), can promote emotional recovery by allowing reflection and meaning-making (Mooneyham &

Schooler, 2013), and by allowing simulation of future events mind wandering can be a valuable cognitive tool for goal setting and motivation (Seli et al., 2016). And even spontaneous thoughts can be adaptive in learning and real-life problem-solving (Baars, 2010). Therefore, mind wandering requires attention from the working memory research field, with an alternative view being that it is a strategy that our mind employs to deal with the environment we live in, rather than a negative consequence of poor cognitive control.

Trait Worry and Mitigation of External Distraction in Working Memory

Worry can be defined as uncontrollable and excessive thoughts regarding uncertain events in the future, and it is often increased in anxiety (Borkovec et al., 1998). Both worry and rumination, as forms of negative thinking, are significant contributors to mind wandering (Robison et al., 2017). This, in turn, has been shown to impair working memory (Sari et al., 2017; Bruning et al., 2023; Gustavson & Miyake, 2015), indicating reduced cognitive control during memory tasks. To test the role of worry in the ability to elicit cognitive control over memory maintenance during external distraction, in Chapters 2 and 3, I examined the relation between trait worry and the ability to mitigate external distraction.

The findings indicated that the tendency toward excessive worry was unrelated to distractor mitigation ability. However, results from Chapter 2 show a slight trend suggesting that individuals who worry more may experience greater impairment in memory recall precision in the face of external distraction (when static distractors were displayed during memory maintenance compared to unfilled delay). In addition, another trend towards overall memory recall precision and the tendency to worry was found in Chapter 3 (in the absence of TMS). Note, that there were no unfilled delays in the experiment in Chapter 3, a distracting stimulus was always present but varied in the level of visual engagement. Therefore, the overall memory recall precision in this experiment could potentially relate to the ability to mitigate external distraction.

These results can be reconciled by considering Attentional Control Theory (ACT; Eysenck et al., 2007) which posits that anxiety primarily affects processing (the use of cognitive resources), but not processing effectiveness (performance outcomes) suggesting that anxious individuals exert more cognitive effort to achieve the same level of performance as non-anxious individuals (Berggren & Derakshan, 2013). Indeed, this compensatory mechanism can explain the discrepancies in previous studies of trait anxiety and working memory as reviewed by Berggren & Derakshan (2013). Specifically, the cognitive load of the task influences the susceptibility of anxious individuals to distraction; while low and medium loads may not impair performance, high loads can lead to performance decline due to the depletion of cognitive resources (Berggren & Derakshan, 2013). This could potentially explain null findings in the present experiments. Since participants had to memorise only one image, the load of the task was not high enough to consume all the cognitive resources resulting in no significant reduction in memory recall precision for individuals with higher levels of trait worry. However, the trends suggest that distractors could potentially impact the memory performance of these individuals by increasing the cognitive load of the task. Future studies could explore varying cognitive load levels and the role of different types of distractors (e.g., emotional, visual, or auditory) to better understand the threshold at which anxiety impairs performance, and how these factors interact to affect cognitive effort.

Furthermore, no trend was present in the session when TMS was applied over dlPFC in Chapter 3. Since this region plays an important role in anxiety (Basten et al., 2011; Sagliano et al., 2019), one possible interpretation is that suppression of activity in this region may have reduced cognitive control, disrupting the compensatory mechanisms individuals with higher trait worry rely on, causing both types of distractors to negatively impact performance.

Indeed, this possibility is supported by the neuroimaging results demonstrating the increased activation of dlPFC in anxious individuals (Basten et al., 2011). This activation is related to the compensatory mechanism induced by the connections from the anterior cingulate cortex (ACC) which monitors performance and signals dlPFC and other frontoparietal regions (part of MD network) to increase cognitive engagement (Eysenck et al., 2023). Therefore, TMS over dlPFC could disrupt this mechanism and lead to memory interference by both types of external distractors further supporting the necessity of cognitive control over working memory. Future studies should clarify this interpretation by exploring the role of dlPFC in managing resources for resolving both internal distractions, caused by worry, and external distractors in working memory. For instance, analysing dlPFC connectivity with regions implicated in excessive worry and anxiety (Eysenck et al., 2023), along with its interactions with MD network during external distraction, could provide valuable insights into the interplay of these networks and how dlPFC-driven cognitive control effectively manages both types of distractions.

The connection between anxiety and cognitive control highlights the broader role of dlPFC in both anxiety and working memory. Indeed, dlPFC is involved in managing anxiety (Basten et al., 2011) and the activity in this region in highly anxious individuals is increased as a result of engaging compensatory mechanisms (Basten et al., 2011). Consequently, dlPFC has become a key target for TMS-based therapies, which aim to reduce activity in this region to alleviate anxiety symptoms, as reviewed by Sagliano et al. (2019). Interestingly, the link between anxiety and working memory is further supported by the reduction of anxiety symptoms following working memory training, which is considered a promising intervention for managing this condition (Roughan & Hadwin, 2011; Beloe & Derakshan, 2020). Therefore, the role of dlPFC in working memory should also be examined through the lens of emotional regulation in future studies since these types of cognition are tightly intertwined (Pessoa, 2008).

Finally, the study of the link between working memory and worry is important for educational settings since academic performance depends not only on working memory performance but is highly reliant on the emotional regulation of students (Usán Supervía & Quílez Robres, 2021; Pe et al., 2013).

In sum, the present findings point to an intricate connection between external distraction in working memory and trait worry, both managed by the cognitive control exerted by dlPFC. However, further research is required to clarify how this region orchestrates these functions, particularly through a network analysis approach that could reveal the interactions between multiple brain regions involved in these processes.

Pupillometry Application to Tracking Cognitive Control and Effort Allocation in Working Memory During External Distraction

Pupil dilation reflects the current state of arousal and can be used to track cognitive effort during the task at hand (Van der Wel et al., 2018). Specifically, tonic pupil activity reflects overall task engagement and can be used to evaluate fluctuations in cognitive effort with larger fluctuations denoting weaker cognitive control (Unsworth & Robison, 2015; Robison & Unsworth, 2018), whereas phasic pupil activity reflects the immediate cognitive effort (Naber et al., 2013). Phasic pupil response during memory encoding has been shown to predict subsequent memory recall (Cronin et al., 2023; Kucewicz et al., 2018; Miller & Unsworth, 2020; Naber et al., 2013), as well as track active memory maintenance by predicting the number of recalled items (Robison & Unsworth, 2019) or reflecting the content of the maintained memoranda (Zokaei et al., 2019).

