Investigating working memory capacity: stability, contextual cueing and attention

TURNER, COURTNEY,LEIGH

How to cite:

TURNER, COURTNEY,LEIGH (2025). *Investigating working memory capacity: stability, contextual cueing and attention* , Durham e-Theses. http://etheses.dur.ac.uk/16402/

Investigating working memory capacity: stability, contextual cueing and attention

Courtney Turner

A thesis submitted for the degree of Doctor of Philosophy in the Department of Psychology

2025

Abstract

Visual working memory (vWM) is a system responsible for temporarily storing visual information, yet its precise nature and capacity limits remain unclear. This thesis presents three studies examining whether vWM capacity is stable across perceptual conditions, can be increased through learned regularities and how it constrains the allocation of attention during search. Study 1 (Chapter 2) comprised five behavioural experiments manipulating perceptual parameters: stimulus density (Experiment 1), stimulus eccentricity (Experiment 2), stimulus organisation (Experiment 3), stimulus type (Experiment 4), and spatial variability of stimuli (Experiment 5). Absolute capacity estimates (K values) differed significantly across conditions but showed consistent individual differences. Study 2 (Chapter 3) comprised two electrophysiological experiments investigating the effects of implicit learning on vWM. Experiment 1 employed a change detection task with low vs. high memory load conditions. Results showed that repeated configurations led to higher K values and reduced contralateral delay activity (CDA) amplitudes. Experiment 2 compared visuospatial (colour-location) and non-spatial (colour-only) conditions, revealing that both K increases and CDA reductions were specific only to the visuospatial condition. Study 3 (Chapter 4) was a single electrophysiological experiment examining the relationship between vWM capacity and visual search performance. Results from change detection and search tasks revealed that individuals with higher vWM capacity exhibited greater search accuracy, stronger attentional selection, and reduced performance costs under higher search loads. Together, these studies demonstrate that VWM capacity is not fixed in absolute terms but is a reliable relative measure across individuals. This work advances our understanding of how vWM capacity interacts with perceptual context, implicit learning, and attentional control.

Contents

List of Figures .... *v*

Declaration .... *viii*

Statement of Copyright .... *ix*

Acknowledgments .... *x*

Chapter 1: General Introduction *1*

1.1 Introduction to working memory 2

1.2 Working memory models 3

1.3 Working memory capacity 6

1.4 Visual working memory capacity measurements 11

1.5 Visual working memory and contextual learning 14

1.6 Visual working memory and attentional templates 15

1.7 Overall thesis aims and overview 17

Chapter 2: Individual working memory capacity measures depend on visuo-perceptual stimulus parameters *21*

2.1 Introduction 22

2.2 General Methods 26

2.3 Experiment 1: Stimulus density 30

2.3.1 Methods 31

2.3.2 Results 32

2.3.3 Discussion of Experiment 1 35

2.4 Experiment 2: Stimulus eccentricity 36

2.4.1 Methods 37

2.4.2 Results 38

2.4.3 Discussion of Experiment 2 39

2.5 Experiment 3: Stimulus organisation 42

2.5.1 Methods 43

2.5.2 Results 44

2.5.3 Discussion of Experiment 3 45

2.6 Experiment 4: Stimulus type 47

2.6.1 Methods ... 48
2.6.2 Results ... 49
2.6.3 Discussion of Experiment 4 ... 50
2.7 Experiment 5: Spatial variability of stimuli ... 52
2.7.1 Methods ... 53
2.7.2 Results ... 54
2.7.3 Discussion of Experiment 5 ... 55
2.8 General Discussion ... 58
Chapter 3: Contextual cueing: implicit learning improves visual working memory capacity ... 62
3.1 Introduction ... 63
3.2 Experiment 1 ... 69
3.2.1 Methods ... 69
3.2.2 Results ... 75
3.2.3 Discussion of Experiment 1 ... 78
3.3 Experiment 2 ... 81
3.3.1 Introduction ... 81
3.3.2 Methods ... 82
3.3.3 Results ... 84
3.3.4 Discussion of Experiment 2 ... 88
3.4 General discussion ... 89
Chapter 4: Individual working memory capacity predicts search performance in multiple colour search ... 94
4.1 Introduction ... 95
4.2 Methods ... 101
4.3 Results ... 111
4.3.1 Working memory – Change detection task ... 111
4.3.2 Attention – Visual search task ... 118
4.3.3 Cross task correlations (change detection and visual search task) ... 123
4.3.4 Rapid serial probe presentation (RSPP) ... 125
4.4 Discussion ... 133
Chapter 5: General Discussion ... 140
5.1 Summary of key findings across studies ... 141

5.2 Theoretical implications for models of vWM capacity 143
5.3 Limitations and future directions 149
5.4 Conclusion 151
References 153

List of Figures

Chapter 1: General Introduction

Figure 1 Schematic overview of theoretical working memory models 4

Chapter 2: Individual working memory capacity measures depend on visuo-perceptual stimulus parameters

Figure 2 Schematic illustration of the stimuli and time course of events in the change detection tasks of Experiments 1 to 5..29
Figure 3 Schematic illustration of stimuli and results from Experiment 1 34
Figure 4 Schematic illustration of stimuli and results from Experiment 2 40
Figure 5 Schematic illustration of stimuli and results from Experiment 3 45
Figure 6 Schematic illustration of stimuli and results from Experiment 4 50
Figure 7 Schematic illustration of stimuli and results from Experiment 5 55

Chapter 3: Contextual cueing: implicit learning improves visual working memory capacity

Figure 8 Schematic illustration of the stimuli and time course of events in the two-colour and four-colour tasks of Experiment 1 and the colour-location and colour-only tasks of Experiment 2 73
Figure 9 Behavioural performance (error rates & K) across the two-colour and four-colour tasks for repeated and novel configurations of Experiment 1..................76
Figure 10 Grand-averaged ERPs and CDA difference waveforms elicited in repeated and novel stimulus configurations in the two-colour and four-colour tasks of Experiment 1 78
Figure 11 Behavioural performance (error rates & K) across the colour-location and colour-only tasks for repeated and novel configurations of Experiment 2...........85
Figure 12 Grand-averaged ERPs and CDA difference waveforms elicited in repeated and novel stimulus configurations in the colour-location and colour-only tasks of Experiment 2 87

Chapter 4: Individual working memory capacity predicts search performance in multiple colour search

Figure 13 Schematic illustration of the stimuli and time course of events in the change detection task........104

Figure 14 Schematic illustration of the stimuli and time course of events in the visual search task........107

Figure 15 Mean working memory capacity (K) in all load trials of the change detection task........111

Figure 16 Mean working memory capacity (K) in all load trials of the change detection task for High-K and Low-K groups........113

Figure 17 CDA difference waveforms and mean amplitudes in all load trials of the change detection task........114

Figure 18 Scatterplots showing the relationship between working memory capacity (K) and CDA in all load trials........115

Figure 19 CDA difference and load costs waveforms and mean amplitudes in all load trials of the change detection task for High-K and Low-K groups........117

Figure 20 Error rates and reaction times in all load trials of the visual search task.....119

Figure 21 N2pc difference waveforms and mean amplitudes in all load trials of the visual search task........120

Figure 22 Scatterplots showing the relationship between working memory capacity (K) and N2pc in all load trials........121

Figure 23 N2pc difference and load costs waveforms and mean amplitudes in all load trials of the visual search task for High-K and Low-K groups........123

Figure 24 Scatterplots showing the relationships between changes in N2pc, CDA and working memory capacity (K) across all loads. 125

Figure 25 N2pc difference waveforms for target colour probes in the one-colour search for High-K and Low-K groups........129

Figure 26 N2pc difference waveforms for target colour probes in the two-colour search for High-K and Low-K groups........130

Figure 27 N2pc difference waveforms for target colour probes in the three-colour search for High-K and Low-K groups 131
Figure 28 N2pc difference waveforms for distractor colour probes in all load trials for High-K and Low-K groups 133

Declaration

The author confirms that none of the material presented in this thesis has been submitted elsewhere for any other qualification and is the author’s own work unless referenced otherwise.

Statement of Copyright

Acknowledgments

First and foremost, I would like to thank Professor Anna Grubert. From very first meeting you, I knew you would play a big part in my life. You've been so much more than a supervisor - part mentor, part coach and occasionally part stand-up comedian (whether you realised it or not). You've had an uncanny ability to know exactly when I needed a motivational pep talk and your patience has been saint like. You've not only guided me through the challenges of this PhD, but you've also taught me how to think critically, stay curious and persevere when things don't go to plan. I will always be grateful for your guidance and kindness, and I know the lessons I've learned from you will stay with me far beyond this PhD. I would also like to thank Professor David Sanderson and Professor Anthony McGregor for your help and support throughout my PhD journey.

A big thank you to Ziyi, Rai and Mikel (my lab buddies) for the experiments, brainstorming and most importantly, the laughter. I'll never forget how easily we could go from serious research mode to uncontrollable laughter in about three seconds and I'm so grateful to have shared this chapter of my life with you.

To all my PGR and PDRA friends in Durham, the real OG's and to Linda. From first year nerves to final year chaos, you've been there with tea, advice and laughter, turning ordinary days into some of my favourite memories.

To my amazing best friends outside of academia, Emily, Gemma and Becca. Thank you for putting up with me disappearing during busy times, for listening to me explain ‘life’ and for finding a way to make me laugh when I needed it most.

To my parents, thank you for a lifetime of unconditional love and encouragement, even if your most common question over the past few years has been “So… what is it you’re doing again?”. You may not have understood the ins and outs of my research (and honestly, I don’t blame you), but you’ve understood me and that has made all the difference.

Finally, to Chris, thank you for being my constant support through every stage of this journey. You’ve done it with patience, humour and a genuine interest (or at least a very convincing impression of it anyway). You’ve kept me grounded when I’ve started to spiral, reminded me to take breaks when I was running on empty and celebrated every tiny victory as if it was the finish line. Your love and belief in me have helped me finish this thesis and I’m so lucky to have had you by my side through it all.

Chapter 1: General Introduction

1.1 Introduction to working memory

The ability to temporarily maintain information actively in mind and manipulate that information is a central function of human cognition (Baddeley, 1986, 1996). This cognitive system, known as working memory (WM), allows us to maintain relevant information in an active state to support complex behaviours such as planning, problem solving, learning and goal-directed behaviour (Baddeley 1996). Unlike long-term memory (LTM), which holds vast amounts of information over long periods of time, WM is limited in both capacity and duration (Cowan, 2001, 2008). Despite this, WM plays a fundamental role in everyday life. Whether locating a friend in a crowd or mentally rehearsing a list of shopping items, we depend on the ability to keep information temporarily accessible and actively maintained. Over the past few decades, the study of WM has occupied a prominent position in cognitive psychology (Baddeley, 2012). Research through models and concepts has demonstrated that WM is not merely a passive storage system, but rather an active and flexible component that supports a wide range of high-level processes (Baddeley, 2003; Cowan, 2017). While I return to these models in greater detail later in the thesis, here I focus on the functional role of WM and more specifically visual working memory (vWM) as the basis for my research.

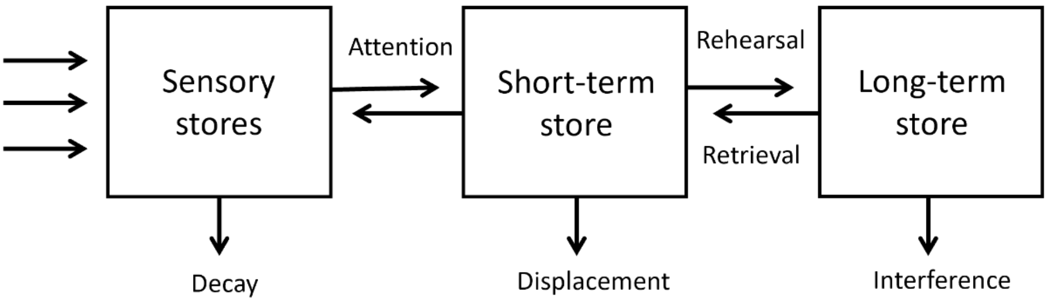
Within WM, researchers have identified domain-specific subsystems: verbal and visual working memory, which retain different types of information (Engle et al., 1999; Repovs & Baddeley, 2006). My research focuses on vWM, which is responsible for temporarily storing visual features such as colour, shape and spatial location. It plays a crucial role in everyday tasks such as remembering locations and it also operates closely with attention to prioritise goal-relevant features (Awh et al., 2006). Despite its flexibility, vWM has a limited capacity, with most individuals typically able to hold around three to four visual items at once (Cowan, 2001; Luck & Vogel, 1997). This has prompted an extensive investigation in the literature into how the capacity of vWM is defined, measured and essentially employed.

The present thesis focuses on K (Cowan, 2001), a behavioural estimate of vWM capacity that quantifies how many items an individual can store, as a central construct. Across three empirical studies, I examine how K varies across individuals and perceptual conditions, how it is influenced by learning and spatial regularities and how it shapes the guidance of attention during visual search. I aim to clarify the nature of vWM capacity as both a trait-like individual difference and a state-like, context sensitive system that is adaptable depending on environmental needs.

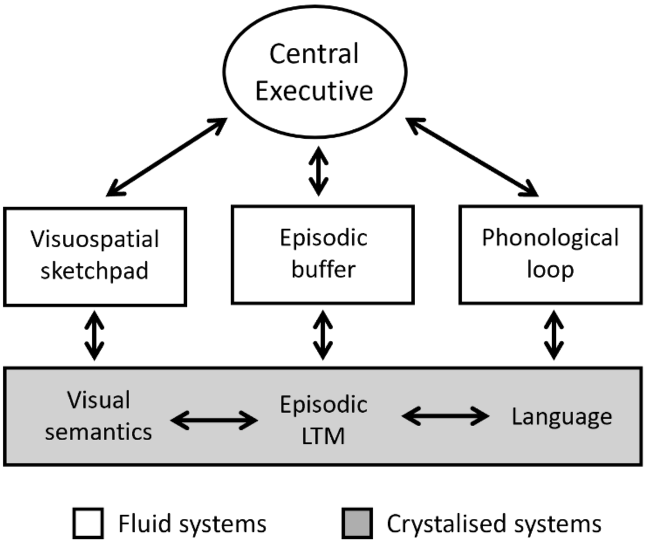
1.2 Working memory models

Several competing WM models have been proposed over decades to explain its structure, function and capacity. Each of these models all differ in how they conceptualise the relationship between short-term memory and LTM, but each has significantly contributed to our understanding of how visual information is maintained and used in goal-related behaviours. In this section I provide three complementary models on WM function: the modal model (Atkinson & Shiffrin, 1968; Broadbent, 1958), the multi-component model (Baddeley & Hitch, 1974; Baddeley, 2000) and the embedded-processes model (Cowan, 1988, 1999).

A) The Modal model (Broadbent, 1958; Atkinson & Shiffrin, 1968)



B) The Multi-component model (Baddeley & Hitch, 1974; 2000)



C) The Embedded-processes model (Cowan, 1988; 1999)

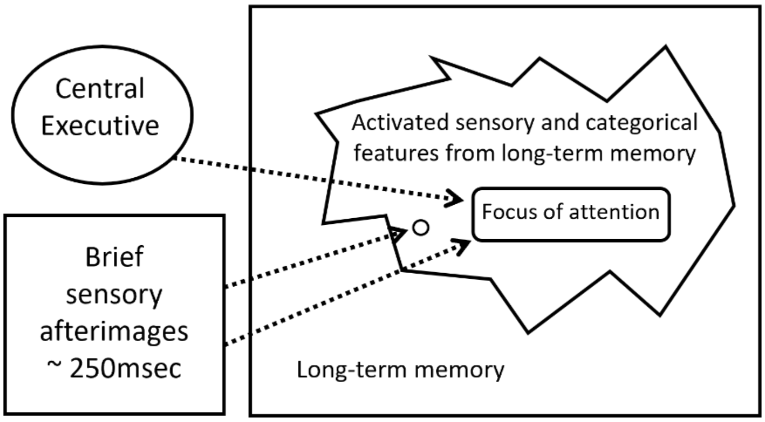


Figure 1 Schematic overview of theoretical working memory models

The first system postulated regarding general memory was the modal model (Atkinson & Shiffrin, 1968; Broadbent, 1958). This first attempt model assumed that there was a short-term store system that mediated incoming sensory information from the external environment with information in long-term storage. Although this model introduced the idea of a temporary memory buffer, it treated short-term storage as a relatively passive system. This model, as seen in Figure 1, panel A, was proposed by Broadbent (1958) and later expanded by Atkinson and Shiffrin (1968). The modal model provided a basic understanding of how WM adapts to the environment laying the foundations for subsequent models that redefined this short-term store as more active than passive, now recognised as WM.