Building on this evidence, I used the pupillometry approach in Chapters 2 and 3 to investigate how fluctuations of tonic pupil response (which reflect the strength of cognitive control) predicted distractor mitigation ability in working memory or overall performance memory in working memory. In line with previous studies (Unsworth & Robison, 2015; Robison & Unsworth, 2019), in Chapter 3 I found that fluctuations in tonic pupil response predicted memory recall precision for objects irrespective of distractor type (dynamic or static noise) presented during the delay period. This result demonstrates that cognitive control over the sustained deployment of efforts throughout the experiment appears to be important for overall task performance and managing distractions. Additionally, fluctuations in tonic pupil response predicted memory recall precision even in the session when stimulation was applied over dlPFC, suggesting that dlPFC may not be involved in the type of cognitive control reflected by tonic pupil response. Indeed, pupillary responses are likely regulated by the locus-coeruleus noradrenergic arousal system (LC-NE; Robison et al., 2023) which exerts control over dlPFC (Grueschow et al., 2022; Grueschow et al., 2020). Therefore, future studies need to address the interplay between the LC-NE system and dlPFC in the mitigation of distraction in working memory to reveal the contribution of each control type to this process.

However, in a slightly different paradigm in Chapter 2, no relationship between tonic pupil fluctuation and overall memory performance with or without distractors was found. Although experiments in Chapters 2 and 3 were different, the types of memory targets, the measure of memory performance (object recall precision) and the analysis of pupillary response were the same. Thus, it is hard to attribute this discrepancy to the differences between the studies, and further investigation is needed to replicate the result found in Chapter 3. Note that in previous studies, pupil fluctuations were found to predict memory capacity (Unsworth & Robison, 2015; Robison & Unsworth, 2019), a measure of working memory performance distinct from the memory recall precision used in the present work. This highlights the differences in how working memory performance is assessed, with memory capacity focusing on the quantity of information retained, while recall precision emphasizes the accuracy of the details remembered. Therefore, these differences could lead to distinct results.

Despite the discrepancy we found, the findings in Chapter 3 offer further evidence supporting the use of pupillometry to track fluctuations in cognitive effort through an alternative approach—measuring working memory performance via recall precision for naturalistic objects. This encourages further exploration of tonic pupil fluctuations in capturing individual differences in cognitive control during working memory tasks, particularly through the use of continuous recall paradigms.

While phasic pupil response to a stimulus may reflect the depth of its processing as indicated by previous pupillometry studies (Naber et al., 2013; Koevoet et al., 2023). To evaluate the amount of cognitive effort exerted on distractor processing, in Chapter 2, I compared the phasic pupil response to different types of distracting stimuli (in terms of the strength of their visual engagement) presented during the delay period. The results revealed that all distractor types – fixation cross, static or dynamic abstract vibrant shapes – evoked the same phasic pupil response. This finding was surprising since these stimuli are very different in their visual features and would presumably require different amounts of cognitive effort for visual processing. However, the distractors did not lead to lower object recall precision, suggesting that cognitive resources were not significantly consumed by their further processing, also supported by the pupillometry findings. This might not be surprising since participants were instructed to ignore any distractors and no response to them was required. Indeed, this lack of a significant pupillary response across distractors is consistent with the findings in this thesis and previous studies in which relevant information is the focus of cognitive control (Feredoes et al., 2011; Jackson et al., 2018). When faced with distraction, cognitive control may minimise interference with memory by deploying control to relevant memory targets and minimising further processing of distractors.

Furthermore, in Chapter 3, I examined the effect of dlPFC TMS on phasic pupil responses to distractors and found that stimulation reduced pupil dilation during the later stage of the distractor phase (2500 ms after distractor onset), which likely reflected active memory maintenance (Beatty, 1982; Robison & Unsworth, 2019; Strauch et al., 2022; Zhou et al., 2022), regardless of distractor type. In addition, TMS over dlPFC did not impact phasic pupil response at the earlier stages of the distractor phase. These findings suggest that suppressing dlPFC-driven cognitive control primarily reduced the overall cognitive resources allocated to memory maintenance, but not the processing of distracting stimuli. This result aligns with the neural decoding evidence discussed earlier: MD regions, including dlPFC, exert cognitive control over memory representations in the face of interference by enhancing those representations, potentially reflected in the sustained pupil response during the delay period. When dlPFC activity was suppressed with an inhibitory TMS protocol, this enhancement was diminished, leaving memory representations more vulnerable to interference.

Previous studies propose that phasic pupil response during the delay period reflects active memory maintenance (Robison & Unsworth, 2019; Zokaei et al., 2019). Therefore, in Chapter 2, I applied this idea to assess the impact of visual distractors on memory maintenance by comparing phasic pupil responses during the delay period before and after the distractor presentation. The results revealed that the presentation of different distracting stimuli did not interfere with the phasic pupil response during the delay period. Since the distractors in that study did not impact memory recall precision, it is expected that they did not impact pupil response corresponding to memory maintenance. Additionally, I found that the pupil response increased by the end of the delay period which could reflect an increase in cognitive effort required to actively maintain memory as a preparation for subsequent recall (Piquado et al., 2010).

Altogether, pupil response serves as a valuable physiological marker for understanding how cognitive resources can be allocated during working memory, particularly when distractions are present. While fluctuations in cognitive effort may be linked to memory precision, further research is required to validate this connection across a broader range of working memory tasks, especially those using continuous recall measures that can report memory precision. Pupillometry can be used to track the online allocation of cognitive resources in working memory, and in resolving distractions, in particular. Future studies should test these ideas by examining phasic pupil responses to distracting stimuli that specifically impact memory performance or recall precision.

Practical Implications of This Work

To begin, the present findings have several practical implications for clinical application. First, understanding how dlPFC manages cognitive control can lead to targeted interventions for conditions characterised by poor working memory and high distractibility, such as attention deficit hyperactivity disorder (ADHD; Holmes et al., 2014) or schizophrenia (Forbes et al., 2009). Techniques such as cognitive training (Holmes et al., 2014; Forbes et al., 2009) or neuromodulation, e.g., TMS (Luber & Lisanby, 2014) could be optimised to enhance dlPFC function and improve cognitive control and memory; for instance, by applying excitatory TMS protocols (e.g. 5 Hz) over dlPFC during cognitive training involving distraction in working memory will enhance the effect of both interventions. For conditions like Alzheimer's disease or other forms of dementia in which working memory is compromised, understanding the neural mechanisms of dlPFC-driven cognitive control could aid in developing targeted diagnostic and treatment interventions, potentially focusing on enhancing cognitive control mechanisms by combining neuromodulation and cognitive training (Vecchio et al., 2022; Thams et al., 2020). Importantly, research on the role of MD network in cognitive control is vital for advancing these interventions, as further studies on network interactions are necessary to deepen our understanding of these complex neurological conditions for more efficient clinical interventions (Sporns, 2013). In addition, by identifying biomarkers specific to MD network dysfunction, researchers could potentially detect early signs of cognitive decline in neuropsychiatric conditions which affect cognitive control; this could lead to earlier, more proactive treatments, such as combined cognitive and neuromodulation therapies, aimed at preserving MD network integrity and preventing further cognitive decline.

Since the ability to protect working memory from constant interference in our environment is key to successful performance in studying, work and everyday life, the present results also offer valuable insights for enhancing performance in these areas. For instance, we know that interruptions impact working memory more than distractors, therefore we need to minimise this specific type of interference in our environment, e.g. by switching off the phone when finishing an important piece of work because the messages will disrupt your focus, or going to the library to study if you know that your cat jumps onto your laptop and prevents you from preparing for exams. If you know you are prone to mind wandering, you could make your tasks slightly more challenging to help maintain focus and stay engaged with them more easily. For instance, when revising a topic for an exam, do not just read through all the materials, but introduce quizzes or explain difficult concepts to a classmate to make the process more engaging and challenging at the same time. The finding on increased worry being related to lower cognitive control highlights the importance of combining working memory assessments with emotional well-being in contexts reliant on working memory, such as educational settings. In everyday life, we may notice how excessive worrying consumes our energy and, thus, try to direct our cognitive efforts on the task at hand to reduce the impact of worry on working memory.

As outlined in this thesis, cognitive control plays a central role in regulating various cognitive processes. Consequently, enhancing cognitive control can result in widespread improvements across multiple cognitive functions. Exercising mental practices such as mindfulness and meditation for this purpose recently received a lot of attention. For instance, two-week training in mindfulness, i.e. the ability to focus on the present moment (Black, 2011), improved working memory performance and scores on university examinations in undergraduate students (Mrazek et al., 2013). Furthermore, neuroimaging revealed that in experienced individuals with mental practices, both neural efficiency and brain plasticity are enhanced, leading to improved cognitive control and resulting in advancements across various cognitive domains (Slagter et al., 2011; Bailey et al., 2020).

Future Directions

Interaction of MD Regions in Working Memory Protection

The decoding results of the present work revealed that interrupter task difficulty modulated memory representations in MD regions. However, this modulation differed in its magnitude (decoding accuracy percentage) suggesting that these regions may play distinct roles in memory protection. Therefore, future analysis of this data will aim to evaluate these roles by revealing interactions among them. Specifically, Dynamic Causal Modelling (DCM) (Friston et al., 2003) will be applied to examine the *effective connectivity* between MD regions and their connection with the early visual cortex. This type of connectivity models the causal influence between hidden neuronal states, describing how the activity in one brain region affects another over time, and providing the directionality of this influence (Friston et al., 2003). Comparing the models of interactions among MD regions, and the early visual areas between each of the interrupter task conditions (varied in difficulty) will elucidate how network dynamics change as task demands increase. One prediction would be that the difficulty of the interrupter task would increase the connection strength of prefrontal MD regions with parietal and sensory regions, as a mechanism of cognitive control elicited to protect a memorandum from interference. Monkey neurophysiology studies offer more nuanced insights into the mechanisms of distractor mitigation, specifically highlighting the communication between frontal regions—particularly the dlPFC—and parietal regions (Suzuki & Gottlieb, 2013; Jacob & Nieder, 2014). Specifically, the dlPFC may implement cognitive control by modulating memory representations to align with the task at hand. It achieves this by preserving low-dimensional memory codes in the presence of distractions (Parthasarathy et al., 2017; Parthasarathy et al., 2019) and enhancing coupling with posterior regions that maintain robust memory representations (Suzuki & Gottlieb, 2013). These representations may then be retrieved and used by the dlPFC to adapt dynamically to current task demands. Yet, more studies are needed to confirm this mechanism in the human population. Moreover, while these studies have focused on memory protection from distractors, further research is needed to investigate the frontoparietal interaction in mitigating the impact of interrupters on WM. The planned analysis of the acquired dataset will offer a valuable contribution to advancing our understanding of this topic.

Another sophisticated approach that can be applied to the current dataset is multivariate pattern dependence (MVPD), which is a method for analysing the statistical relationships between brain regions by examining the multivariate relations between their patterns of responses (Anzellotti & Coutanche, 2018). Applying this approach will provide insights into how working memory representations are transferred between regions of MD network and visual cortex during interference.

Future studies could address the dynamics of the representational flow between MD regions and sensory cortices by adopting the approach used by Goddard et al., (2016). Namely, the study used MEG combined with multivariate pattern analysis to measure how object-related information is processed in peri-occipital and peri-frontal areas of the brain and found that the occipital cortex coded object identity as early as 80 ms after stimulus onset, while representations in the frontal regions emerged at 265 ms (Goddard et al., 2016). Such temporal resolution of representational flow between areas will help to understand how MD regions and sensory areas code information at memory encoding, during the delay period and at recall, but also when and through which regions the memory codes are enhanced in the presence of an interfering stimulus or task.

Since distractors and interrupters appear to affect working memory in different ways (Clapp et al., 2011), distinct cognitive control mechanisms may be engaged by MD system, in keeping with its adaptive role. Future research could apply the decoding approach described above to investigate the representational dynamics of MD network and sensory cortices during both interruptions and external distractions in working memory, helping to clarify these potentially distinct mechanisms.