In contrast to the modal model, and perhaps the most influential, the multi-component model proposed by Baddeley and Hitch (1974), suggests WM is an active system. Shown in Figure 1, panel B, it includes a central executive that coordinates the phonological loop, which deals with speech-based information and the visuospatial sketchpad, responsible for visual and spatial information. Later revisions of their model also incorporated an episodic buffer (Baddeley, 2000), which integrates multimodal information and provides an interface with LTM. Of particular interest to the current research is the visuospatial sketchpad which aligns with vWM. Specifically, the model allows for investigation of capacity limitations within vWM by separating components (i.e. phonological loop & visuospatial sketchpad) based on the type of information they store.

A different perspective is that of the embedded-processes model, proposed by Cowan (1988, 1999). Rather than there being separate memory stores like in the previous models, they see WM as a state of activated LTM. This model, depicted in panel C of Figure

1, conceptualises the limited amount of information that is readily available as being the focus of attention. Simply, LTM representations become temporarily activated and a smaller amount enters the focus of attention, posited as a capacity limited component. This model places attention at the centre of WM, emphasising its role in deciding which representations are actively maintained. Although this model suggests a change in the central executive's role in the memory process, it complements the multi-component model by offering a more attention-driven explanation for why capacity varies and directly underpins K as a measure used throughout this thesis.

Together, these models present important insight into the dynamic structure of WM. While the modal model introduced the concept of temporary storage, both the multi-component and embedded-processes model has helped conceptualise WM as a limited but adaptable system, highlighting the importance of attention. In this thesis, I provide a view that draws on both the domain and attention elements of the multi-component and embedded-processes model respectively. The following sections explore how vWM capacity is conceptualised and measured, with K serving as the main measure across all three studies.

1.3 Working memory capacity

The most well-known contribution into how much information WM can hold at once came from Miller (1956). Using tasks involving digit span and verbal sequences, they established that verbal working memory consisted of 'chunks' limited to 7 ± 2 items. This view sparked a surge of research focusing on quantifying the constraints of WM storage (Cowan, 2001; Luck & Vogel, 1997). One of the main reasons for this focus is because individual differences in WM are often correlated with differences in reading comprehension, fluid intelligence and academic achievement (Alloway & Alloway, 2010;

Fukuda et al., 2010b). Hence, there has been strong motivation to better understand these limits as it could provide insight into cognitive function more generally. Subsequent research, especially in the visual domain, has revised this estimate considerably, suggesting that capacity is typically limited to around three to four items for vWM (Cowan, 2010). This view remains the most prominent in the literature to date. In this section, I outline the main theoretical approaches to vWM capacity, focusing on the discrete-slot model, the fixed resource model, hybrid accounts and relevant neural correlates. These views will help to interpret the methods and findings of my three studies to follow.

The discrete-slot model suggests that vWM capacity is limited by a discrete number of ‘slots’, meaning any stored item must be assigned to one of a limited number of slots (Cowan, 2001; Fukuda et al., 2010a). Once this limit is exceeded, additional items are not stored. Evidence supporting this model has come from behavioural studies using change detection tasks showing that when set size increases beyond capacity, performance declines (Awh et al., 2007; Balaban et al., 2019; Luck & Vogel, 1997; Rouder et al., 2008; Vogel et al., 2001). For example, Luck and Vogel (1997) found that participants could reliably detect changes with three to four items, but performance declined beyond that. Crucially, they also explored whether this capacity depends on the number of features or the number of objects. Accuracy was measured across set sizes, 4, 8 or 12 under three conditions: colour-only changes, orientation-only changes and a conjunction condition, where either colour or orientation could change. Using a set size of four, results showed that performance in the single feature conditions (colour-only & orientation-only) was almost identical to performance in the conjunction condition, concluding that capacity is determined by the number of objects and not the number of features, consistent with the discrete-slot model of vWM capacity. Despite the discrete-slot model being the most prevailing model used in terms of vWM capacity (Cowan, 2001; Luck & Vogel, 2013) and its support from empirical work, it has been argued that it's all or nothing nature may oversimplify the mechanisms underlying encoding, maintenance and retrieval in vWM.

In contrast, the alternative flexible-resource model suggests vWM capacity is limited by a pool of resources that can be distributed across an unlimited number of items, prioritising some over others (Bays & Husain, 2008; Wilken & Ma, 2004). Thus, all items can be stored, but with declining precision as set size increases. Evidence to support this model shows that response errors increase gradually with load (Wilken & Ma, 2004), as opposed to stopping at a certain set size, like in the discrete-slot model. Not only has the flexible-resource model gained evidence from studies assessing how many items are stored, but also how similar, complex or salient items are (Alvarez & Cavanagh, 2004). For example, using five stimuli classes and set sizes of 4, 8 and 12 in the change detection task, Alvarez and Cavanagh (2004) found that both information load and number of objects affect capacity limits on vWM.

While discrete-slot and flexible-resource models offer contrasting accounts of how vWM capacity is constrained, an emerging body of research has argued for hybrid models that incorporate aspects of both. One prominent hybrid account which blurred the distinction between discrete slots and flexible resources was developed by Zhang and Luck (2008). They suggest that when fewer items are presented than the number of available slots (e.g., 2 items, 4 slots), multiple slots can be assigned to a single item simultaneously to enhance its representational precision. In other words, when the memory load is low, the 'extra' unused slots are not wasted. Similarly, Chong and Treisman (2005a, 2005b) argue that capacity limitations may not just apply to individual items, but to abstract representations of the entire visual field. From this perspective, the nature of what is stored (e.g. contexts, features, objects), and how it is represented can vary depending on task demands and attentional priorities. Overall, it seems there is a bridge between discrete-slot and flexible-resource models showing characteristics of both types of models.

Further hybrid accounts of capacity arise from interference-based models, which challenge the need for fixed capacity systems altogether. For example, Oberauer and Kliegl (2001, 2006) proposed that capacity limitations arise from the mutual interference between stored items, rather than from a fixed number of available slots or resources. They administered an updating paradigm whereby participants maintained digits in spatially specific locations while executing arithmetic updates. Performance decreased with increasing load and ceiling accuracy was shown up to around four items, like results from change detection tasks (Luck & Vogel, 1997). However, the critical observation here was that they attributed this limit to item interference rather than a strict capacity ceiling, where interference results from overlapping feature representations, making it harder to separate items as more are stored.

Further support for interference models comes from Davelaar et al. (2005) who developed the general-plus-specific interference model. They suggested that all newly encoded items introduce competition within WM regardless of similarity. When related stimuli are present, this interference is more pronounced, but even unrelated stimuli can present interference. This account offers a mechanism by which both capacity limits and representational similarity limit performance, consistent with both discrete slot and flexible resource restrictions.

Lastly, the dual-store model (Unsworth & Engle, 2007) offers a complementary account. This model appoints individual differences in WM performance to both primary memory (capacity-limited and attentionally sustained) and secondary memory (longer-term storage that requires controlled retrieval). This view suggests capacity limits are observable when task demands separate primary memory, i.e., in change detection paradigms, providing a framework in which capacity limitations (such as those estimated by K) are separated from retrieval variance.

In addition to behavioural models of vWM capacity, neural markers have also been identified to reflect the number of items actively maintained in memory. One of the most widely studied electrophysiological correlates is the contralateral delay activity (CDA), a sustained negative wave observed over posterior PO7/PO8 scalp regions, contralateral to the hemifield where the memory items are presented (Ikkai et al., 2010; McCollough et al., 2007; Vogel & Machizawa, 2004). In line with slot-based predictions, the CDA increases with set size, often levelling off at around three to four items (Vogel et al., 2005). However, interpretations of the CDA are not restricted to discrete-slot models and thus not entirely capacity driven. Some researchers have argued that the CDA may reflect attentional allocation, resource distribution or encoding strength, making it relevant to flexible-resource and hybrid models too (Luria et al., 2016). For example, if items vary in complexity, the CDA can represent the overall demand of these items, not just how many there is (Roy & Faubert, 2023). This interpretation stresses the importance of using both behavioural and neural measures when assessing capacity. Here, I use the CDA to complement behavioural measures of capacity (K) to explore how vWM capacity is modulated by stimulus context (Study 2) and linked to attentional control (Study 3).

The range of capacity models presents three key features: vWM capacity is limited, variable across individuals and sensitive to task demands. The diversity of the models reflects the complexity of vWM, thus my aim in this thesis is not to test these models directly against each other, but rather to explore how vWM capacity is expressed under different cognitive and environmental conditions. I examine how K varies with perceptual structure (Study 1), how it is enhanced through implicit learning (Study 2) and how it correlates with attentional control efficiency (Study 3).

1.4 Visual working memory capacity measurements

The models outlined previously form the foundation for how capacity is conceptualised. In this section, I provide a description of the tasks and tools used to measure capacity in experimental research, focusing on the task I use in all three of my studies: the change detection task. I also introduce the rationale for Study 1 which directly addresses the issue of vWM capacity measurements, in that capacity estimates are sensitive to the design and contexts of task displays.

WM capacity can be tested in a variety of tasks. A commonly used measure are complex span tasks which combine a memory span with a concurrent processing task. Daneman and Carpenter (1980) invented the first version of this task, named the ‘reading span’, where individuals read several sentences and tried to remember the last word of each sentence. While they believed that the combination of storage and processing was needed to measure WM capacity, we now know that WM capacity can be measured with WM tasks that have no additional processing involved (Colom et al., 2008; Unsworth & Engle, 2007). However, the question of what features a task must have to qualify as a measure of WM capacity remains an open question.

Used in the present three studies and one of the most widely used paradigms for estimating vWM capacity is the change detection task (Phillips, 1974), which has been well established by researchers (Luck & Vogel, 1997; Vogel et al., 2001). In a typical change detection task, participants briefly view a memory display of visual items e.g. coloured squares and remember these across a short delay. Participants are then presented with a test display where either one of the items has changed (change trial) or stayed the same (no-change trial). By manipulating the set size in the memory arrays, researchers can infer individual memory capacity from accuracy rates, also known as K (Cowan, 2001). There are two formulae used to measure K in the literature which are both dependant on whether the task is a whole (Pashler, 1988) or single probe display (Cowan, 2001). Since all three studies presented here all have whole displays and following the suggestions of Rouder et al. (2011), K was calculated according to Pashler's equation (Pashler, 1988) in all three of my studies. The formula to calculate K is:

$$K = N\left(\frac{h-f}{1-f}\right)$$

where N is the number of memory items (set size), and *h* and *f* are the measured hit and false alarm rates, respectively. Hit rates are calculated as

$$h = \frac{hits}{hits + misses}$$

and false alarm rates are calculated as

$$f = \frac{false\ alarms}{false\ alarms + correct\ rejections}$$

Although alternative approaches to measure vWM capacity have been developed (Bays & Husain, 2008), given the structure of my studies, particularly the use of electrophysiological measures in Study 2 and 3, here I adopt a change detection task paired with K.

Although K is widely used to estimate vWM capacity, the reliability of this measure is influenced by parameters of the experimental design (Vogel et al., 2001). Thus, despite their efficiency, K values are not always constant across tasks or contexts. Manipulations such as item complexity, spatial organisation, or attentional demands can significantly affect K values (Alvarez & Cavanagh, 2004; Oberauer & Eichenberger, 2013). For example, items grouped visually may be remembered with higher accuracy, even when the number of items remains the same (Woodman et al., 2003). This raises a fundamental question: does K reflect a stable trait-like memory limit, or does it fluctuate based on how the task supports encoding and prioritisation? Thus, capacity might not only represent how much an individual can remember, but also how efficiently they can extract, coordinate, and maintain visual input under different environmental settings.

To address this, Study 1 was designed to investigate whether K is stable across perceptual conditions. While K is treated as a robust measure in many studies, I specifically wanted to challenge this assumption by seeing if it is consistent across systematic variations in the visual structure of a task. I had two goals for Study 1: firstly, to determine whether absolute K values change across these contexts and secondly, to assess whether individual stability is preserved, i.e., do individuals maintain their relative vWM capacity even when perceptual task parameters change.

1.5 Visual working memory and contextual learning

Following the discussion of how K can vary with changes in perceptual parameters, a related question now arises: can capacity also be shaped by learning? Can learned regularities in the environment enhance the efficiency of visual information encoding and maintenance? While Study 1 focuses on how perceptual task structure can affect K, Study 2 and the present section focus on the role of contextual learning: how repeated exposure to visual displays may lead to improved memory performance. I begin by introducing the contextual cueing phenomena. I then discuss how implicit learning may influence vWM, leading to the rationale of Study 2.

In our everyday life, we use contextual regularities to guide our attention, for example, knowing where to look for milk in the kitchen relies on learned spatial associations. One concept that has been utilised to explain this phenomenon is the contextual cueing effect (Chun & Jiang, 1998), where repeated spatial configurations improve visual search efficiency over time (Chua & Chun, 2003; Jiang & Chun, 2001; Olson & Chun, 2002). Chun and Jiang (1998) first demonstrated this using a visual search task. This requires participants to search for a target letter ‘T’ amongst a configuration of distractors ‘L’s. Unknown to the participants, half of the displays are repeated distractor configurations that consistently predict a target location while the other half of displays are novel configurations. Results show significantly faster response times for repeated configurations than randomly generated novel configurations, despite individuals being unaware of the repetition. This effect has since been replicated across many different visual search tasks (Chun, 2000; Geyer et al., 2010; Pollmann, 2019; Yang et al., 2021) with a consistent finding: repeated global contexts support faster and more accurate search.

Chun and Jiang (1998, 1999) proposed that repeated exposure to a specific search array leads to the formation of an implicit memory which reflects learned associations between the targets location and the surrounding visual context. Once these associations are formed, they are automatically retrieved and used to guide attention when the same configuration appears. Crucially, contextual cueing occurs without explicit awareness of the repetition, reinforcing its implicit nature. Evidence from neuroimaging studies support this view by showing that contextual cueing engages medial temporal lobe structures, such as the hippocampus, which are typically involved in associative learning (Greene et al., 2007).

While contextual cueing is typically displayed in visual search, very few studies have explored whether this implicit learning can enhance vWM directly. The idea is that repeated configurations in the environment may support more efficient organisation of visual information, leading to an increase in K. This idea challenges the view that contextual cueing operates solely through attentional mechanisms and suggests the possibility that this could occur at the level of memory encoding and maintenance. Thus, Study 2 was designed to investigate whether contextual learning can directly influence vWM capacity, as indexed by K. While contextual cueing is often attributed to attention, I aimed to test whether repeated spatial configurations could also enhance vWM directly. I had two main goals for Study 2: first, to determine whether repeated spatial configurations improve behavioural estimates of capacity (K) and second, to assess whether these improvements are reflected at the neural level, through modulations of the CDA.

1.6 Visual working memory and attentional templates

While I have previously discussed how perceptual and contextual factors may alter the efficiency of vWM capacity, it is important to recognise how capacity limitations may shape the way vWM is used in real time operations. Representations stored in vWM can act as attentional templates, guiding the selection of relevant information in the visual environment, thus understanding this relationship is crucial for the confines of capacity. In this section, I focus on the role of vWM in attentional guidance, specifically, how stored representations function as attentional templates, setting out the rationale for Study 3.

vWM not only enables the short-term storage of visual information, but it also plays a role in attentional control. One way in which this occurs is through attentional templates, which are internal representations of target features that are maintained in vWM (Desimone & Duncan, 1995; Duncan & Humphreys, 1989; Olivers et al., 2011). These templates act as filters that enhance sensitivity to stimuli that match goal-relevant features. For example, when looking for your blue suitcase at the airport baggage carousel you hold a mental image of its colour and shape in mind. This template allows your visual system to prioritise blue square objects and effectively ignore irrelevant items i.e., yellow backpacks. Research at the behavioural level shows that search is more efficient when a single template is maintained, shown by faster reaction times and greater accuracy rates (Carlisle et al., 2011; Olivers et al., 2011). However, when multiple potential target templates must be maintained at the same time, i.e., search for one of three colours, performance typically declines (Beck et al., 2012). These costs may be reflected by the limited capacity of vWM, which varies across individuals.

Neural mechanisms underlying attentional template use have been studied extensively using EEG, particularly the N2-posterior-contralateral (N2pc) component. The N2pc component of the event-related potential (ERP) is a negative going deflection observed approximately 200-300ms after stimulus onset at posterior scalp electrodes contralateral to the attended visual field. It serves as an electrophysiological marker of attentional selection (Eimer, 1996; Luck & Hillyard, 1994). The N2pc has been widely used to track the time course of attentional deployment during visual search tasks. Typically, when a search display contains a lateralised target that matches a maintained template, an N2pc is observed during search. To more precisely track attentional processes prior to and during search, Grubert and Eimer (2018) developed the rapid serial probe presentation (RSPP) paradigm in combination with a visual search task. In this design, not only are N2pcs measured in response to the target in the search display, but also time-locked to lateralised probes presented during the preparatory phase, before target onset. The idea is that when a probe matches an item stored in vWM, and that template is active, it captures attention thus eliciting an N2pc, allowing monitoring of the activation state of vWM representations over time. Critically, this then raises the question as to whether both search- and probe-evoked N2pcs are sensitive to individual differences in capacity. That is, does an individual's capacity, as measured behaviourally by K and neurally via CDA, predict how many templates they can maintain and activate in parallel?