How Regions and Networks Outside MD System Contribute to Working Memory Protection from Interference

Despite the central role of dlPFC in eliciting cognitive control over a memorandum to protect it from interference, this control is implemented by interaction not only within MD regions but also with other brain areas outside MD system. The results from the current work support this idea. First, the present neurostimulation finding suggests that dlPFC impacts (spontaneous) mind-wandering, potentially by eliciting inhibitory control (Smallwood et al., 2012) over DMN which engages during mind-wandering episodes (Weissman et al., 2006; Christoff et al., 2009). Since mind-wandering impacts working memory and its protection, future research should address the question of how the MD and DMN interact during interference in working memory. Namely, future neuroimaging studies could explore whether immersion in internally generated thoughts shields cognitive processing from external distraction. This could be done by assessing DMN involvement and individuals' ability to mitigate distractions in working memory tasks on a trial-by-trial basis (Zhang et al., 2022). Second, since pupil response reflecting the active memory maintenance during the delay period was also modulated, i.e. decreased, by dlPFC stimulation, it would be interesting to investigate the interaction between dlPFC-driven cognitive control and the locus-coeruleus noradrenergic arousal system (LC-NE) that is likely to be the source of control over the state of arousal reflected in pupil dilation (Robison et al., 2023). Previous research found that the LC-NE system exerts control over dlPFC when more cognitive resources must be deployed over a task at hand (Grueschow et al., 2022; Grueschow et al., 2020). In contrast, the dlPFC has also been shown to mediate the activation of the LC-NE system, suggesting a bidirectional relationship (Tomassini et al., 2022). Thus, exploring the interplay of these two control systems during working memory protection may shed light on how the allocation of cognitive resources is managed during this process.

Second, previous research points to another mechanism of protecting working memory by gating the irrelevant information from interfering with the memorandum, in which basal ganglia (BG) has been proposed to act as a gate for selecting what information is stored (input gating), what is retrieved for use (output gating, and what is cleared from working memory reallocation (Chatham & Badre, 2015). This process is implemented by close interaction of dlPFC and BG where BG gates the access of dlPFC to information, ensuring cognitive flexibility and the ability to focus on relevant tasks (D'Ardenne et al., 2012). However, future research needs to address the precise interplay of these regions during interruption and external distraction since these types of interference may require different mechanisms of memory protection (Clapp et al., 2010). BG is also involved in affective processing (Pierce & Péron, 2020); thus, future studies may explore how the interplay of dlPFC and BG resolves distractions with affective content in working memory. Importantly, the role of dlPFC-BG interaction in working memory protection should be tested in individuals with BG dysfunction (Hallett, 1993; Utter & Basso, 2008) to provide unique insights into the input gating mechanism and reveal what might be the compensatory mechanisms supporting the working memory of these individuals.

Another region that contributes to working memory and, possibly its protection, is the hippocampus. Although this region is considered crucial for long-term memory (Jeneson & Squire, 2012), recent evidence points to its involvement in working memory as discussed by Husain (2024) and Leszczynski (2011). Namely, hippocampus might support working memory through binding together relational and spatial information, helping to maintain complex associations between objects and their locations (Husain, 2024; Leszczynski, 2011). Yet, how these functions help to mitigate the impact of interference needs to be addressed by further studies. In addition, another structure of the medial temporal lobe (MTL) - parahippocampal gyrus - is involved in memory protection from interference (Sakai & Passingham, 2004). Researchers demonstrated a double dissociation between parahippocampal gyrus and dlPFC. Specifically, activity within dlPFC increased in response to greater interference but not for reactivating stored information, while activity within the parahippocampal gyrus increased for memory reactivation but remained unaffected by interference levels (Sakai & Passingham, 2004). However, it is unclear how these regions interact, and if dlPFC has control over this reactivation; hence, more research is needed to clarify this mechanism.

These ideas collectively highlight future directions for investigating the intricate networks of brain regions, some of which are not considered 'traditional' working memory areas, that interact with dlPFC to resolve interference in working memory, emphasizing the need for a holistic network approach to account for the complexity of this process.

Advancing the Study of Interference within Working Memory: Future Prospects

In this section, I will explore several directions for future research on interference mitigation in working memory, along with insights from my work.

Moving Towards Ecological Validity

How often do you need to memorise the exact colour shade of an object? Or a certain orientation of a line? Perhaps, not very often. Yet, research on working memory is largely focused on testing participants' memory for these simple features, while real-life objects are more complex. While these measures allow precise and systematic control over variables, they are potentially at some distance from the information we encounter in 'real life'. Indeed, meaningful memory targets, i.e. realistic complex objects, are thought not to be memorised the same way as simple features due to semantic information which enriches a memorandum. For instance, Brady et al. (2022) demonstrated that WM capacity for real-life objects is higher than for simple features such as coloured squares or shapes that are typically used as memory stimuli. Researchers suggest that real-world objects carry more detailed and conceptually rich information compared to simple stimuli; this additional information facilitates the ability to distinguish between different objects and maintain them over time in working memory (Brady et al., 2022). Veldsman et al. (2017) provided further support for this idea by using neural decoding to compare recognisable versus non-recognisable morphed objects. The study found that recognisable objects were recalled with higher precision and produced more variable neural patterns across trials compared to unrecognisable objects (Veldsman et al., 2017). This indicates that recognisable objects might be supported by a richer and more diverse set of neural representations, even though these differences did not result in stronger or additional brain region activity (Veldsman et al., 2017). Therefore, the use of naturalistic memory targets could lead to a more accurate understanding of how working memory is represented in the brain and how it functions in everyday life.

To further demonstrate the necessity of more realistic measures of working memory, researchers investigated the relationship between working memory and natural behaviour, specifically during an object-copying task in a virtual reality (VR) environment (Draschkow et al. 2021). Participants were required to copy object arrangements by selecting objects from a resource area and placing them into a workspace, while periodically referencing a model display for guidance. The model's location varied, requiring different levels of locomotive effort to view, allowing researchers to measure how much participants relied on holding instructions in working memory versus accessing them externally in order to complete the task. The results revealed that reliance on working memory was surprisingly low: participants preferred to gather information externally from the environment rather than relying on their WM; this was true even during effortful tasks that required locomotion and in spite of their cognitive capacity to hold more information (Draschkow et al. 2021). Importantly, these insights offer a deeper understanding of how working memory is employed in real-world tasks, contrasting with traditional lab-based studies that may overestimate the actual use of working memory capacity (Williams & Störmer, 2021).

This approach has also been applied to the issue of distraction in working memory. In a study by Stokes et al. (2022), a VR classroom environment was created, with concurrent eye-tracking to study how children with Attention-Deficit Hyperactivity Disorder (ADHD) performed various tasks, such as school-like mathematical exercises, the Stroop test, and the continuous performance test (CPT) while facing classroom-like distractions (e.g., a pedestrian passing by the window, a student avatar sneezing, or a phone ringing on a student's desk) (Stokes et al., 2022). The presence of these distractors significantly disrupted on-task eye gaze and impaired task performance. Children with ADHD struggled to return to tasks after interruptions, leading to prolonged disengagement, even when the distractors were brief or did not fully capture their attention (Stokes et al., 2022). Therefore, VR environment can provide a realistic environment to capture real-world distractions and study their impact on working memory in individuals with ADHD.

In another study, researchers explored how visual distractions impact various cognitive processes, such as encoding, visual search, working memory usage, and decision-making in a VR environment. Participants were asked to replicate a model display by selecting objects from a resource pool of 24 cubes (Kumle et al., 2024). Eight cubes matched the target objects in the model display, while the remaining 16 served as distractors with varying levels of distractibility. Some distracting cubes were more transparent and easily distinguishable from the targets, while others had the same opacity as the target objects (Kumle et al., 2024). The study found that difficulty managing distractions led to decreased reliance on working memory, which in turn slowed task performance and increased the physical effort required to complete the object-copying task.

Notably, cognitive tests conducted in VR environments with immersive, colourful backgrounds and 3D depth have been shown to yield results comparable to simpler cognitive tasks performed on computers (Redlinger et al., 2022). This indicates that VR-based assessments could serve as reliable tools for investigating working memory and mitigating distractions in future research.

The Importance of an Individual Differences Approach in Exploring Working Memory Protection

The ability to resolve interference in working memory is different among individuals. These differences are insightful for revealing the complex mechanisms involved in working memory (Jarrold & Towse, 2006). Yet, investigations on interference in working memory often overlook these differences by averaging the responses across participants. In particular, neuroimaging research often averages the responses across participants for a better signal-to-noise ratio of the data acquired, thus not accounting for the individual differences which are also reflected in brain activity and can provide valuable insights on the neural mechanisms underlying working memory. Therefore, more methods to reveal the neural mechanisms in individual sets may help to advance this field.

Experiments discussed in this thesis illustrate how individual differences in cognitive control reflected in excessive mind-wandering impacted the ability to mitigate visual distractors in working memory. Spontaneous (but not deliberate) mind wandering had a surprising protective effect against distraction. These individuals could recall objects with greater precision when more visually engaging stimuli were presented on the screen during the delay period suggesting that these stimuli helped them to stay on task. Such counterintuitive findings reveal how task load and individuals' cognitive control interact with performance in working memory.

The link between worry and distractor mitigation ability was not confirmed but suggested by a trend in the results. However, due to compensatory mechanisms of individuals prone to excessive worry, this connection is not always observed in working memory performance (Berggren & Derakshan, 2013). Given that worry often accompanies anxiety, affecting the daily lives of millions (World Health Organization, 2023), it is crucial to understand the compensatory mechanisms these individuals employ. This knowledge can help develop more effective interventions to alleviate symptoms, e.g. working memory training (Wang et al., 2023)), and enhance performance in daily tasks, as well as shed light on how these mechanisms interact with the ability to maintain and protect information in working memory.

Another example of individual differences impacting performance in working memory is the use of strategies (McNamara & Scott, 2001). Strategies play a significant role in enhancing task performance through efficient use of cognitive resources. Research suggests that individuals who use effective strategies, such as chunking, grouping, or rehearsal, can improve their working memory capacity and performance on complex cognitive tasks (Dunning & Holmes, 2014). Participants are also likely to employ different strategies to deal with interference in working memory, but these strategies are largely unexplored. Identifying them can pinpoint those which result in better task performance, and reveal how the brain strategically controls available neural resources to improve the resolution of interference in working memory.

It is noteworthy that there are multiple factors playing a role in individual differences in working memory performance such as motivation, alertness, mood, and personality traits (Robison et al., 2020). However, future studies should approach this challenge to uncover the intricate links underlying our ability to hold and protect information in working memory. Importantly, the link between individual differences and working memory discussed above is bi-directional; individual differences, e.g. tendency to worry or mind wander, impact working memory, but improving working memory through training and using efficient strategies could improve those cognitive functions.

No Distractor Effect and Improvement Following Interruption Are Not Null Results

It is well acknowledged that interference is a significant cause of forgetting information from working memory (Lorenc et al., 2021). However, in the laboratory environment, by virtue of the need to control for as many variables as possible, and to maximise statistical detection of experimental manipulations, the impact of introducing interference in a lab setting often has the effect of reducing its efficacy. For example, to have enough trials to reliably detect an effect of a distractor on task performance, there needs to be many repeated trials with distractors, which has the effect of making the distraction predictable and easier to deal with. Hence, resolving distractions within an experimental task at hand is easier than performing real-life tasks where distraction can is often varied and unpredictable. In addition, participants are not naïve to the experiment because they are informed about the study aim of investigating interference, and can therefore adopt a strategy accordingly. Thus, more cognitive control can be exerted for a task which may explain improvements with interruptions that have been observed (Zickerick et al., 2022; Kiyonaga et al., 2017) or no effect of distraction in working memory (Rademaker et al., 2019; Yoon et al., 2006; Postle et al., 2004). It is therefore important to consider that these *null* results might be reflecting underlying *successful* cognitive control strategies for protecting working memory from interference, which would benefit from further investigation from this perspective.

Notably, the null behavioural results I described are particularly useful along with neuroimaging or neurostimulation since they can provide insights into the brain mechanisms involved in successfully dealing with interference. For instance, in the present work, visual distractors did not impair the object recall precision. However, when the activity in dlPFC was suppressed, memory precision dropped in the presence of distractors. This example illustrates that participants were successfully dealing with distraction, which impacted memorandum only when dlPFC-driven cognitive control was reduced.