To address this, Study 3 was designed to investigate whether individual differences in K predict the efficiency of attentional selection under conditions of increasing load. While it is widely assumed that vWM supports the maintenance of attentional templates (Olivers et al., 2011), I specifically aimed to test whether these templates are constrained by an individual's vWM capacity (K).

1.7 Overall thesis aims and overview

This thesis aims to advance our understanding of vWM capacity through the lens of K, a widely used behavioural estimate of the number of items an individual can store in WM.

While traditional accounts have conceptualised K as a fixed, trait-like marker of capacity, the work presented in this thesis suggests K may be sensitive to perceptual, contextual and strategic elements. In this thesis, I measure K alongside electrophysiological markers to explore the dynamic interplay between trait and state aspects of vWM capacity.

The three studies presented in this thesis are organised into three empirical chapters designed to examine different, but interconnected, aspects of vWM capacity. Chapter 2 presents Study 1 which investigates the stability of K across perceptual task conditions. Chapter 3 presents Study 2 which examines whether K can be improved through implicit contextual learning. Chapter 4 presents Study 3 which explores whether individual differences in K predict how many attentional templates can be used simultaneously during visual search.

In Chapter 2 (Study 1), I challenge the assumption that K represents a stable, trait-like capacity limit. While K is often used to capture individual vWM ability, it's not clear whether these estimates are consistent across variations in perceptual structure. To investigate this, I conducted five behavioural experiments using a change detection task. A different perceptual stimulus property was manipulated in each of the five experiments: stimulus density (Experiment 1), stimulus eccentricity (Experiment 2), stimulus organisation (Experiment 3), stimulus type (Experiment 4), and spatial variability of stimuli (Experiment 5). By systematically varying perceptual characteristics of the memory displays, I aimed to determine whether K remains consistent across conditions or whether it is restricted to perceptual features. Study 1 was motivated by the idea that capacity may not be fixed and challenges the interpretation of K as a purely intrinsic measure.

In Chapter 3 (Study 2), I investigate whether K can be enhanced through contextual learning. By incorporating repeated visual configurations into a change detection task, I tested whether contextual cueing effects, typically found in visual search, can be observed directly at the level of vWM. Experiment 1 examined whether repeated spatial configurations improve vWM capacity under different memory loads (two and four) using a change detection task. In Experiment 2, I tested whether contextual learning enhances capacity specifically when spatial information is relevant. Using only a set size of three, I used a similar change detection task with two conditions: visuospatial (context was defined by colour and location) and non-spatial (context was defined by colour only). This allowed me to detach the spatial element of contextual cueing.

Finally, in Chapter 4 (Study 3), I examine individual differences in K and whether they predict the number of templates that can be maintained simultaneously in visual search. Here I test whether capacity limits constrain the number of templates that can be maintained during visual search. To address this, I incorporate two tasks (change detection task and visual search task), each with three levels of load (1, 2 or 3), to systematically manipulate the number of items to be remembered or the number of templates to be activated in parallel. The change detection task was used to measure individual vWM capacity behaviourally (K) and neurally (CDA). The visual search task combined with the RSPP paradigm was used to measure accuracy rates, and target- and probe-evoked N2pcs. The overall aim here was to determine whether individuals with higher capacity are better able to maintain and employ multiple attentional templates.

Taken together, the work in this thesis seeks to support a functional and flexible view of vWM capacity. Rather than viewing K as a fixed trait, I propose that capacity emerges through the dynamic interplay of perceptual limitations, contextual learning, and individual attentional mechanisms. By integrating behavioural and neural measures across varying paradigms, this research builds a comprehensive picture of how vWM capacity functions, adapts, and varies, advancing our understanding of not just what K is, but how it works.

Chapter 2: Individual working memory capacity measures depend on visuo-perceptual stimulus parameters

2.1 Introduction

vWM refers to the cognitive system responsible for the temporary storage and manipulation of visual information that is no longer physically available in the visual environment (Baddeley, 2000; Baddeley & Hitch, 1974). It plays a crucial role in guiding attention, perception, and goal-directed behaviour by holding a limited amount of visual information available for short periods. A common method for measuring the capacity of vWM is the change detection task (Luck & Vogel, 1997; Phillips, 1974), in which participants are briefly shown a display of visual items, followed by a short blank delay, and then a second stimulus display. Participants' task is to indicate whether a change has occurred between the two stimulus displays, or whether they are identical. Accuracy rates are then used to calculate a working memory (WM) capacity coefficient K (Cowan, 2001; Pashler, 1988; Rouder et al., 2011), which estimates the number of items an individual can maintain in vWM (Luck & Vogel, 1997).

A core concept of vWM is that its capacity is limited and that these capacity limitations are individually different. On average, participants can maintain three to four objects in vWM (Cowan, 2001), but individual differences have been observed (Vogel & Awh, 2008) and linked to individual visual search (Anderson et al., 2011; Luria & Vogel, 2011) and multiple object tracking performance (Drew et al., 2011). A study also revealed that individuals with higher vWM capacity are more efficient at excluding unnecessary items during task performance (Vogel et al., 2005). Understanding these capacity limits is particularly important because they are thought to be stable within individuals and are predictive of a broad range of other cognitive functions. For example, prior research has demonstrated that vWM capacity is correlated with syllogistic and spatial reasoning (Capon et al., 2003), mathematical skills (Kyttälä & Lehto, 2008), creativity (Maor et al., 2023), academic achievement (Aspanani et al., 2023; Tsubomi & Watanabe, 2017), and fluid intelligence (Engel de Abreu et al., 2010; Fukuda et al., 2010b). Since researchers use an estimation of individual vWM capacity to account for individual differences in other cognitive functions, it is essential to evaluate the reliability and validity of vWM capacity estimation based on the change detection task (Dai et al., 2019). If vWM capacity is indeed a stable trait, an individual's capacity measure should remain consistent across different implementations of the change detection task, regardless of specific task parameters. In other words, the overall ability of a person to store more or fewer items in vWM should be a generalisable and task-invariant property, reinforcing the utility of vWM capacity as a robust measure in cognitive and psychological research.

Several empirical studies have provided positive evidence supporting this hypothesis. For instance, Fukuda et al. (2015) found that individuals K estimates did not systematically change over the time course of an experiment, suggesting that K is a stable and reliable metric of individual differences in information processing within a single test session. In a dedicated learning study, Xu et al. (2018) measured individual vWM capacity across multiple test sessions and observed that individual K values increased over the sessions, reflecting improvement in overall change detection performance. However, even though they found that K was not stable across test session, the reliability of K measures was not compromised, not even with extensive practice, i.e., participants with higher K values always had higher K values, regardless of the test session. Along similar lines, Dai et al. (2019) measured vWM capacity with a change detection task in two groups with a three-day or 16-day delay, respectively, and found reasonably high test-retest correlations of K values in both groups.

These converging findings suggest that K is a reliable measure of vWM capacity within a given task and across time. The empirical literature thus provides strong foundational support for the assumption that vWM capacity, as captured by K, reflects a stable and meaningful cognitive trait which can serve as a predictive index for higher-order cognition, such as reasoning, problem-solving, academic performance, and fluid intelligence.

However, a factor that has not been systematically accounted for in the measurement of vWM capacity is the influence of the visual properties of the stimuli to be remembered. Although the change detection task is widely used to estimate K, the assumption that WM capacity measures are purely cognitive and independent of perceptual input is problematic. In reality, the formation of a memory trace is preceded by and dependent on visual perception, and if perceptual quality varies, so too might the fidelity of the resulting memory representation. Consequently, variations in perceptual conditions within the same task may influence K values, thereby challenging the assumption that they reflect a stable, task-invariant trait. Indeed, prior research has demonstrated that visual perception is not uniform across the visual field. For example, visual processing is enhanced at the fovea compared to peripheral vision (Larson & Loschky, 2009), in sparse versus crowded visual scenes (Whitney & Levi, 2011), and for bilaterally presented objects compared to those confined to a single visual hemifield (Cavanagh & Alvarez, 2005). These findings suggest that the quality of the memory encoding, which depends on perceptual acuity and unambiguity, might vary significantly depending on stimulus location, density, and spatial configuration. Since every vWM representation is formed based on an initial visual percept, it follows that a clearer perceptual input may support more accurate memory encoding, just as for example visual acuity and appropriate letter spacing have been shown to enhance reading performance (Chung, 2004).

One conceptual framework that directly links perceptual quality to vWM performance is the sensory recruitment hypothesis, which proposes that vWM storage relies on the same sensory neural systems that initially process the stimulus (Emrich et al., 2013; Pasternak & Greenlee, 2005; Serences, 2016; Sreenivasan et al., 2014). Evidence from neuroimaging showed that memorised visual items can be decoded from activation patterns in primary visual cortex (V1), even in the absence of any sustained activity during the retention period (Harrison & Tong, 2009; Serences et al., 2009). This suggests that the fidelity of WM representations is constrained by the resolution and precision of the underlying sensory codes, making perceptual conditions a direct determinant of memory capacity. Given this interplay between perception and memory, the validity of K as a stable measure of vWM capacity may depend, at least in part, on the perceptual characteristics of the stimuli used in the task. Yet, surprisingly few studies have systematically investigated whether and how perceptual variations affect the stability of K estimates across conditions. In the following five experiments, we set out to close this gap in the literature: are individual differences in vWM capacity robust to changes in visual perception, or are K values modulated by the perceptual quality of the to-be-memorised information?

To address this research question, the present study systematically manipulated five perceptual factors across five experiments: stimulus density, stimulus eccentricity, stimulus organisation, stimulus type, and spatial variability of stimuli. In each experiment, we measured K values to assess whether estimated vWM capacity remained stable across differing perceptual conditions. Additionally, we examined correlations of K values across task conditions to evaluate whether relative memory ability within the sample (i.e., the rank ordering of individuals by K) was preserved, even when perceptual input changed. This approach allows us to test the assumption that K reflects a perceptual-invariant trait, and to clarify the degree to which vWM capacity estimates depend on early visual processing.

2.2 General Methods

2.2.1 Participants

In all five experiments of this study the same fifteen participants were tested, but each experiment was tested in a separate test session with a minimum of three days and a maximum of seven days between test sessions. Participants were paid at an hourly rate of £10. The experiments were approved by the Ethics Committee of the Psychology Department at Durham University and were conducted in accordance with the Declaration of Helsinki. Participants gave informed written consent prior to testing. Participants were between 18 and 22 years of age (M_{age} = 20.5, SD_{age} = 1.1). Nine participants were female and six were male. One participant was left-handed; the others were right-handed. All participants had normal or corrected-to-normal vision and normal colour vision (as verified with the Ishihara colour vision test; Ishihara, 1972). G*Power 3.1 (Faul et al., 2007) indicated that given our sample size, an alpha of .05 and power of .80, we could detect differences in F-tests with partial η^2 as small as 0.1 (effect size f=.33) with a minimal $F(2,28)=3.3$.

2.2.2 Stimuli and procedures

Participants were sat in a dimly illuminated chamber with an approximately 90cm viewing distance from the monitor. Stimuli were presented on a 22-inch MSI Optix G272 LCD monitor with a 100-Hz refresh rate and a resolution of 1920×1080 pixels. MATLAB and the Cogent 2000 toolbox were used on an LG Pentium PC running under Windows 10 to control stimulus presentation, timing, and response collection. Stimuli were presented on a black background. A central grey fixation point (CIE x, y colour coordinates: .305/.314; 0.2°x 0.2° of visual angle) was presented throughout each experimental block. All five experiments employed the same change detection task (see Figure 2). Each trial started with the presentation of a memory display for 200ms, which was followed by an 800ms blank display (resulting in a 1000ms retention time), and a test display which was presented until a manual response was recorded. During the inter-trial interval (between the offset of the test display in trial n-1 and the onset of the memory display in trial n), a blank display was presented for 300ms plus a random temporal jitter between 0-200ms. The memory and test displays in each trial contained a certain number of coloured squares which were presented at the same locations in both displays (see Experiments 1-5 and Figures 3-7 for stimulus details). In 50% of all trials, the stimulus colours were identical in the memory and test displays (no-change trial). In the other 50% of the trials, one of the squares in the test display randomly changed colour (change trial). Change and no-change trials were equiprobable but were presented randomly in an intermixed fashion within each block. Participants were instructed to press the ‘C’ key on a standard keyboard when they detected a colour change, and the ‘N’ key when they felt that the colours in the test display were the same as in the memory display. The order of the five experiments was randomised for each participant, ensuring that no fixed sequence was imposed. A different perceptual stimulus property was manipulated in each of the five experiments of this study: stimulus density (Experiment 1), stimulus eccentricity (Experiment 2), stimulus organisation (Experiment 3), stimulus type (Experiment 4), and spatial variability of stimuli (Experiment 5). Each of these perceptual variables had three equiprobable levels (e.g., 3, 5, and 7 memory items in Experiment 1; see methods of the respective experiments for details) which were randomly presented in an intermixed fashion in each block. All experiments were tested in 15 blocks of 36 trials, resulting in a total of 540 trials per experiment. Before each experiment, participants practised the respective task for one full block of 36 trials. These practice data were not recorded or analysed.

2.2.3 Data analysis

Individual WM capacity K was calculated for each participant separately for each task condition. Following the suggestions of Rouder et al. (2011), K was calculated according to Pashler's equation (Pashler, 1988):

$$K = N\left(\frac{h - f}{1 - f}\right)$$

where N is the number of memory items (set size), and *h* and *f* are the measured hit and false alarm rates, respectively. Hit rates were calculated as

$$h = \frac{hits}{hits + misses}$$

and false alarm rates were calculated as

$$f = \frac{false\ alarms}{false\ alarms + correct\ rejections}$$

To test whether mean K values changed between perceptual task conditions, K values were submitted to repeated-measures ANOVAs. Effect sizes are reported as Cohen's *d* (Cohen, 1988) with a CI of 95% for *t*-tests, and as partial eta squared (η_p^2) for *F*-tests. When necessary, Greenhouse-Geisser corrections were applied to *F*-tests, and Bonferroni corrections to *t*-tests. All *t*-tests were two-tailed. In Experiments 2-5, K values were first calculated separately for different set sizes (3,4, and 5) but were then averaged into a more reliable $K_{(mean)}$ value for each participant (see Dai et al., 2019, for similar procedures). To test whether individual WM capacity was correlated between task conditions, individual K or

$K_{(mean)}$ values, respectively, were furthermore analysed by means of Pearson's correlations (r). All statistical analyses were conducted with JASP statistical software (version 0.95.0.0).

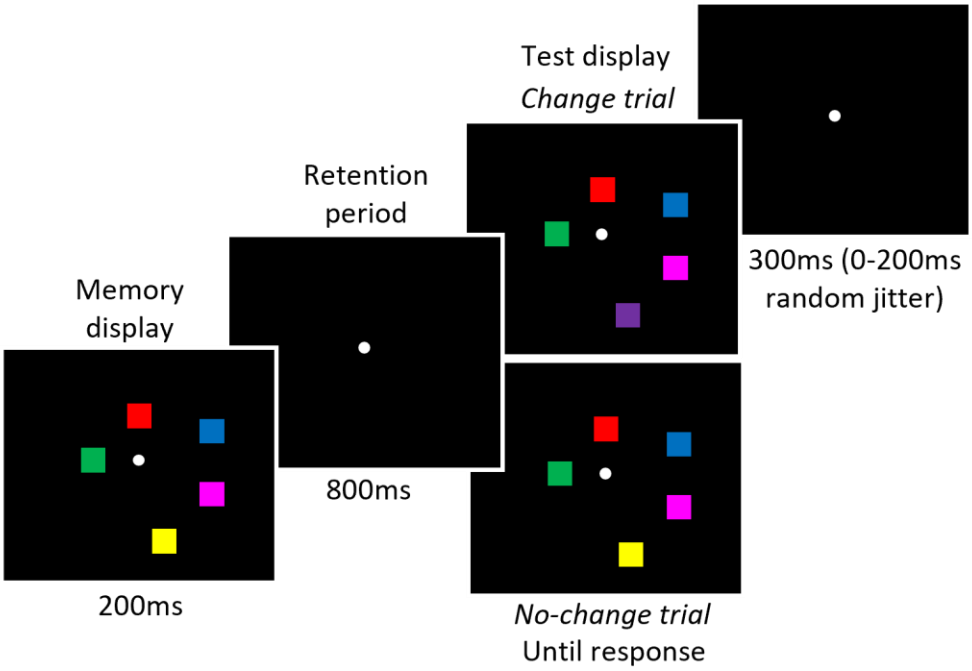


Figure 2 Schematic illustration of the time course of events and example stimuli used in the change detection tasks of Experiments 1 to 5. Each trial contained a memory display followed by a blank retention period and a test display in which the memory items were either identical (no-change trial) or one of them was changed (change trial).