Conclusion

In conclusion, this thesis has provided insights into the neural mechanisms of resolving interference – external distraction or interruption – in working memory, particularly focusing on the role of cognitive control. I demonstrated that the dlPFC plays a causal role in mitigating external distraction in working memory. This supports the established understanding of its involvement in cognitive control processes. Additionally, by examining the involvement of MD network, this research has expanded current knowledge by showing that memory representations in MD regions are enhanced by task difficulty, highlighting the potentially key role of MD system in managing interruptions. This research offers significant potential for both clinical applications and improving the daily lives of healthy individuals. In a wider perspective, this work elucidates the neural underpinnings of the cognitive control mechanisms of maintaining flexible yet stable behaviour, essential for decision-making, reasoning, and adapting to changing environments.

References

Anzellotti, S., & Coutanche, M. N. (2018). Beyond functional connectivity: investigating networks of multivariate representations. Trends in Cognitive Sciences, 22(3), 258269.

Assem, M., Glasser, M. F., Van Essen, D. C., & Duncan, J. (2020). A domain general cognitive core defined in multimodally parcellated human cortex. Cerebral Cortex, 30(8), 43614380.

Badre, D. (2008). Cognitive control, hierarchy, and the rostro–caudal organization of the frontal lobes. Trends in Cognitive Sciences, 12(5), 193200.

Bailey, N. W., Freedman, G., Raj, K., Spierings, K. N., Piccoli, L. R., Sullivan, C. M., ... & Fitzgerald, P. B. (2020). Mindfulness meditators show enhanced accuracy and different neural activity during working memory. Mindfulness, 11, 17621781.

Basten, U., Stelzel, C., & Fiebach, C. J. (2011). Trait anxiety modulates the neural efficiency of inhibitory control. Journal of Cognitive Neuroscience, 23(10), 31323145.

Beatty, J. (1982). Task evoked pupillary responses, processing load, and the structure of processing resources. Psychological Bulletin, 91(2), 276292.

Beloe, P., & Derakshan, N. (2020). Adaptive working memory training can reduce anxiety and depression vulnerability in adolescents. Developmental Science, 23(4), e12831.

Berggren, N., & Derakshan, N. (2013). Attentional control deficits in trait anxiety: Why you see them and why you don't. Biological Psychology, 92(3), 440446.

Bettencourt, K. C., & Xu, Y. (2016). Decoding the content of visual shortterm memory under distraction in occipital and parietal areas. Nature Neuroscience, 19(1), 150157.

Black, D. S. (2011). A brief definition of mindfulness. Behavioral Neuroscience, 7(2), 109.

Bruning, A. L., Mallya, M. M., & LewisPeacock, J. A. (2023). Rumination burdens the updating of working memory. Attention, Perception, & Psychophysics, 85, 1452–1460.

Chao, L. L., & Knight, R. T. (1998). Contribution of human prefrontal cortex to delay performance. Journal of Cognitive Neuroscience, 10(2), 167177.

Chatham, C. H., & Badre, D. (2015). Multiple gates on working memory. Current Opinion in Behavioral Sciences, 1, 2331.

Coggan, D. D., Allen, L. A., Farrar, O. R., Gouws, A. D., Morland, A. B., Baker, D. H., & Andrews, T. J. (2017). Differences in selectivity to natural images in early visual areas (V1–V3). Scientific Reports, 7(1), 2444.

Cowan, N. (2011). The focus of attention as observed in visual working memory tasks: Making sense of competing claims. Neuropsychologia, 49(6), 14011406.

Cronin, S. L., Lipp, O. V., & Marinovic, W. (2023). Pupil dilation during encoding, but not type of auditory stimulation, predicts recognition success in face memory. Biological Psychology, 178, 108547.

Curtis, C. E. (2006). Prefrontal and parietal contributions to spatial working memory. Neuroscience, 139(1), 173180.

D'Ardenne, K., Eshel, N., Luka, J., Lenartowicz, A., Nystrom, L. E., & Cohen, J. D. (2012). Role of prefrontal cortex and the midbrain dopamine system in working memory updating. Proceedings of the National Academy of Sciences, 109(49), 1990019909.

D'Esposito, M., & Postle, B. R. (2015). The cognitive neuroscience of working memory. Annual Review of Psychology, 66(1), 115142.

Draschkow, D., Kallmayer, M., & Nobre, A. C. (2021). When natural behavior engages working memory. Current Biology, 31(4), 869–874.

Duncan, J. (2010). The multipledemand (MD) system of the primate brain: mental programs for intelligent behaviour. Trends in Cognitive Sciences, 14(4), 172179.

Dunning, D. L., & Holmes, J. (2014). Does working memory training promote the use of strategies on untrained working memory tasks?. Memory & Cognition, 42, 854862.

Esslinger, C., Schuler, N., Sauer, C., Gass, D., Mier, D., Braun, U., ... & Meyer Lindenberg, A. (2014). Induction and quantification of prefrontal cortical network plasticity using 5 Hz rTMS and fMRI. Human Brain Mapping, 35, 140–151.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. Emotion, 7, 336–353.

Eysenck, M. W., Moser, J. S., Derakshan, N., Hepsomali, P., & Allen, P. (2023). A neurocognitive account of attentional control theory: how does trait anxiety affect the brain's attentional networks?. Cognition and Emotion, 37(2), 220237.

Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. Proceedings of the National Academy of Sciences, 110(41), 1661616621.

Feredoes, E., Heinen, K., Weiskopf, N., Ruff, C., & Driver, J. (2011). Causal evidence for frontal involvement in memory target maintenance by posterior brain areas during distracter interference of visual working memory. Proceedings of the National Academy of Sciences, 108(42), 1751017515.

Forbes, N. F., Carrick, L. A., McIntosh, A. M., & Lawrie, S. M. (2009). Working memory in schizophrenia: a metaanalysis. Psychological Medicine, 39(6), 889905.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proceedings of the National Academy of Sciences, 102(27), 96739678.

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. NeuroImage, 19(4), 12731302.

Goddard, E., Carlson, T. A., Dermody, N., & Woolgar, A. (2016). Representational dynamics of object recognition: Feedforward and feedback information flows. NeuroImage, 128, 385397.

Grueschow, M., Kleim, B., & Ruff, C. C. (2020). Role of the locus coeruleus arousal system in cognitive control. Journal of Neuroendocrinology, 32(12), e12890.

Grueschow, M., Kleim, B., & Ruff, C. C. (2022). Functional coupling of the locus coeruleus is linked to successful cognitive control. Brain Sciences, 12(3), 305.

Gustavson, D. E., & Miyake, A. (2016). Trait worry is associated with difficulties in working memory updating. Cognition and Emotion, 30(7), 12891303.

Hallenbeck, G. E., Sprague, T. C., Rahmati, M., Sreenivasan, K. K., & Curtis, C. E. (2021). Working memory representations in visual cortex mediate distraction effects. Nature Communications, 12(1), 4714.

Hallett, M. (1993). Physiology of basal ganglia disorders: an overview. Canadian Journal of Neurological Sciences, 20(3), 177183.

Hebart, M. N., Bankson, B. B., Harel, A., Baker, C. I., & Cichy, R. M. (2018). The representational dynamics of task and object processing in humans. Elife, 7, e32816.

Hemond, C., Brown, R. M., & Robertson, E. M. (2010). A distraction can impair or enhance motor performance. Journal of Neuroscience, 30(2), 650654.

Holmes, J., Hilton, K. A., Place, M., Alloway, T. P., Elliott, J. G., & Gathercole, S. E. (2014). Children with low working memory and children with ADHD: same or different?. Frontiers in Human Neuroscience, 8, 976.

Husain, M. (2024). The human hippocampus contributes to shortterm memory. Brain, 147(8), 25932594.

Jackson, J. B., Feredoes, E., Rich, A. N., Lindner, M., & Woolgar, A. (2021). Concurrent neuroimaging and neurostimulation reveals a causal role for dlPFC in coding of taskrelevant information. Communications Biology, 4(1), 588.

Jackson, J. B., & Woolgar, A. (2018). Adaptive coding in the human brain: Distinct object features are encoded by overlapping voxels in frontoparietal cortex. Cortex, 108, 2534.

Jackson, J., Rich, A. N., Moerel, D., Teichmann, L., Duncan, J., & Woolgar, A. (2024). Domain general frontoparietal regions show modalitydependent coding of auditory and visual rules. bioRxiv, 202403.

Jackson, J., Rich, A. N., Williams, M. A., & Woolgar, A. (2017). Featureselective attention in frontoparietal cortex: multivoxel codes adjust to prioritize taskrelevant information. Journal of Cognitive Neuroscience, 29(2), 310321.

Jeneson, A., & Squire, L. R. (2012). Working memory, longterm memory, and medial temporal lobe function. Learning & Memory, 19, 1525.

Jarrold, C., & Towse, J. N. (2006). Individual differences in working memory. Neuroscience, 139(1), 3950.

Koevoet, D., Naber, M., Strauch, C., Somai, R. S., & Van der Stigchel, S. (2023). Differential aspects of attention predict the depth of visual working memory encoding: Evidence from pupillometry. Journal of Vision, 23(6), 9.

Kiyonaga, A., Dowd, E. W., & Egner, T. (2017). Neural representation of working memory content is modulated by visual attentional demand. Journal of Cognitive Neuroscience, 29(12), 20112024.

Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008). Representational similarity analysisconnecting the branches of systems neuroscience. Frontiers in Systems Neuroscience, 2, 249.

Kucewicz, M. T., Dolezal, J., Kremen, V., Berry, B. M., Miller, L. R., Magee, A. L., ... & Worrell, G. A. (2018). Pupil size reflects successful encoding and recall of memory in humans. Scientific Reports, 8(1), 4949.

Leszczynski, M. (2011). How does hippocampus contribute to working memory processing?. Frontiers in Human Neuroscience, 5, 168.

Lorenc, E. S., Mallett, R., & LewisPeacock, J. A. (2021). Distraction in visual working memory: Resistance is not futile. Trends in Cognitive Sciences, 25(3), 228239.

Lorenc, E. S., Sreenivasan, K. K., Nee, D. E., Vandenbroucke, A. R., & D'Esposito, M. (2018). Flexible coding of visual working memory representations during distraction. Journal of Neuroscience, 38(23), 52675276.

Luber, B., & Lisanby, S. H. (2014). Enhancement of human cognitive performance using transcranial magnetic stimulation (TMS). NeuroImage, 85, 961970.

Machizawa, M. G., Goh, C. C., & Driver, J. (2012). Human visual shortterm memory precision can be varied at will when the number of retained items is low. Psychological Science, 23(6), 554559.

McNamara, D. S., & Scott, J. L. (2001). Working memory capacity and strategy use. Memory & Cognition, 29, 1017.

Mrazek, M. D., Franklin, M. S., Phillips, D. T., Baird, B., & Schooler, J. W. (2013). Mindfulness training improves working memory capacity and GRE performance while reducing mind wandering. Psychological Science, 24(5), 776781.

Naber, M., & Nakayama, K. (2013). Pupil responses to high level image content. Journal of Vision, 13(6), 7.

Nagaraj, N. K. (2021). Effect of auditory distraction on working memory, attention switching, and listening comprehension. Audiology Research, 11(2), 227243.

Oliveri, M., Turriziani, P., Carlesimo, G. A., Koch, G., Tomaiuolo, F., Panella, M., & Caltagirone, C. (2001). Parietofrontal interactions in visual object and visual spatial working memory: Evidence from transcranial magnetic stimulation. Cerebral Cortex, 11, 606–618.

Pe, M. L., Raes, F., & Kuppens, P. (2013). The cognitive building blocks of emotion regulation: Ability to update working memory moderates the efficacy of rumination and reappraisal on emotion. PLOS One, 8(7), e69071.

Pessoa, L. (2008). On the relationship between emotion and cognition. Nature Reviews Neuroscience, 9(2), 148158.

Pierce, J. E., & Péron, J. (2020). The basal ganglia and the cerebellum in human emotion. Social Cognitive and Affective Neuroscience, 15(5), 599613.

Piquado, T., Isaacowitz, D., & Wingfield, A. (2010). Pupillometry as a measure of cognitive effort in younger and older adults. Psychophysiology, 47(3), 560569.

Postle, B. R., Ferrarelli, F., Hamidi, M., Feredoes, E., Massimini, M., Peterson, M., Alexander, A., & Tononi, G. (2006). Repetitive transcranial magnetic stimulation dissociates working memory manipulation from retention functions in the prefrontal, but not posterior parietal, cortex. Journal of Cognitive Neuroscience, 18, 1712–1722.