2.3 Experiment 1: Stimulus density

Human perception of the visual world is constrained by the ability to distinguish its individual elements, and spatial resolution is one of the fundamental limitations in object recognition. Object recognition is hampered in dense as compared to sparse stimulus displays and this effect is known as visual crowding. Visual crowding is essentially the inability to isolate individual objects from a cluttered visual scene, i.e., crowded objects are perceived as fuzzy and tangled together. Even though crowding is assumed to be an impairment of object identification, not of object detection (Levi et al., 2002), it nonetheless imposes a fundamental constraint on conscious visual perception (see Whitney & Levi, 2011, for a comprehensive review). Although certain cognitive mechanisms, such as attention (Gong et al., 2023; Scolari et al., 2007) and perceptual learning (Zhu et al., 2016), can override crowding effects, crowding remains a strong and reliable effect. It has been consistently demonstrated in a variety of experiments, including those using artificial stimuli such as Gabor patches (Poder & Wagemans, 2007), Vernier stimuli (Chicherov et al., 2014), basic shapes such as alphanumerical characters (Strasburger et al., 1991), and naturalistic objects and scenes (Ringer et al., 2021). Accordingly, crowding effects have been found between basic features, such as colour and orientation (e.g., Greenwood & Parsons, 2020), between different parts of the same object (e.g., Martelli et al., 2005), between whole different objects (e.g., Wallace & Tjan, 2011), and within global configurations across all elements in the entire visual field (e.g., Herzog & Manassi, 2015). Taken together, the occurrence of crowding under the variety of occasions suggests that there is no single bottleneck of crowding, but that crowding can occur at multiple levels of the visual analysis. With this respect, it would not be surprising if crowding, in the sense of more densely populated memory and test displays, would affect vWM performance in change detection tasks. Two recent studies provided evidence for this assumption (Tamber-Rosenau et al., 2015; Yoruk et al., 2020). Both studies used a standard crowding paradigm and measured memory accuracy as a function of crowding, i.e., the distance between memory items. In both studies, vWM performance decreased with decreasing distance between the memory items, suggesting that visual perception and vWM share the same spatial resolution.

In our Experiments 1a and 1b, we took the reversed approach and manipulated crowding, i.e., the number of items presented in a given visual space, in a standard change detection task allowing us to calculate and correlate K as a measure of vWM capacity. In Experiment 1a we tested the stimulus density hypothesis by presenting observers with set sizes of 3, 5, and 7 items. Our initial findings in Experiment 1a revealed that K values were significantly lower at set size 7 than set size 3, suggesting that increased stimulus density hampered stimulus identification and therefore affected vWM performance negatively. However, because stimulus density was confounded with memory load in Experiment 1a, we tested Experiment 1b to isolate the effect of density. Memory load was kept constant at 3 items but task irrelevant distractors were added to increase stimulus density in the memory and test displays.

2.3.1 Methods

Memory and test displays in both versions of Experiment 1 contained either three, five, or seven squares (0.5° x 0.5°). Set size 3, 5, and 7 trials were equiprobable but presented in an intermixed fashion in each block. There were 12 possible stimulus locations, which were arranged on two imaginary rings. The inner ring, presented at an eccentricity of 0.7° from central fixation, contained four possible stimulus locations at the 12, 3, 6, and 9 o'clock positions of an imaginary clock face. The outer ring, presented at an eccentricity of

1.5° from central fixation, contained eight equally spaced possible stimulus locations, two in each quadrant. In each trial, the three, five, or seven stimulus locations, respectively, were randomly chosen from the 12 possible stimulus locations with the restrictions that each ring contained at least one item and that there was at least one item in each hemifield. In Experiment 1a (Figure 3A), each memory item had a different colour which was randomly chosen (without replacement) from a set of nine possible stimulus colours: red (.583/.304), pink (.396/.254), orange (.396/.346), green (.282/.628), lime (.310/.499), yellow (.400/.422), blue (.153/.081), cyan (.188/.152), and purple (.260/.136). In Experiment 1b (Figure 3B), all set size conditions contained three coloured squares, chosen randomly from the above colour set. The additional two and four items in the set size 5 and set size 7 conditions were uniformly grey (.305/.314). All colours were equiluminant (~9.3cd/m2). In both experiments, participants were instructed to memorise the colours of the squares in the memory display and compare them with the colours shown in the test display. In no-change trials (50%), the colours in the test display were identical to the colours used in the memory display. In change trials (50%), one of the colours changed in the test display. This replacement colour was randomly chosen from the remaining colours that were not used for any of the memory items. In each block, Experiment 1a and 1b each contained six trials for each combination of set size (3, 5, and 7 memory items) and trial type (change, and no-change trial), resulting in 36 trials per block.

2.3.2 Results

K values were submitted to a 2 (Experiment: 1a, 1b) x 3 (Set size: 3, 5, 7) repeated-measures ANOVA. There was no main effect of experiment, $F(1,14)=2.3$, $p=.151$, $\eta_p^2=.14$. However, there was a significant main effect of set size, $F(2,28)=11.5$, $p<.001$, $\eta_p^2=.45$. In

Experiment 1a, K values were 2.3 (set size 3), 2.5 (set size 5), and 2.0 (set size 7). In Experiment 1b, K values were 2.2, 2.2, and 2.0 for set sizes 3, 5, and 7, respectively. The Experiment × Set size interaction was not significant, $F(2,28)=2.1$, $p=.144$, $\eta_p^2=.13$. Follow-up *t*-tests showed that K values were similar at set sizes 3 and 5 (collapsed means: $M_3=2.3$; $M_5=2.4$), $t(14)=.9$, $p=1.000$, $d=.16$. In contrast to this, K values were significantly smaller in the set size 7 ($M_7=2.0$), as compared to the set size 3, $t(14)=3.8$, $p=.006$, $d=.37$, and set size 5 conditions, $t(14)=4.7$, $p<.001$, $d=.48$ (see Figure 3C). Pearson's correlations revealed that individual K values were strongly positively correlated across the three set size conditions for both Experiment 1a: all $r(13)>.69$, $p<.004$ (Figure 3, panels D-F) and Experiment 1b: all $r(13)>.90$, $p<.001$ (Figure 3, panels G-I).

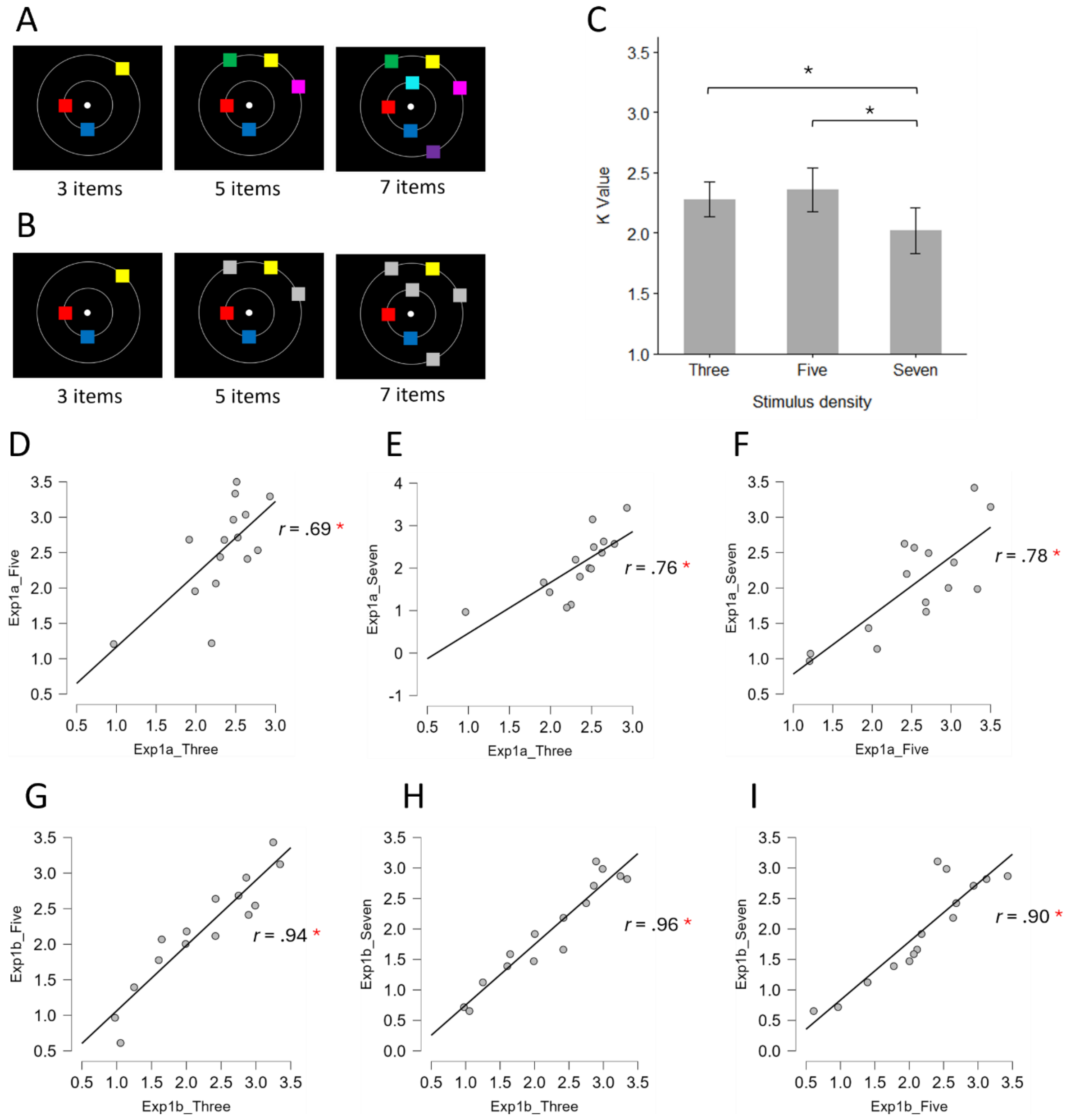


Figure 3 Method stimuli and results of Experiment 1. A-B) Example memory displays of the set size 3, 5 and 7 conditions of Experiment 1a and 1b, respectively; outline circles illustrate the possible stimulus locations in the experiment, but were not visible during the task. C) Mean K values as a function of set size (collapsed across Experiments 1a and 1b). Error bars represent standard error of the mean and significant differences are marked by asterisks. D-F) Scatterplots showing the relationship between K values across the three set size conditions in Experiment 1a. Panels D-F show correlations between K values for the three stimulus pairs: 3 vs 5 (D), 3 vs 7 (E), and 5 vs 7 (F). G-I) Scatterplots showing the relationship between K values across the three set size conditions in Experiment 1b. Panels G-I show correlations between K values for the three stimulus pairs: 3 vs 5 (G), 3 vs 7 (H), and 5 vs 7 (I). Pearson correlation coefficients (*r*) are displayed in each plot, with red asterisks indicating significant correlations (p<.05).

2.3.3 Discussion of Experiment 1

Experiment 1a demonstrated clear crowding effects at the largest set size. When given the opportunity to memorise five rather than three items, participants produced slightly higher K values at set size 5 than at set size 3. However, their overall memory performance did not increase substantially (given the average K of 2.5 which is not yet at ceiling). Importantly, Experiment 1b also revealed a crowding related decline in performance, specifically from increased item density, rather than increased memory load. Across all conditions, participants always memorised three items, but at set sizes 5 and 7, two and four additional task-irrelevant distractors were added to the displays, respectively. As in Experiment 1a, K values were statistically comparable between set sizes 3 and 5, suggesting that participants could perceptually handle five items within the available space. However, at set size 7, the presence of four additional distractors impaired perceptual resolution, reducing vWM performance relative to set sizes 3 and 5. These findings suggest that perception and vWM share common spatial limitations.

Previous work has frequently employed change detection tasks with varying set sizes to assess vWM capacity at the group level. Typically, accuracy remains stable, and K values increase as more items must be retained, but once vWM capacity is reached (around three to four items), accuracy declines and K values plateau or decrease (Luck & Vogel, 1997). Our mean findings from Experiments 1a and 1b align with this pattern. However, individual-level comparisons of K across set sizes have yielded mixed results in the past. For example, Pailian and Halberda (2015) measured individual K values at set sizes below, at, and above vWM capacity (2, 4, and 8 items, respectively) and found no reliable correlations, i.e., an individual’s capacity estimate at one set size did not predict their capacity at another. In contrast, Balaban et al. (2019), analysing change detection data from over 3800 participants, reported reliable correlations between capacity estimates at set sizes 4 and 8, even though capacity estimates varied between set sizes. Notably, Balaban et al. presented items within a constant spatial region, whereas Pailian and Halberda used randomly distributed arrays, spreading out memory items across space. It is therefore possible that they did not observe any systematic changes in K across set sizes because participants used different spatial strategies at the different set sizes to memorise the available items. Our results from Experiments 1a and 1b mirror Balaban et al.'s findings. Even though absolute capacity estimates varied across set sizes, individual differences in K were preserved. This supports the view that individual capacity estimates are not fixed across different set sizes but nonetheless reflect stable individual differences in vWM.

2.4 Experiment 2: Stimulus eccentricity

In Experiment 1, set size was manipulated to introduce crowding effects. In Experiments 2–5, we used set sizes of three, four, and five items (corresponding to just below, at, and just above typical vWM capacity; Cowan, 2001) to avoid ceiling effects and to accommodate individual differences in capacity. K values for the different set sizes in Experiments 2-5 were pooled before statistical data analysis ($K_{(mean)}$). The key manipulation in Experiment 2 concerned the eccentricity at which stimuli were presented in the memory and test displays. Because photoreceptor density and distribution vary across the retina, perceptual acuity declines with increasing distance from fixation (Duncan & Boynton, 2003; Green, 1970). Visual acuity is highest in the fovea, where cone density is approximately 40 times greater than in the periphery, and foveal cones are oversampled by ganglion cells at about four times the rate of peripheral cones (Curcio et al., 1987; Curcio & Allen, 1990). This foveal advantage is further amplified in later processing stages. In the lateral geniculate nucleus (LGN), there are four times as many cells per ganglion input from the fovea compared to the periphery, and in primary visual cortex, there are roughly ten times more cells per foveal LGN projection (Connolly & Van Essen, 1984).

Despite these well-established perceptual asymmetries, relatively few studies have investigated vWM performance as a function of eccentricity. One exception is a multiple object tracking study showing reduced accuracy and slower responses with increasing eccentricity (Vater et al., 2017). To test the role of eccentricity in a change detection paradigm, we presented memory items at three eccentricities. In Experiment 2a, K values decreased with increasing eccentricity, but because stimulus size was held constant, retinal eccentricity was confounded with the size of the cortical representation of the stimuli. Experiment 2b addressed this limitation by scaling stimuli according to cortical magnification, thus equating cortical representation across eccentricities.

2.4.1 Methods

Stimuli and procedures were as described in the General Methods and as in Experiment 1a with the following exceptions. In Experiment 2, memory and test items were either presented on an imaginary ring with a near (0.7°), intermediate (1.5°), or far (3.0°) eccentricity from fixation. There were eight possible equidistant stimulus locations on each ring: four at the 12, 3, 6, and 9 o’clock positions and four more in between these positions (one in each quadrant of the visual field). Three set size conditions (3, 4, and 5 memory items) were tested in each of these eccentricity conditions. The three, four, or five stimulus locations, respectively, were determined randomly in each trial with the limitation that each hemifield contained at least one item. Near, intermediate, and far eccentricity trials with set sizes of 3, 4, and 5 memory items were equiprobable but presented randomly in an intermixed fashion in each block. In Experiment 2a (Figure 4A), all coloured squares, irrespective of their eccentricity from fixation, were identical in size (0.5° x 0.5°). In Experiment 2b (Figure 4B), the stimulus sizes were adjusted in line with the cortical magnification equation by Rovamo and Virsu (1979). Precisely, stimulus sizes were scaled based on the averaged magnification factor M of the nasal ($_N$) and temporal ($_T$) visual field (see Beck & Lavie, 2005, for similar procedures) which were calculated as:

$$M_N = 1 + 0.33 * E + 0.00007 * E^3$$

$$M_T = 1 + 0.29 * E + 0.000012 * E^3$$

where E refers to the eccentricity in degrees of visual angle. This resulted in stimulus sizes of 0.6° x 0.6°, 0.7° x 0.7° and 1.0° x 1.0° in the near, intermediate, and far eccentricity conditions, respectively. Experiment 2a and 2b each contained two trials for each combination of stimulus eccentricity (near, intermediate, and far eccentricity), set size (3, 4, and 5 memory items) and trial type (change, and no-change trial), resulting in 36 trials per block.