Rademaker, R. L., Chunharas, C., & Serences, J. T. (2019). Coexisting representations of sensory and mnemonic information in human visual cortex. Nature Neuroscience, 22(8), 13361344.

Robison, M. K., & Unsworth, N. (2018). Cognitive and contextual correlates of spontaneous and deliberate mindwandering. Journal of Experimental Psychology: Learning, Memory, and Cognition, 44(1), 85.

Robison, M. K., Gath, K. I., & Unsworth, N. (2017). The neurotic wandering mind: An individual differences investigation of neuroticism, mindwandering, and executive control. Quarterly Journal of Experimental Psychology, 70(4), 649663.

Robison, M. K., Miller, A. L., & Unsworth, N. (2020). A multifaceted approach to understanding individual differences in mindwandering. Cognition, 198, 104078.

Roughan, L., & Hadwin, J. A. (2011). The impact of working memory training in young people with social, emotional and behavioural difficulties. Learning and Individual Differences, 21(6), 759764.

Sagliano, L., Atripaldi, D., De Vita, D., D'Olimpio, F., & Trojano, L. (2019). Noninvasive brain stimulation in generalized anxiety disorder: a systematic review. Progress in NeuroPsychopharmacology and Biological Psychiatry, 93, 3138.

Sakai, K., & Passingham, R. E. (2004). Prefrontal selection and medial temporal lobe reactivation in retrieval of shortterm verbal information. Cerebral Cortex, 14(8), 914921.

Sari, B. A., Koster, E. H., & Derakshan, N. (2017). The effects of active worrying on working memory capacity. Cognition and Emotion, 31(5), 9951003.

Slagter, H. A., Davidson, R. J., & Lutz, A. (2011). Mental training as a tool in the neuroscientific study of brain and cognitive plasticity. Frontiers in Human Neuroscience, 5, 17.

Sporns, O. (2013). Structure and function of complex brain networks. Dialogues in Clinical Neuroscience, 15(3), 247262.

Stojanoski, B., Emrich, S. M., & Cusack, R. (2019). Representation of semantic information in ventral areas during encoding is associated with improved visual shortterm memory. bioRxiv, 201912.

Strauch, C., Wang, C. A., Einhäuser, W., Van der Stigchel, S., & Naber, M. (2022). Pupillometry as an integrated readout of distinct attentional networks. Trends in Neurosciences, 45(8), 635647.

Thams, F., Kuzmina, A., Backhaus, M., Li, S. C., Grittner, U., Antonenko, D., & Flöel, A. (2020). Cognitive training and brain stimulation in prodromal Alzheimer's disease (ADStim)—study protocol for a doubleblind randomized controlled phase IIb (monocenter) trial. Alzheimer's Research & Therapy, 12, 112.

Tomassini, A., Hezemans, F. H., Ye, R., Tsvetanov, K. A., Wolpe, N., & Rowe, J. B. (2022). Prefrontal cortical connectivity mediates locus coeruleus noradrenergic regulation of inhibitory control in older adults. Journal of Neuroscience, 42(16), 34843493.

Unsworth, N., & Robison, M. K. (2017). A locus coeruleusnorepinephrine account of individual differences in working memory capacity and attention control. Psychonomic Bulletin & Review, 24, 12821311.

Usán Supervía, P., & Quílez Robres, A. (2021). Emotional regulation and academic performance in the academic context: The mediating role of selfefficacy in secondary education students. International Journal of Environmental Research and Public Health, 18(11), 5715.

Van der Wel, P., & Van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. Psychonomic Bulletin & Review, 25, 20052015.

Vecchio, F., Quaranta, D., Miraglia, F., Pappalettera, C., Di Iorio, R., L'Abbate, F., ... & Rossini, P. M. (2022). Neuronavigated Magnetic Stimulation combined with cognitive training for Alzheimer's patients: An EEG graph study. Geroscience, 114.

Veldsman, M., Mitchell, D. J., & Cusack, R. (2017). The neural basis of precise visual shortterm memory for complex recognizable objects. NeuroImage, 159, 131145.

VuCheung, K., Ester, E. F., & Sprague, T. C. (2021). Spatial working memory representations in human cortex are robust to a taskirrelevant interrupting stimulus. bioRxiv, 202109.

Williams, J., & Störmer, V. S. (2021). Working memory: how much is it used in natural behavior?. Current Biology, 31(4), R205R206.

Woolgar, A., Jackson, J., & Duncan, J. (2016). Coding of visual, auditory, rule, and response information in the brain: 10 years of multivoxel pattern analysis. Journal of Cognitive Neuroscience, 28(10), 14331454.

Woolgar, A., Thompson, R., Bor, D., & Duncan, J. (2011). Multivoxel coding of stimuli, rules, and responses in human frontoparietal cortex. NeuroImage, 56(2), 744752.

World Health Organization. (2023). Anxiety disorders. https://www.who.int/newsroom/factsheets/detail/anxietydisorders

Xu, Y. (2018). The posterior parietal cortex in adaptive visual processing. Trends in Neurosciences, 41(11), 806822.

Zanto, T. P., & Gazzaley, A. (2009). Neural suppression of irrelevant information underlies optimal working memory performance. Journal of Neuroscience, 29(10), 30593066.

Zanto, T. P., Rubens, M. T., Thangavel, A., & Gazzaley, A. (2011). Causal role of the prefrontal cortex in topdown modulation of visual processing and working memory. Nature Neuroscience, 14(5), 656661.

Zhang, H., Yang, S. Y., Qiao, Y., Ge, Q., Tang, Y. Y., Northoff, G., & Zang, Y. F. (2022). Default mode network mediates low-frequency fluctuations in brain activity and behavior during sustained attention. Human Brain Mapping, 43(18), 54785489.

Zhou, C., Lorist, M. M., & Mathôt, S. (2022). Is categorization in visual working memory a way to reduce mental effort? A pupillometry study. Cognitive Science, 46(9), e13194.

Zickerick, B., Thönes, S., Kobald, S. O., Wascher, E., Schneider, D., & Küper, K. (2020). Differential effects of interruptions and distractions on working memory processes in an ERP study. Frontiers in Human Neuroscience, 14, 84.

Zokaei, N., Board, A. G., Manohar, S. G., & Nobre, A. C. (2019). Modulation of the pupillary response by the content of visual working memory. Proceedings of the National Academy of Sciences, 116(45), 2280222810.