2.4.2 Results

$K_{(mean)}$ values were submitted to a 2 (Experiment: 2a, 2b) x 3 (Eccentricity: near, intermediate, far) repeated-measures ANOVA. There was no main effect of experiment, $F(1,14)=.01$, $p=.941$, $\eta_p^2=.00$. However, there was a significant main effect of eccentricity, $F(2,28)=7.9$, $p=.002$, $\eta_p^2=.36$. $K_{(mean)}$ values showed the same pattern across both sessions. In Experiment 2a, $K_{(mean)}$ values were 2.4 (near), 2.3 (intermediate), and 2.2 (far). In Experiment 2b, $K_{(mean)}$ values were 2.4 (near), 2.4 (intermediate), and 2.2 (far). Thus, K decreased systematically with increasing eccentricity in both experiments. The Experiment ×

Eccentricity interaction was not significant, *F*(2,28)=.08, *p*=.922, η_p^2=.01. Follow-up *t*-tests showed that $K_{(mean)}$ values were similar at the near and intermediate eccentricities (collapsed means: M_{Near}=2.4; $M_{Intermediate}$=2.4), *t*(14)=.1.5, *p*=.452, *d*=.11. In contrast to this, $K_{(mean)}$ values at the far eccentricity (M_{Far}=2.2), were significantly lower than both the near, *t*(14)=3.5, *p*=.010, *d*=.27, and intermediate eccentricities, *t*(14)=3.0, *p*=.031, *d*=.16 (Figure 4C). Pearson's correlations revealed that individual $K_{(mean)}$ values were strongly positively correlated across the three set size conditions for both Experiment 2a: all *r*(13)>.91, *p*<.001 (see Figure 4, panels D-F) and Experiment 2b: all *r*(13)>.92, *p*<.001 (see Figure 4, panels G-I).

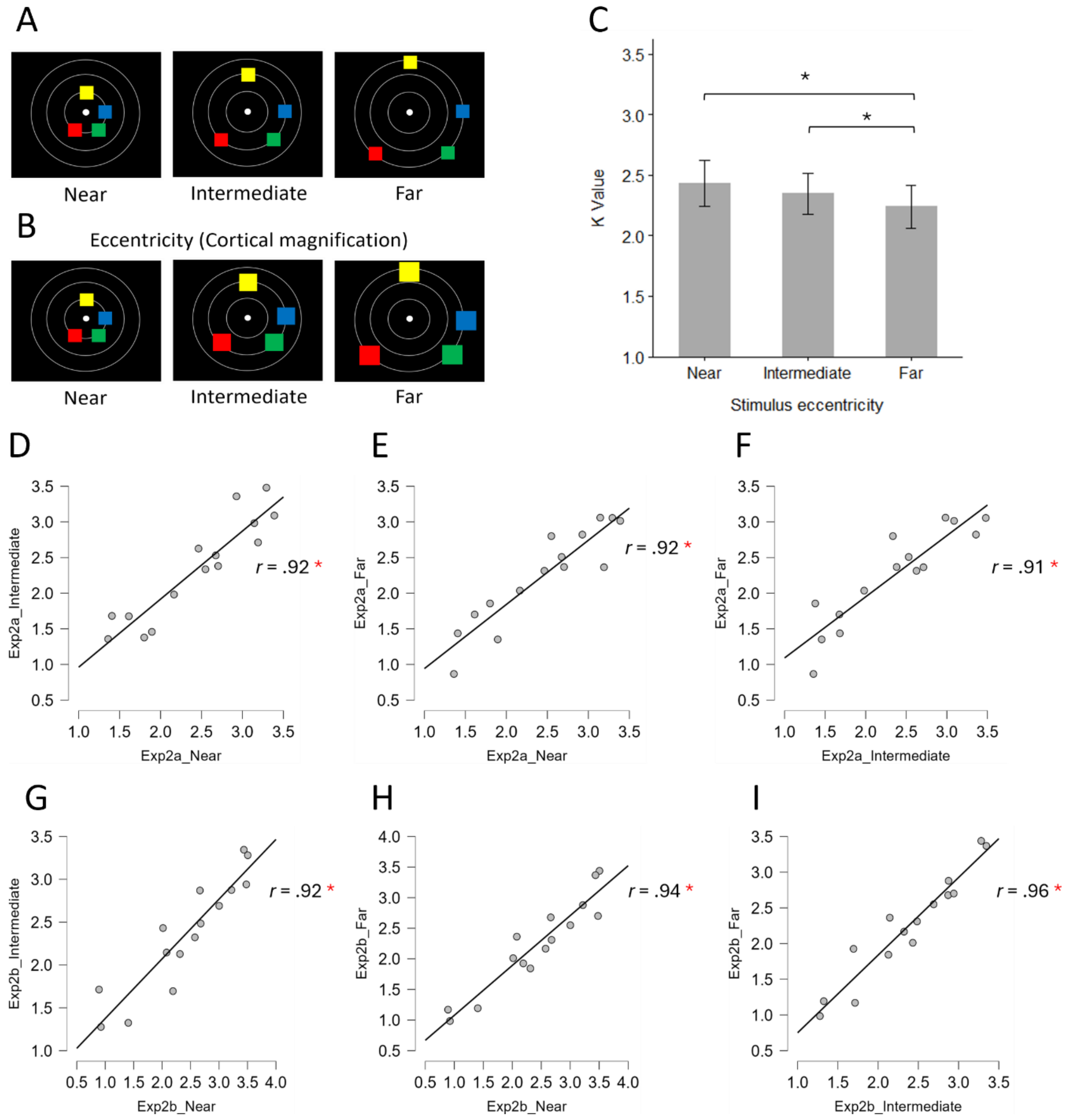


Figure 4 Method stimuli and results of Experiment 2. A-B) Example memory displays of the set size 4 trials with eccentricity near, intermediate and far conditions of Experiment 2a and 2b, respectively; outline circles illustrate the possible stimulus locations in the experiment, but were not visible during the task. C) Mean K values as a function of eccentricity (collapsed across Experiments 2a and 2b). Error bars represent standard error of the mean and significant differences are marked by asterisks. D-F) Scatterplots showing the relationship between K values across the three eccentricity conditions in Experiment 2a. Panels D-F show correlations between K values for the three stimulus pairs: near vs intermediate (D), near vs far (E), and intermediate vs far (F). G-I) Scatterplots showing the relationship between K values across the three eccentricity conditions in Experiment 2b. Panels G-I show correlations between K values for the three stimulus pairs: near vs intermediate (G), near vs far (H), and intermediate vs far (I). Pearson correlation coefficients (*r*) are displayed in each plot, with red asterisks indicating significant correlations ($p<.05$).

2.4.3 Discussion of Experiment 2

The combined results of Experiment 2a and 2b support the hypothesis that vWM performance is modulated by the eccentricity at which memory items are presented. Capacity estimates (K values) declined systematically across eccentricity conditions, being highest at near and lowest at far eccentricity. Crucially, there was no main effect of experiment or interaction, indicating that performance was similar in both 2a and 2b. Thus, the eccentricity effect persisted even when stimuli were size scaled. This pattern suggests that perceptual limitations, such as reduced peripheral colour vision, directly constrained vWM performance.

Only a handful of previous studies have examined eccentricity effects on vWM, and, to our knowledge, none have used a change detection task. Nevertheless, our findings align with related work. For example, Vater et al. (2017) showed in a multiple object tracking task that detection rates and response times varied with eccentricity, with superior performance at near compared to far locations. Interestingly, these effects were more pronounced for form than for motion changes, which the authors attributed to reduced spatial acuity in the periphery and the relatively preserved motion sensitivity of peripheral vision. Similarly, Metha et al. (1994) reported that detection sensitivity for colour changes declined steeply with eccentricity beyond focal points.

A further relevant study was conducted by Velisavljević and Elder (2008), who tested short-term memory for coherent versus scrambled natural scenes. They found that for coherent scenes, memory performance declined steadily with target eccentricity, whereas for scrambled scenes, performance remained above chance but showed no eccentricity effects. The authors argued that low-level factors (e.g., acuity) can therefore not fully account for memory performance across the visual field, and that high-level factors, such as scene context, also play an important role. However, it remains unclear whether such high-level factors are themselves constrained by reduced peripheral resolution. In our study, where no global scene configurations were present, the evidence points more clearly to low-level factors, particularly visual resolution, as direct contributors to vWM performance, yielding different capacity estimates across eccentricity conditions.

Together, Experiments 2a and 2b provide converging evidence that the eccentricity of stimuli imposes a reliable constraint on vWM capacity. Even when perceptual inputs were scaled to compensate for cortical magnification, performance continued to decline with eccentricity.

2.5 Experiment 3: Stimulus organisation

There is ample evidence that participants can increase their vWM performance by pooling information across memory items, a process referred to as chunking (see Nassar et al., 2018, for a review). One possibility to facilitate chunking might be the pooling of memory items based on their perceptual organisation. Gestalt psychologists have long argued that specific organisational patterns of visual input facilitate perceptual grouping and segmentation, enabling the visual system to combine elements into coherent objects and to divide scenes into distinct regions (Wagemans et al., 2012a, 2012b). Evidence from change detection experiments demonstrated improved vWM performance when memory items could be grouped by Gestalt principles such as closure (illusory contours; Li et al., 2018), proximity and connectedness (Peterson et al., 2015; Woodman et al., 2003), similarity (Morey, 2019; Peterson & Berryhill, 2013; Prieto et al., 2022), common fate (Balaban & Luria, 2016; Luria & Vogel, 2014; Luck & Vogel, 1997), and Figure-Ground organisation (Kalamala et al., 2017; Xu & Chun, 2007), even when the tested feature was grouping irrelevant. In fact, in the original study of the change detection paradigm, checkerboard patterns (grids with black and white squares) were used with varying levels of complexity (Phillips, 1974). The results proposed that when stimuli are structured or follow a pattern, they are easier to encode and recall as compared to when the stimuli are less structured or random. Building on this work, in Experiment 3 we compared vWM performance across conditions in which memory items were either presented at random locations or arranged to allow grouping by continuation (items aligned in consecutive lines) or closure (items forming geometric shapes).

2.5.1 Methods

The stimuli and procedures were as described in the General Methods and as in Experiment 1a, only that the stimulus locations in Experiment 3 were no longer chosen from circular arrays, but from a five-by-five cell matrix that was centred at the fixation point (3.0° x 3.0°; with each cell subtending 0.6° x 0.6°). Three, four, and five memory items were used as set sizes, as described in Experiment 2a. There were three stimulus organisation conditions in Experiment 3. In the random (control) condition (Figure 5A, top panel), the three, four, or five stimulus locations in each trial were randomly chosen from the 24 possible stimulus locations of the search matrix (the central position was always occupied by the fixation cross), with the limitation that a maximum of two stimuli were allowed in adjacent cells (horizontally, vertically, or diagonally). In the continuation condition (Figure 5A, middle panel), the three, four, or five stimuli in each trial were always presented in horizontal, vertical, or diagonal adjacent cells of the search matrix. The three continuation directions were equiprobably within each block but were randomly used across the three set sizes. In the closed figure condition (Figure 5A, lower panel), the stimuli were presented in adjacent cells to form a triangle (set size 3), a square (set size 4), or a cross (set size 5). The locations of the continuous stimulus strings and the closed figures within the search matrix were picked randomly in each trial. Experiment 3 contained two trials for each combination of stimulus organisation (random, continuation, and closed figure), set size (3, 4, and 5 memory items) and trial type (change, and no-change trial), resulting in 36 trials per block.

2.5.2 Results

The one-way repeated-measures ANOVA on $K_{(mean)}$ values in Experiment 3 (Figure 5B) failed to produce a main effect of stimulus organisation, $F(2,28)<1$, $p=.613$, $\eta_p^2=.03$. $K_{(mean)}$ values in the random ($K_{(mean)}=2.1$), continuation ($K_{(mean)}=2.1$), and closed figure ($K_{(mean)}=2.2$) conditions were statistically identical. In line with the previous Experiments 1 and 2, Pearson’s correlations produced strong positive correlations between the individual $K_{(mean)}$ values across the three stimulus organisation conditions, $r(13)>.94$, $p<.001$ (see Figure 5, panels C-E).

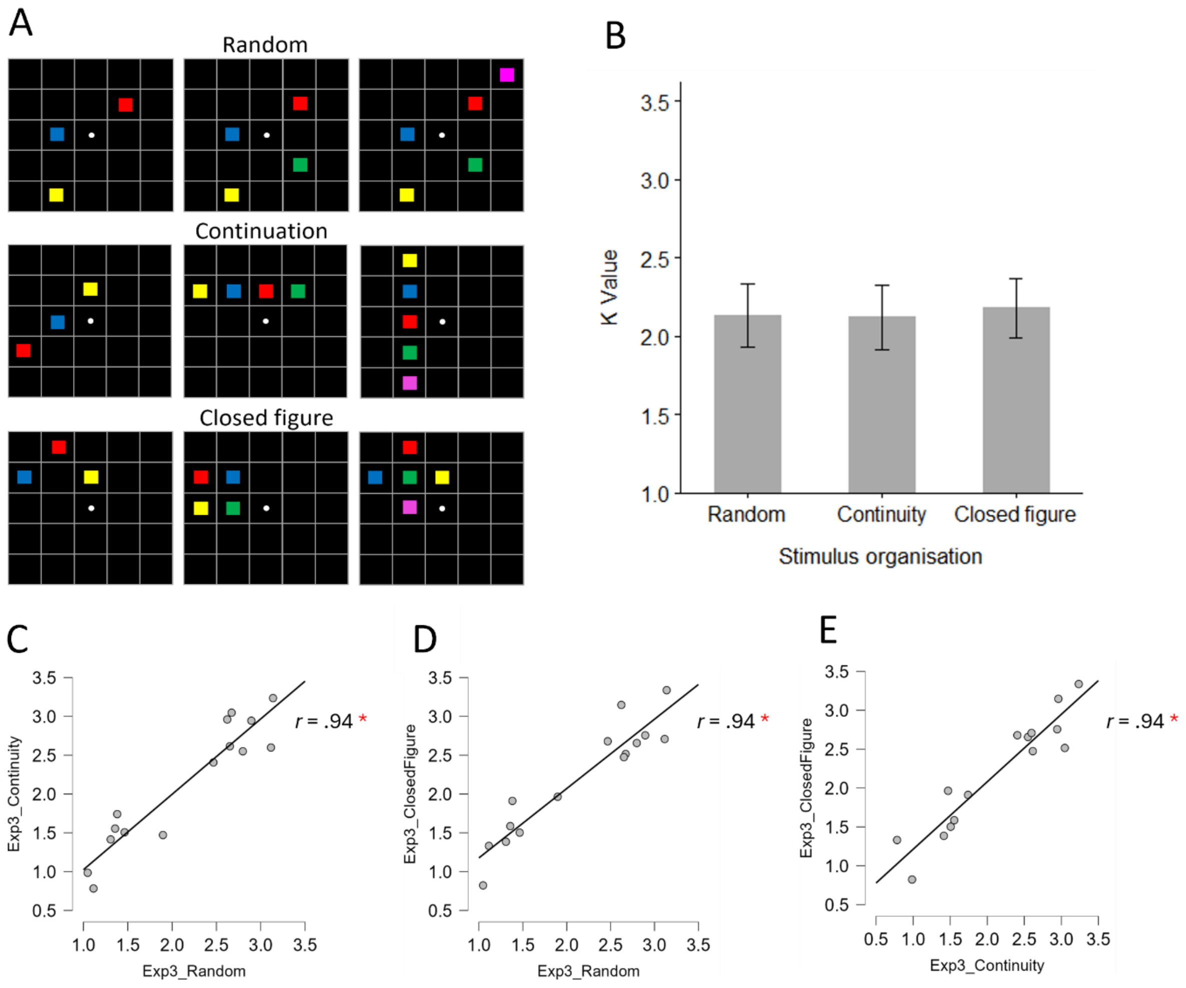


Figure 5 Method stimuli and results of Experiment 3. A) Example memory displays of the set size 3, 4 and 5 trials in the random, continuation and closed figure conditions of Experiment 3; outline grid squares illustrate the possible stimulus locations in the experiment, but were not visible during the task. B) Mean K values as a function of stimulus organisation. Error bars represent standard error of the mean and significant differences are marked by asterisks. C-E) Scatterplots showing the relationship between K values across the three conditions in Experiment 3. Panels C-E show correlations between K values for the three stimulus pairs: random vs continuity (C), random vs closed figure (D), and continuity vs closed figure (E). Pearson correlation coefficients (*r*) are displayed in each plot, with red asterisks indicating significant correlations ($p<.05$).

2.5.3 Discussion of Experiment 3

Contrary to our predictions, and somewhat surprisingly, given the ample evidence for beneficial effects of stimulus grouping in vWM, Experiment 3 did not reveal differences in capacity estimates across stimulus organisation conditions. Based on prior findings, we expected larger K values in the continuation and closed figure conditions compared to randomly organised displays, reflecting enhanced memory performance for grouped stimuli. An important consideration is that the task-relevant feature in our study was colour, whereas the grouping manipulation was based on spatial organisation. This mismatch between grouping and task-relevant dimensions could have attenuated potential effects. However, several previous studies employing similar feature dissociations nonetheless reported grouping benefits in change detection tasks (e.g., Kalamala et al., 2017; Morey, 2019; Woodman et al., 2003; Xu & Chun, 2007). Thus, the absence of effects in our study cannot be fully attributed to this factor alone.

It is also worth noting that not all grouping principles appear to facilitate vWM equally. A recent meta-analysis (Li et al., 2018) demonstrated that grouping based on common fate, common region, or symmetry, often failed to yield reliable improvements in memory performance. Considering this, one possibility is that one of the specific principles we examined, i.e., continuity, represents a relatively weak or less prevalent grouping cue in vWM. To our knowledge, no prior studies have systematically tested continuity as a grouping mechanism in this context, making our findings potentially indicative of principle-specific limits. Furthermore, boundaries between grouping principles are often conceptually blurred. For instance, one might argue that stimulus configurations in our closed figure condition were more consistent with symmetry rather than with the closed figure principle. While the latter has previously been shown to enhance vWM performance, the former has not (Li et al., 2018). If participants predominantly perceived our displays as instances of symmetry rather than closure, this could account for the absence of measurable grouping effects. So, it is possible that the participants in our Experiment 3 simply did not group the stimuli and therefore produced null effects.

Taken together, our null results highlight an important point and show that grouping benefits in vWM are not universal but may depend critically on the specific principle employed. Some principles, i.e., continuity and symmetry in this case, may not be robust cues for improving memory capacity in change detection tasks. Future research will have to disentangle the relative strength of different Gestalt principles in vWM capacity improvement by systematically comparing them within the same paradigm and participant sample.

2.6 Experiment 4: Stimulus type

Not all stimulus features are processed with equal speed. For instance, behavioural evidence (Moutoussis & Zeki, 1997) and electrophysiological findings (Lee et al., 2018) indicate that colour-based attentional guidance often takes precedence over shape-based guidance. This interpretation is consistent with early anatomical research (e.g., Barlow, 1972) and more recent fMRI results demonstrating a clear neural segregation between colour and shape processing (e.g., Lafer-Sousa et al., 2016). Wolfe and Horowitz (2004) argued that colour is an undoubtable attribute to guide attention, shape is a probable attribute, but alphanumerical category is only a doubtful case. But it must be noted that there is evidence suggesting that processing speed differences across feature dimensions can be minimised with appropriate tuning. For example, Grubert et al. (in press) and Jimenez et al. (2024) demonstrated that attentional guidance by colour and shape, or colour and location, respectively, can operate at comparable speeds under specific conditions. In the absence of such dedicated manipulations, however, the question arises whether vWM capacity depends on the feature dimension of the stimuli to be retained. If colour benefits from faster anatomical processing and hence from more efficient attentional engagement, it may lead to enhanced vWM performance compared to shape or alphanumerical stimuli. We directly tested this hypothesis in Experiment 4.

Previous research has shown that vWM capacity decreases with increasing stimulus complexity. Alvarez and Cavanagh (2004) reported higher capacity for simple colours compared to Chinese characters or shaded cubes. However, Awh et al. (2007) argued that this effect might reflect high sample-test similarity rather than complexity per se, as the changed test items may not have been sufficiently distinct from the original memory items. Supporting this interpretation, Jackson et al. (2015) used polygons in a change detection task to disentangle stimulus complexity from similarity and found that similarity, not complexity, determined vWM performance. Building on this, Experiment 4 employed basic colours, shapes, and alphanumerical stimuli with sufficient sample-test dissimilarity to avoid such confounds.

2.6.1 Methods

The stimuli and procedures were as described in the General Methods and as in Experiment 1a, but there were three different stimulus types in Experiment 4 (Figure 6A, top panel). Participants were therefore instructed to memorise the objects (as opposed to the colours) presented in the memory display and compare them with the objects shown in the test display. The colour condition was identical to the colour task of Experiment 1a, only that set sizes were three, four, and five memory items as described for Experiment 2a. In the shape and letter conditions, memory and test items were grey shapes (square, circle, triangle, pentagon, star, heart, cross, diamond, gate; Figure 6A, bottom panel) or letters (G, H, K, L, N, P, S, X, Y), respectively. The three, four, or five memory items in each trial were chosen randomly (without replacement) from the sets of nine colours, nine shapes, or nine letters, respectively. The replacement colour, shape, or letter in change trials was randomly chosen from the set of colours, shapes, or letters, respectively, that were not used for any of the memory items. Experiment 4 contained two trials for each combination of stimulus type (colours, shapes, and letters), set size (3, 4, and 5 memory items) and trial type (change, and no-change trial), resulting in 36 trials per block.

2.6.2 Results

$K_{(mean)}$ values, as shown in Figure 6B, were submitted to a one-way repeated-measures ANOVA, which produced a main effect of stimulus type, $F(2,28)=62.9$, $p<.001$, $\eta_p^2=.82$. Follow-up *t*-tests uncovered that $K_{(mean)}$ values were comparable for colour ($K_{(mean)}=2.7$) and letter stimuli ($K_{(mean)}=3.0$), $t(14)=1.7$, $p=.116$, $d=.29$, but in comparison to these, they were substantially smaller for shape stimuli ($K_{(mean)}=1.6$), both $t(14)>9.3$, $p<.001$, $d=1.71$. Pearson's correlations produced strong positive correlations between the individual $K_{(mean)}$ values across the three stimulus type conditions, $r(13)>.77$, $p<.001$ (see Figure 6, panels C-E).

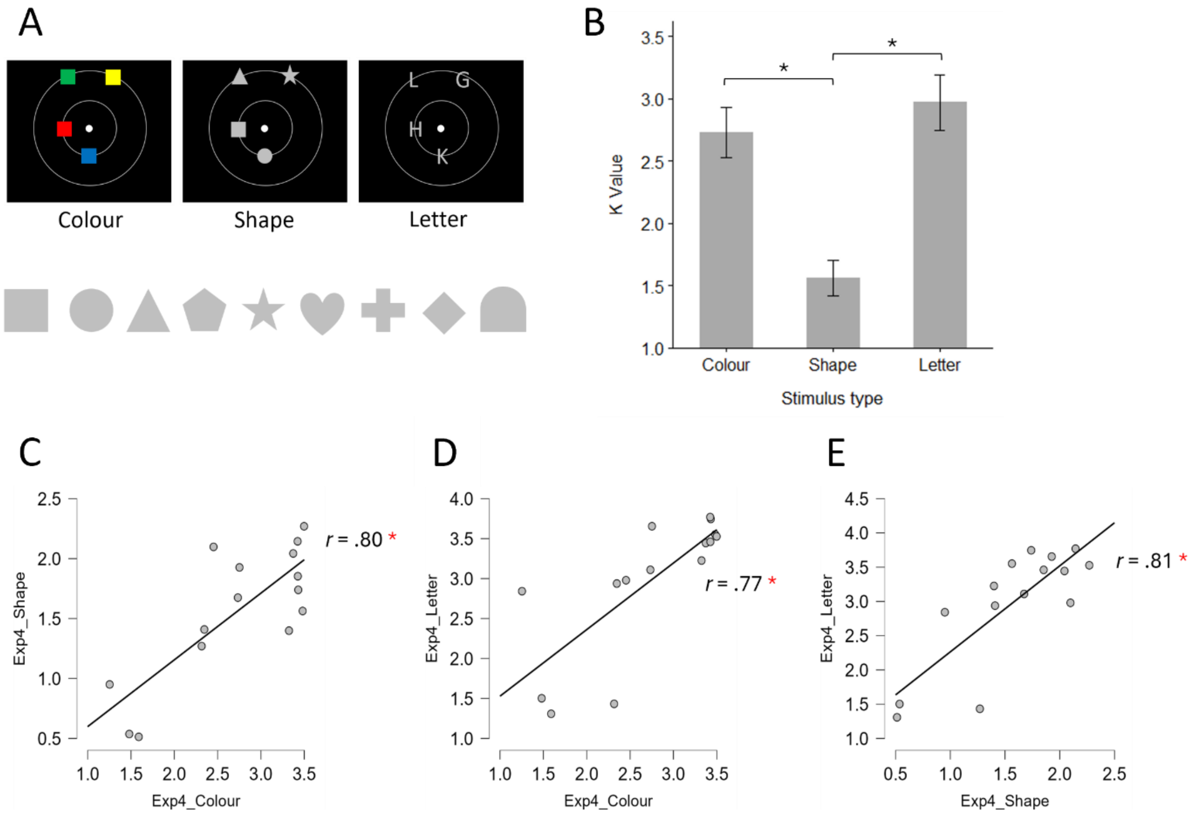


Figure 6 Method stimuli and results of Experiment 4. A) Example memory displays of the set size 4 trials in the colour, shape and letter conditions of Experiment 4; outline circles illustrate the possible stimulus locations in the experiment, but were not visible during the task. Underneath illustrates the nine possible shapes of the shape condition. B) Mean K values as a function of stimulus type. Error bars represent standard error of the mean and significant differences are marked by asterisks. C-E) Scatterplots showing the relationship between K values across the three conditions in Experiment 4. Panels C-E show correlations between K values for the three stimulus pairs: colour vs shape (C), colour vs letter (D), and shape vs letter (E). Pearson correlation coefficients (*r*) are displayed in each plot, with red asterisks indicating significant correlations (*p*<.05).

2.6.3 Discussion of Experiment 4

Experiment 4 partly confirmed our hypothesis: vWM capacity was higher for colour than for shape, consistent with the notion of superior colour processing relative to shape processing (Lee et al., 2018; Moutoussis & Zeki, 1997) and with evidence for the neural segregation of colour and shape pathways (Barlow, 1972; Lafer-Sousa et al., 2016). However, contrary to predictions based on Wolfe and Horowitz (2004), capacity estimates were in fact largest in the letter condition, even exceeding colour memory.

Contemporary theories of visual attention typically posit that early visual analysis is feature-based (Bundesen, 1990; Huang & Pashler, 2007; Itti & Koch, 2001; Mueller & Krummenacher, 2006; Wolfe, 2021). For instance, Guided Search 6.0 (Wolfe, 2021) proposes that attentional deployment is controlled by a feature-driven priority map that dynamically signals the most likely target location. Once a candidate location is selected, features are bound together and compared against stored object representations in long-term memory (see also Cunningham & Wolfe, 2014). Assuming that attention is required to select items for encoding into vWM, one would expect letters to be processed at the object level rather than the feature level, incurring an additional processing step and potentially reducing vWM performance relative to simpler features such as colour.

Yet the opposite was observed. One possible explanation lies in the debate over whether vWM is fundamentally feature-based or object-based. Classic change detection studies have shown that vWM capacity remains constant whether participants retain simple features or feature conjunctions, suggesting successful integration into coherent object representations (e.g., Luck & Vogel, 1997; Vogel et al., 2001). By contrast, electrophysiological studies have provided more nuanced evidence, indicating that conjunctive stimuli may still be represented as separate features (Eimer & Grubert, 2014; Grubert et al., in press), or that binding in vWM is contingent on stimulus parameters and task demands (Balaban & Luria, 2016; Berggren & Eimer, 2018). Letters may constitute a special case since they are highly familiar and nameable, which could facilitate both perceptual encoding and mnemonic retention. Supporting this view, Alvarez and Cavanagh

(2004) found in a multiple object tracking paradigm that while colours yielded the best detection performance, letters followed closely behind.

Taken together, Experiment 4 indicates that perceptual differences between stimulus categories lead to substantial differences in vWM capacity, with some categories (i.e., colours, letters) being more efficiently memorised than others (i.e., shapes). It is worth noting that prior work suggests that these differences cannot be attributed solely to perceptual encoding speed. Eng et al. (2005), who compared several stimulus types in change detection while manipulating encoding time (1000-3000ms), found that longer exposure reduced, but did not eliminate capacity differences. Even after 3000ms (15 times longer than in our paradigm), the differences in K values persisted. This suggests that, beyond perceptual factors, higher-level influences that may benefit rehearsal during longer exposure times, such as nameability, familiarity, or the degree of stimulus abstractness, also play a critical role in shaping vWM performance.

2.7 Experiment 5: Spatial variability of stimuli

Stimulus identity (e.g., colour or shape) and stimulus location are separable attributes that can be selectively encoded and maintained in memory (e.g., Jimenez et al., 2024). At the same time, every feature necessarily occupies a position in space, and several theories propose that spatial location is automatically encoded alongside an object’s featural identity (Logie et al., 2011). Feature Integration Theory (Treisman & Gelade, 1980; see also Treisman & Zhang, 2006, for binding specifically in vWM), for example, assumes that spatial location provides the common reference frame through which features are bound into coherent objects. Without this spatial scaffold, features remain unbound, giving rise to illusory conjunctions (e.g., perceiving a red X when a red O and a green X are presented). In this sense, location may also serve as a retrieval cue in change detection tasks, particularly when items in memory and test displays are presented at identical positions, as is typically the case.

However, the existing literature on the role of spatial location in vWM is inconsistent. Some studies have shown that task-irrelevant changes in stimulus positions between memory and test displays reduce vWM accuracy compared to fixed-location conditions (e.g., Jiang et al., 2000; Logie et al., 2011). These results suggest not only that spatial information contributes to object identification but also that it may be encoded obligatorily, even when detrimental to performance. By contrast, other findings indicate that change detection accuracy is largely unaffected when objects shift locations between sample and test displays (Woodman et al., 2012). In Experiment 5, we addressed these discrepant results by systematically comparing conditions in which both stimulus colours and locations remained constant, colours were randomised while locations were fixed, and locations (and therefore colour-location pairings) were randomised between memory and test displays.

2.7.1 Methods

Stimuli and procedures were as described in the General Methods and as in Experiment 1a with the following exceptions. There were three spatial variability conditions in Experiment 5. The colour/location repetition condition was identical to the colour task of Experiment 1a (Figure 7A, left panel), only that set sizes were three, four, and five memory items as described for Experiment 2a. In this condition, the colours and locations of the squares were identical in the memory and test displays (apart from the colour of the square that changed colour in change trials). In the randomised colour condition (Figure 7A, middle panel), the locations of the squares were identical in the memory and test displays, but all memory colours were randomly shuffled to a new square (no colour repetitions were allowed). In the randomised location condition (Figure 7A, right panel), the coloured squares of the memory display were all presented at a new location in the test display (no location repetitions were allowed). Experiment 5 contained two trials for each combination of spatial variability condition (colour/location repetition, randomised colours, and randomised locations), set size (3, 4, and 5 memory items) and trial type (change, and no-change trial), resulting in 36 trials per block.

2.7.2 Results

The one-way repeated-measures ANOVA on $K_{(mean)}$ values in Experiment 5 revealed a main effect of spatial variability, $F(2,28)=5.7$, $p=.009$, $\eta_p^2=.29$. $K_{(mean)}$ values were virtually identical in the colour/location repetition and randomised colours conditions (both $K_{(mean)}=2.5$), $t(14)=1.0$, $p=.325$, $d=.13$. However, $K_{(mean)}$ values were substantially smaller in the randomised locations condition ($K_{(mean)}=2.2$) as compared to both the colour/location repetition, $t(14)=3.1$, $p=.008$, $d=.49$, and the randomised colours condition, $t(14)=2.3$, $p=.037$, $d=.31$ (Figure 7B). Pearson's correlations produced strong positive correlations between individual $K_{(mean)}$ values across the three spatial variability conditions, $r(13)>.82$, $p<.001$ (see Figure 7, panels C-E).

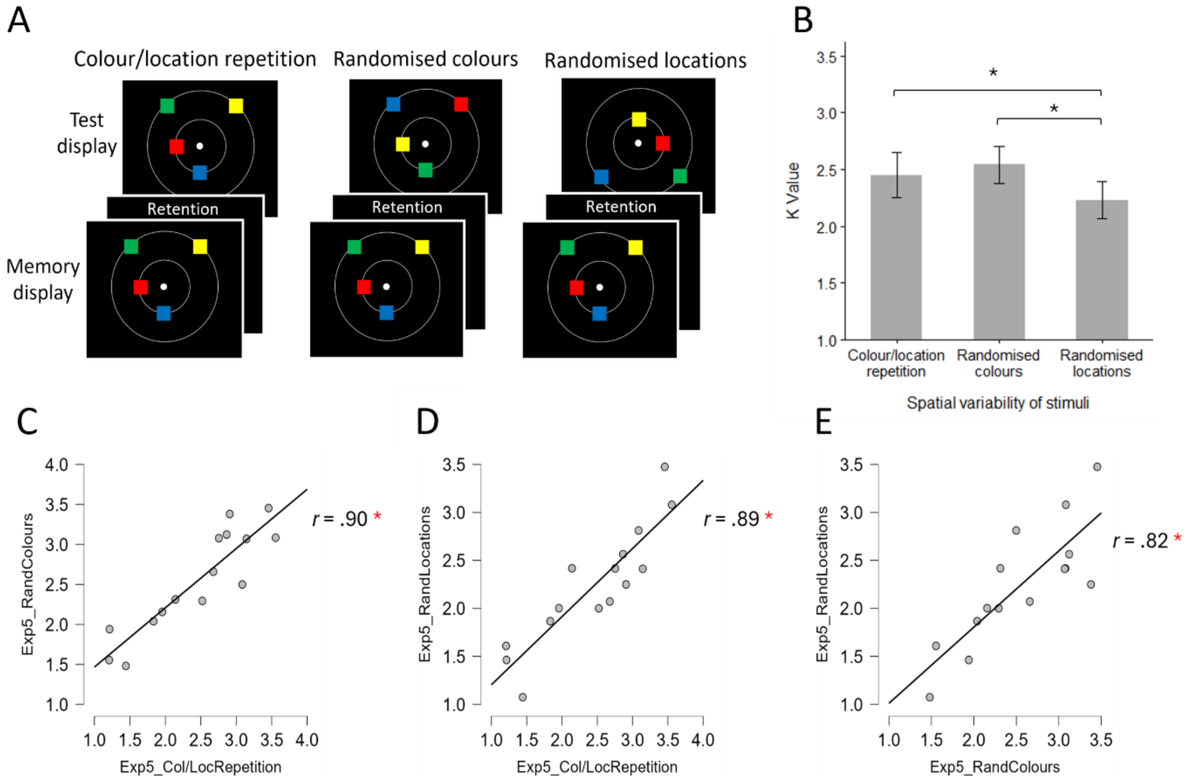


Figure 7 Method stimuli and results of Experiment 5. A) Example memory and test displays of set size 4 trials in the colour/location repetition, randomised colours, and randomised locations conditions; outline circles illustrate the possible stimulus locations in the experiment, but were not visible during the task. B) Mean K values as a function of stimulus spatial variability. Error bars represent standard error of the mean and significant differences are marked by asterisks. C-E) Scatterplots showing the relationship between K values across the three conditions in Experiment 5. Panels C-E show correlations between K values for the three stimulus pairs: colour/location repetition vs randomised colours (C), colour/location repetition vs randomised locations (D), and randomised colours vs randomised locations (E). Pearson correlation coefficients (*r*) are displayed in each plot, with red asterisks indicating significant correlations (*p*<.05).

2.7.3 Discussion of Experiment 5

Experiment 5 revealed improved vWM performance in change detection tasks when stimulus locations remained fixed between memory and test arrays. Crucially, the specific pairing of colours with locations did not matter, i.e., K values were comparable between the colour/location repetition condition and the randomised colours condition. What mattered was spatial predictability, i.e., when spatial consistency was removed entirely (randomised locations condition), capacity estimates dropped significantly. These findings are fully consistent with theories proposing that spatial location is automatically encoded alongside featural identity (e.g., Logie et al., 2011; Treisman & Zhang, 2006). These theories extend previous observations of automatic feature binding in vWM into integrated object representations (Luck & Vogel, 1997; Vogel et al., 2001). If location is simply another feature dimension (Jimenez et al., 2024), then colour and location should be bound automatically. Experiment 5 clearly highlights the role of spatial consistency as a powerful retrieval cue for vWM. vWM performance was reliably better when spatial predictability was preserved, even though location was never task-relevant (participants judged only colour changes). However, if colour-location binding were fully automatic, K values should have been equally low in the randomised colours and randomised locations conditions, since in both cases the colour-to-location mapping was disrupted. Yet this was not the case. K values in the randomised colours condition were comparable to the colour/location repetition condition, suggesting that as long as overall locations and colours were preserved across displays (regardless of their pairing), vWM performance remained equally efficient.

This asymmetry between feature and location binding has also been reported previously. Li and Saiki (2015), using a colour change detection task with a single test stimulus (same or different than the respective memory stimulus), found that colour pre-cues had greater effects on vWM performance than retro-cues, while location pre- and retro-cues were equally effective. These findings can be interpreted as evidence that location encoding, and maintenance share a common mechanism, whereas colour is encoded strongly but may decay more rapidly during retention. In line with our findings, Jiang et al. (2000) also found that random location changes between memory and test displays in change detection resulted in severe impairments of colour retrieval, while random colour changes did not comparably affect vWM performance. They argued that disruption of the global spatial configuration, rather than the mapping of specific items to locations, was the critical factor for successful retrieval.

Together, these findings suggest that spatial consistency acts as a structural scaffold for vWM. Maintaining stable configurations supports retrieval, while disruption of spatial predictability undermines performance, even when location itself is irrelevant. Future work should aim to disentangle the contributions of global spatial structure versus local item-location bindings to better understand how space constrains vWM.

2.8 General Discussion

The present study set out to evaluate the extent to which estimates of vWM capacity (K values) reflect a stable, perceptually invariant cognitive trait, or whether they are influenced by perceptual characteristics of the to-be-remembered stimuli. Across five experiments, we systematically manipulated perceptual factors known to affect visual processing, i.e., stimulus density, stimulus eccentricity, stimulus organisation, stimulus type, and spatial variability of stimuli, and examined both absolute capacity estimates and correlations of K values across conditions.

Taken together, the results provide converging evidence that perceptual properties significantly modulate vWM capacity estimates. In four of our five experiments, absolute K values varied systematically with changes in perceptual input. vWM capacity was reduced under crowded conditions (Experiment 1), at far eccentricities (Experiment 2), for more complex stimulus types (Experiment 4), and when spatial predictability was disrupted (Experiment 5). These findings challenge the assumption that K is a fixed estimation of vWM capacity which is independent of perception. Instead, they highlight that the fidelity of vWM capacity is constrained at least in part by the quality of the initial perceptual encoding, consistent with the sensory recruitment hypothesis (Emrich et al., 2013; Pasternak & Greenlee, 2005; Serences, 2016; Sreenivasan et al., 2014).

At the same time, our results also revealed that relative individual differences were preserved across many perceptual manipulations. That is, participants with higher K values in one condition generally produced higher K values in all other conditions, even when absolute performance shifted. This aligns with previous test-retest reliability studies (Dai et al., 2019; Xu et al., 2018) and suggests that K remains to be a reliable estimate of (trait-like) individual vWM capacity, thus it can be said that perception shapes the level of performance achievable in a given change detection task but does not eliminate stable individual differences.

The present findings have several implications for theories of vWM. First, our findings support claims that capacity does not simply reflect the size of a passive storage space (Balaban et al., 2019; Fukuda et al., 2015; Mance & Vogel, 2013), but instead highlight accounts that view vWM as inherently active, placing perception and attention at the core of WM capacity (Emrich et al., 2013; Pasternak & Greenlee, 2005; Serences, 2016; Sreenivasan et al., 2014). Variability in perceptual input directly translated into differences in memory estimates, demonstrating that what appears as a capacity limit may partly reflect perceptual bottlenecks rather than storage constraints alone.

Second, our results support claims that vWM capacity is highly stable at the individual level (Balaban et al., 2019; Xu et al., 2018), and that correlations with other cognitive abilities such as fluid intelligence are valid. Even though absolute capacity estimates varied across conditions, individual differences in K were preserved in all experiments.

Third, our results add further evidence to the debate over whether vWM representations are feature-based or object-based (Luck & Vogel, 1997; Vogel et al., 2001). The asymmetry between colour and location manipulations (Experiment 5) suggests that some features provide stronger scaffolds for retention than others, highlighting the need to specify how, and to what extent, different feature dimensions are integrated and maintained. Moreover, the superior retention of letters relative to colours (Experiment 4) indicates that long-term familiarity, categorical structure, and nameability can substantially strengthen object-based representations in vWM, thereby enhancing overall performance.

Fourth, our findings paint an ambivalent picture with respect to accounts proposing that change detection performance relies on the spatially global structure of memory displays (Jiang et al., 2000). On the one hand, vWM capacity was preserved when the overall spatial layout of memory items was retained across memory and test arrays, even when colour-location pairings were disrupted (Experiment 5). On the other hand, vWM performance did not benefit when displays were perceptually organised according to Gestalt principles, relative to randomly distributed items (Experiment 3). Thus, while stable global configurations appear to provide a critical scaffold for retention, not all forms of perceptual organisation translate into measurable memory advantages.

Our findings also carry methodological implications for the use of change detection as a diagnostic tool. If perceptual conditions can systematically modulate K values, researchers must be cautious in interpreting differences in vWM capacity across studies or populations. Apparent capacity deficits might in some cases reflect perceptual limitations rather than core memory constraints. At the same time, the preservation of relative rank-order stability supports the continued use of K as a reliable index of individual differences, if task parameters are carefully considered and reported. Future research could extend this approach by directly comparing perceptual influences across populations with known perceptual or attentional differences (e.g., ageing, developmental disorders, or visual impairments).

In summary, the present study demonstrates that vWM capacity estimates are shaped by perceptual properties of the stimuli, but that individual differences remain robust across these variations. These results suggest that K reflects both a stable cognitive trait and a measure sensitive to perceptual constraints. Recognising this dual nature is essential for interpreting vWM performance and for situating vWM capacity within broader theories of cognition. Overall, our findings highlight that vWM is influenced by low-level perceptual factors such as stimulus density and eccentricity as well as by higher-level representational factors such as familiarity and spatially global layouts, suggesting that memory performance reflects an interplay of perceptual resolution, attentional abilities, and categorical coding. This hybrid view implies that capacity is best understood not as a single fixed number, but as an emergent property arising from the interaction between stable individual differences and the situational constraints of encoding and retrieval.

Chapter 3: Contextual cueing: implicit learning improves visual working memory capacity

3.1 Introduction

Directing attention in visual search is often supported by statistical regularities in the environment, thus search rarely starts from scratch (Oliva & Torralba, 2007). The ability to extract such regularities is a crucial function of our cognitive system, and extensive research has shown that individuals learn to make use of these repetitions to process visual scenes more efficiently (Chun & Turk-Browne, 2008). Learning to associate repeated spatial configurations with target positions is a phenomenon that has been demonstrated widely in the literature. One concept that has been utilised to explain this phenomenon is the contextual cueing effect (Chun & Jiang, 1998), which demonstrates a reduced search time in repeatedly shown displays compared with novel non-repeated displays (Chua & Chun, 2003; Jiang & Chun, 2001; Olson & Chun, 2002). In a typical experiment, participants search for a target letter 'T' amongst a configuration of distractors 'L's and indicate the orientation of the target (left vs right). Unknown to the participants, half of the displays are repeated distractor configurations that consistently predict a target location while the other half of displays are novel configurations. Results show significantly faster response times for repeated configurations than randomly generated novel configurations. Chun and Jiang (1998, 1999) propose that repeated exposure to a specific search array leads to the formation of an implicit memory which reflects learned associations between the targets location and the surrounding visual context. When an incoming image matches one of these representations, the retrieved memory guides attention to the target. What is interesting though, is that this effect occurs without participants awareness of repeated displays or explicit memory of target positions, thus it is said to be implicit. In fact, studies have shown that even when participants are presented with a recognition test at the end of the experiment, they are unable to recognise repeated displays, even though they have seen them many times during the experiment (Colagiuri & Livesey, 2016; Vadillo et al., 2020).

Though not entirely clear, it is assumed that contextual cueing involves at least two different mechanisms (Chun & Jiang, 1998, 1999). Firstly, a relationship between the contextual information in the environment (distractor arrangement) and the task-relevant object (target position) must be learned (‘learning’). Secondly, this information needs to be retrieved from memory and applied to the present environment (‘expression of learning’). The latter suggests that LTM plays an important role in contextual learning. One way that information from LTM can interact with present experience is via vWM (Baddeley & Hitch, 1974; Baddeley 2012). vWM is thought to enable attention to prioritise task-relevant information for future actions (Soto et al., 2014) and has been shown to play a role in implicit learning (Janacsek & Nemeth, 2013; Martini et al., 2013). Therefore, for repeated contexts, despite occurring without awareness, vWM may support the implicit encoding and retrieval of target-distractor configuration information from LTM, suggesting that contextual learning relies on vWM resources (Manginelli et al., 2013).

Researchers have attempted to disentangle whether visual search in repeated contexts depends on vWM resources (Pollmann, 2019), often using dual-task approaches that typically combine a visual search and change detection task. These approaches are commonly used to investigate the potential overlap between vWM and contextual learning. If these processes share the same resources, vWM engagement should interfere with contextual learning. Several studies have explored this relationship by varying vWM load, task type (spatial vs non-spatial) and task timing (simultaneously vs sequentially) within dual-task paradigms (Annac et al., 2013; Manginelli et al., 2012; Pollmann, 2019; Travis et al., 2013; Vickery et al., 2010). The first study of this type was conducted by Vickery et al. (2010). Across three experiments, participants performed visual search tasks in which some of the display configurations were repeated while simultaneously performing vWM change detection tasks that varied in load (Experiment 1), spatial locations (Experiment 2) and the sequence of stimuli (Experiment 3). Despite these added demands, the contextual cueing effect (faster search in repeated displays) remained robust, indicating that it is unaffected by interference from either spatial or non-spatial WM loads. Despite this, several studies have shown that WM tasks do interfere with contextual learning, with spatial WM load being more detrimental to visual search performance than vWM load (Travis et al., 2013; Woodman & Luck, 2004). Manginelli et al. (2013) examined performance on a visual search task with engagement in either a non-spatial or spatial vWM task. In the non-spatial vWM task, participants memorised the identity of colours, whereas in the spatial vWM task, participants memorised the location of several squares. The vWM task was introduced either during the initial learning phase or during the later testing phase, while the visual search task was performed in both phases. They found that a simultaneous spatial vWM task reduced the contextual cueing effect only when it was performed simultaneously with visual search during the test phase, suggesting that spatial vWM affects the expression of contextual learning, rather than with the learning itself. Annac et al. (2013) also found a reduction of the contextual cueing effect when using a spatial vWM task. Contradictory findings emerged in a similar dual-task approach (Travis et al., 2013), examining presentation style. Contrary to Vickery et al. (2010), Travis et al. (2013) presented both the study and test items sequentially, requiring participants to compare the temporal pattern of the learned stimuli and the test stimuli, creating an additional memory load. They found that recruitment of these vWM resources during visual search attenuated the contextual cueing effect.

Taken together, the behavioural literature shows mixed evidence for the role of vWM in contextual cueing, with some studies reporting diminished contextual cueing (Travis et al., 2013) while others find intact contextual cueing (Vickery et al., 2010) or mixed results (Annac et al., 2013; Manginelli et al., 2013). In these studies, vWM demands were either applied throughout the entire testing session or restricted to a learning phase, leaving unclear exactly when during contextual learning, the engagement of vWM resources is most critical. To address this, studies have used event-related potentials (ERPs) to investigate the timing of contextual cueing in visual search (Johnson et al., 2007; Zinchenko et al., 2020). ERP studies have focused on components such as the N2pc, P3, stimulus-locked lateralised readiness potential (sLRP), and response-locked lateralised readiness potential (rLRP). Across the literature, it is suggested that attention shifts more efficiently to a target location when it is presented in a repeated display (Johnson et al., 2007; Olson et al., 2001). For instance, Olson et al. (2001) found a significantly larger N210 in a repeated as opposed to novel condition and Johnson et al. (2007) observed a significantly larger N2pc for repeated than novel displays. Similarly, Zinchenko et al. (2020) found an enhanced N2pc for repeated than novel displays, alongside an earlier N1pc reflecting an orienting response to salient stimuli, influencing the N2pc. However, contradictory to the above findings, Schankin and Schubo (2009) found no significant differences between novel and repeated displays for the N2pc, rLRP or sLRP. Additionally, Kojouharova et al. (2023) have reported mixed ERP results, concluding that contextual cueing mechanisms depend on age. In younger adults, an early and intermediate locus was detected, symbolising effective attentional selection and stimulus categorisation, whereas in older adults, a late locus was identified, suggesting more efficient response organisation led to faster reactions. Given the extensive literature on ERP components related to attentional processes, one would expect substantial ERP research highlighting vWM processes (e.g., the contralateral delay activity (CDA) component), especially since vWM is thought to play a role in contextual cueing. However, this is not the case. In fact, ERPs have explicitly been measured during visual search tasks and not vWM tasks.

The consensus in the literature is that contextual learning is task specific, occurring primarily in visual search tasks where participants are required to locate the target but are not explicitly asked to retain all locations in memory for later recall. In contrast, a change detection task requires participants to actively hold the memory items in mind for subsequent comparison. Since no items are specifically marked as distractors, all items in the display must be maintained in vWM. This distinction raises the question: can we generalise contextual cueing beyond visual search to vWM tasks? Behavioural research has directly investigated contextual cueing directly in vWM by means of a change detection task (Jiang & Song, 2005; Olson et al., 2005). In Olson et al. (2005) study, participants were presented with two spatial arrays containing 6, 9 or 12 items one after the other which were identical except for one location. Participants task was to detect the difference between the arrays, by clicking on the missing location using a mouse. They found that spatial WM significantly increased when the same location changed across repetitions but showed no improvement when different locations changed. They concluded that learning in vWM primarily helps determine which information is retained, rather than directly increasing vWM capacity. Similarly, Jiang and Song (2005) used a set of black filled circles against a white background. After a brief delay, participants were shown the test display and asked if any object had moved or changed its position. They found higher accuracy and response times when the spatial context was repeated compared to when it was novel. Both studies conclude that contextual learning can occur within vWM. However, spatial change detection tasks may introduce potential confounds. Firstly, participants may become more attentive to spatial cues, simply due to having to detect a location change, therefore relying on spatial memory. Secondly, whilst location-based change detection tasks do incorporate specific loads in the arrays, these are usually a high number of items, needed to 'fill' the arrays. Whilst these approaches can detect surface-level changes, they fall short when it comes to capturing the 'true capacity' of vWM, which usually lies between 3-4 items (Cowan, 2001). A colour-based change detection task, that avoids explicit spatial demands and uses moderate vWM loads may better isolate contextual cueing effects. Importantly, incorporating neural measures, such as the CDA may offer further insight into these contexts. The CDA is sensitive to the number of items maintained in WM (Vogel & Machizawa, 2004) so it could help track whether learning facilitates more efficient memory retention. For instance, the CDA might show a reduced neural response for repeated locations compared to locations that change unpredictably, indicating less cognitive load when maintaining learned information.

The goal of the present study was to explore how contextual cueing (repeated vs novel contexts) influence both behavioural performance and neural processing during a change detection task, with an emphasis on task load (two vs four items) and task type (colour-location vs colour-only). In Experiment 1, we investigate how varying task loads (two-colour vs four-colour) and repeated versus novel contexts affect participants vWM capacity (K) and accuracy. We also examine how these factors influence neural activity, specifically focusing on the CDA, a well-established ERP marker of WM maintenance (Luria et al., 2016). In Experiment 2, we will compare the same behavioural and neural responses across two different tasks, a colour-location task and a colour-only task. If contextual cueing directly relies on the location of vWM, inhibiting location in the colour-only task should reduce the effect of context repetition. Behaviourally, for repeated displays, vWM can not only access the previous perception, but also the LTM trace. However, for novel displays, vWM relies on perception alone. Therefore, one can assume that vWM would be enhanced for repeated displays, as reflected in higher K estimates and improved accuracy. At the neural level, if CDA amplitude varies depending on whether the context is novel or repeated, this may indicate that contextual cues are actively shaping memory processes. When the context is repeated, these cues may make it easier to maintain relevant information, thereby reducing the mental effort required and thus resulting in a smaller CDA.

3.2 Experiment 1

3.2.1 Methods

3.2.1.1 Participants

Thirty participants took part in Experiment 1 and either received £10/hour or course credits as a compensation for their time. All experimental procedures were approved by the Ethics Committee of the Psychology Department at Durham University. Participants gave informed written consent prior to testing. Four participants were excluded from analysis due to excessive eye movement artifacts (>40% of trials lost during artifact rejection; a priori criterion). The remaining 26 participants were between 18 and 43 years of age (M_{age} = 23.8, SD_{age} = 5.7). Eighteen participants were female and eight were male. Twenty-one participants were right-handed and five were left-handed. All participants had normal or corrected-to-normal vision and no known colour deficiency (tested with the Ishihara colour vision test; Ishihara, 1972). The sample size of 26 was calculated by means of an a priori power analysis (G*Power 3; Faul et al., 2007) used to detect an interaction in a 2×2×2 factorial repeated-measures ANOVA (within subjects factors; laterality, load and context) with an assumed alpha of .05, power of .95, and a medium effect size of .30.

3.2.1.2 Stimuli and procedures

Participants were sat in a dimly lit sound attenuated Faraday cage with an approximately 90cm viewing distance from the monitor. Stimuli were presented on a 22-inch MSI Optix G272 LCD monitor with a 100-Hz refresh rate and a resolution of 1920×1080 pixels. MATLAB (Psychophysics Toolbox) was used on an LG Pentium PC running under Windows 10 to control stimulus presentation, timing, and response collection. Stimuli were presented on a black background. A constant central grey fixation cross was shown throughout each experimental block (CIE x, y colour coordinates: .328/.349; 0.2°× 0.2° of visual angle).

Each trial started with the presentation of a memory display for 200ms, which was followed by a blank 800ms retention period and a test display for 200ms (Figure 8, top panels). The inter-trial interval (between the offset of the test display in trial n-1 and the onset of the memory display in trial n) showed a blank screen and was temporally jittered between 2000-2300ms. There were two equiprobable memory load conditions which were presented in a blocked fashion. In the two-colour task, the memory and test displays each contained two differently coloured squares in each hemifield (Figure 8, top left panel). In the four-colour task, they contained four differently coloured squares in each hemifield (Figure 8, top right panel). Squares subtended 0.3 × 0.3 degrees of visual angle. The stimulus areas in each hemifield measured 1.8 × 1.8 degrees of visual angle and they were located at +/-0.1° to +/-1.9° to the right/left of the y-axis, respectively, and 0.9° to -0.9° above and below the x-axis. The locations of the two or four squares in each hemifield were determined independently and randomly in each trial. They could be located anywhere within the stimulus areas (there were no pre-defined stimulus locations), with the exception that they were not allowed to overlap and that there was always at least a 0.1° gap between them.

Participants task was to memorise the colours of the squares on the left or right side of the memory display and compare them to the colour set shown on the same side of the test display. The relevant memory side alternated between blocks and participants were reminded about the upcoming relevant memory side at the beginning of each block. In no-change trials, the colours of the test display were identical to the colours in the memory display. In change trials, one of the squares in the test display had a new colour which was not previously present in the memory display. Participants were instructed to press the up or down arrow key on a standard keyboard if they detected a colour change or felt that the colours in the test display were the same as in the memory display, respectively. The response-to-key mapping (change/no-change response on arrow up/down key) and the hand-to-key mapping (left/right hand on arrow up/down key) was counterbalanced across participants but was kept constant for each participant for the duration of the whole experiment. The different colours for the two or four squares on the relevant memory side were randomly chosen from a set of six possible memory colours (without replacement): red (.616/.320), green (.267/.638), blue (.169/.166), yellow (.438/.490), pink (.474/.244) and cyan (.210/.276). The different colours for the squares on the irrelevant memory side were then determined from the remaining set of memory colours plus brown (.522/.399) and grey (.328/.349), which only ever served as colours for the squares on the irrelevant side of the memory display. All colours were equiluminant (~10.2 cd/m2).

There were two equiprobable and intermixed task conditions that were concerned with the configuration of the squares on the relevant memory side. In novel configuration trials, the colours and locations of the two or four relevant memory squares were chosen randomly from the set of memory colours and available locations in the relevant stimulus area as described above. In repeated configuration trials, the colours and locations of the two or four relevant memory squares were kept constant across all trials of one block. A new repeated configuration was randomly generated at the start of each block. In both novel and repeated configuration trials, the colours and locations of the squares on the irrelevant side of the memory displays were always determined randomly following the rules described above.

Each experimental block condition contained 48 trials, i.e., six trials for each combination of stimulus configuration (novel, and repeated configuration), trial type (change, and no-change trial), and relevant memory side (left, and right side). Ten blocks were tested for each memory load condition (two-colour, and four-colour task), for a total of 960 trials in Experiment 1. Thirteen participants completed the two-colour task first and then the four-colour task, and vice versa for the other 13 participants. Before each load condition, participants practised the two- and four-colour tasks, respectively, until they felt comfortable with them (usually after one or two blocks). These practice data were not recorded.

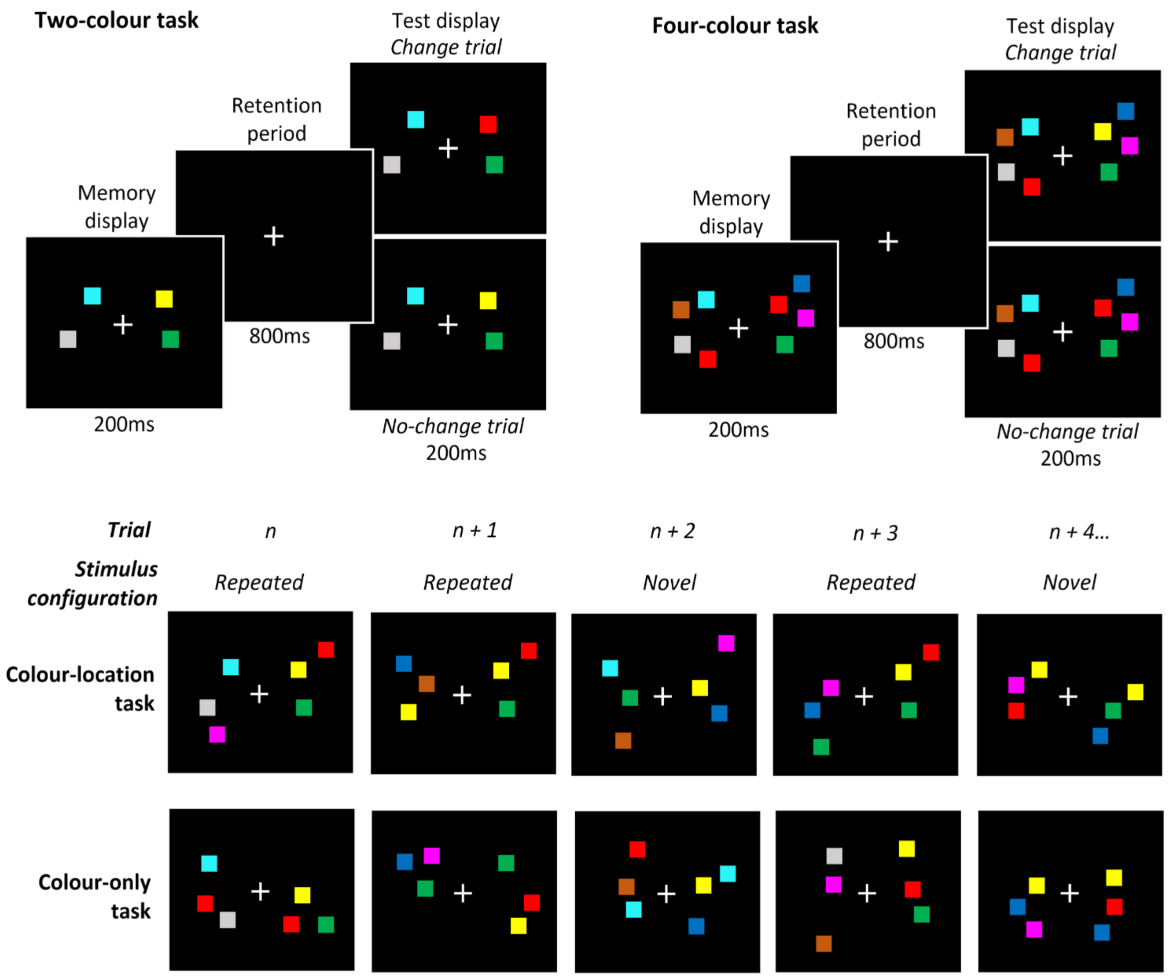


Figure 8 Top panel: Schematic illustration of the stimuli and time course of events in the two-colour (left) and four-colour (right) tasks of Experiments 1. Each trial started with a memory display, followed by a blank retention period, and a test display in which the colours of the memory squares were either identical (no-change trial) or one of them changed (change trial). The relevant memory side alternated between blocks (right in this figure). The stimuli and time course were identical in Experiment 2, only that memory set size was three. Bottom panel: Example sequences of memory displays with novel versus repeated stimulus configurations in Experiment 2. Rows illustrate example configurations only and do not represent the actual sequence of trials. In the colour-location task both the colours and locations of the relevant memory squares were repeated across trials. In the colour-only task, only the colours of the relevant memory squares were repeated, their locations were randomised.

3.2.1.3 Data analyses and EEG recordings

Error rates were used to calculate K values separately for trials with novel and repeated stimulus configurations in both the two- and four-colour tasks. Following the suggestions of Rouder et al. (2011), K was calculated according to Pashler's equation (Pashler, 1988):

$$K = N\left(\frac{h-f}{1-f}\right)$$

where N is the number of memory items (set size), and *h* and *f* are the measured hit and false alarm rates, respectively. Hit rates were calculated as

$$h = \frac{hits}{hits + misses}$$

and false alarm rates were calculated as

$$f = \frac{false\ alarms}{false\ alarms + correct\ rejections}$$

EEG was DC-recorded from 23 scalp electrodes (EasyCap, Brain Products) at standard positions of the extended 10/20 system, and two HEOG electrodes placed at the outer canthi of the eyes. EEG data were recorded using the Brain Vision Recorder (BrainAmp DC amplifier, Brain Products GmbH, Gilching, Germany) at a sampling rate of 500Hz with a 40Hz low-pass filter. No other filters were applied after data acquisition. Impedances were kept below 5kΩ. The left earlobe served as online reference during data acquisition, but all channels were re-referenced offline to linked earlobes. The EEG was segmented into 1100ms time windows including a 100ms pre-stimulus baseline and a 1000ms ERP time window following the onset of the memory displays. Data from incorrect responses were excluded from ERP analyses. Segments that contained eye movements (exceeding ±25µV in the bipolar HEOG channel), blinks (exceeding ±60μV at Fpz), and muscular movements (exceeding ±80μV in all channels) were also excluded. Artefact rejection resulted in an exclusion of 21.1% of all trials in the two-colour task (*SD*=10.5%; ranging between 3.1% and 36.9% across participants) and 19.5% of all trials in the four-colour task (*SD*=8.7%; ranging between 4.2% and 35.2% across participants). The remaining segments were averaged separately for left- and right-side memory squares in trials with novel and repeated stimulus configurations of the two- and four-colour tasks. CDA components were quantified based on ERP mean amplitudes obtained at lateral posterior electrodes PO7 and PO8, contralateral and ipsilateral to the relevant memory side, in the 400-800ms interval after memory display onset. Effect sizes are reported as Cohen's *d* (Cohen, 1988) with a CI of 95% for *t*-tests, and as partial eta squared (η_p^2) for *F*-tests. When necessary, Greenhouse-Geisser corrections were applied to *F*-tests, and Bonferroni corrections to *t*-tests. All *t*-tests were two-tailed.

3.2.2 Results

3.2.2.1 Behavioural results

Error rates were fed into an omnibus repeated-measures ANOVA with the factors memory load (two-colour vs four-colour) and stimulus configuration (novel vs repeated). The ANOVA produced main effects of memory load, $F(1,25)=172.6$, $p<.001$, $\eta_p^2=.87$, and stimulus configuration, $F(1,25)=67.6$, $p<.001$, $\eta_p^2=.73$, since error rates were lower in the two- as compared to the four-colour task, and in trials with repeated versus novel stimulus configurations. There was also a significant interaction, $F(1,25)=7.6$, $p=.011$, $\eta_p^2=.23$, but follow-up *t*-tests confirmed reliable learning effects (i.e., lower error rates for repeated than novel stimulus configurations) both in the two-colour (7.3% vs 8.7%) and four-colour tasks (16.9% vs 21.1%), both $t(25)>2.7$, $p<.012$, $d>.53$ (Figure 9A). K values were subjected to the same ANOVA, which also produced main effects of memory load, *F*(1,25)=118.2, *p*<.001, η_p^2=.83, and stimulus configuration, *F*(1,25)=50.8, *p*<.001, η_p^2=.67, as well as a significant interaction, *F*(1,25)=19.4, *p*<.001, η_p^2=.45. Follow-up *t*-tests demonstrated that the increase in vWM capacity (K) measured in trials with repeated as compared to novel stimulus configurations was small but substantial both in the two-colour (1.8 vs 1.7) and four-colour tasks (2.8 vs 2.5), both *t*(25)>2.4, *p*<.025, *d*>.47 (Figure 9B).

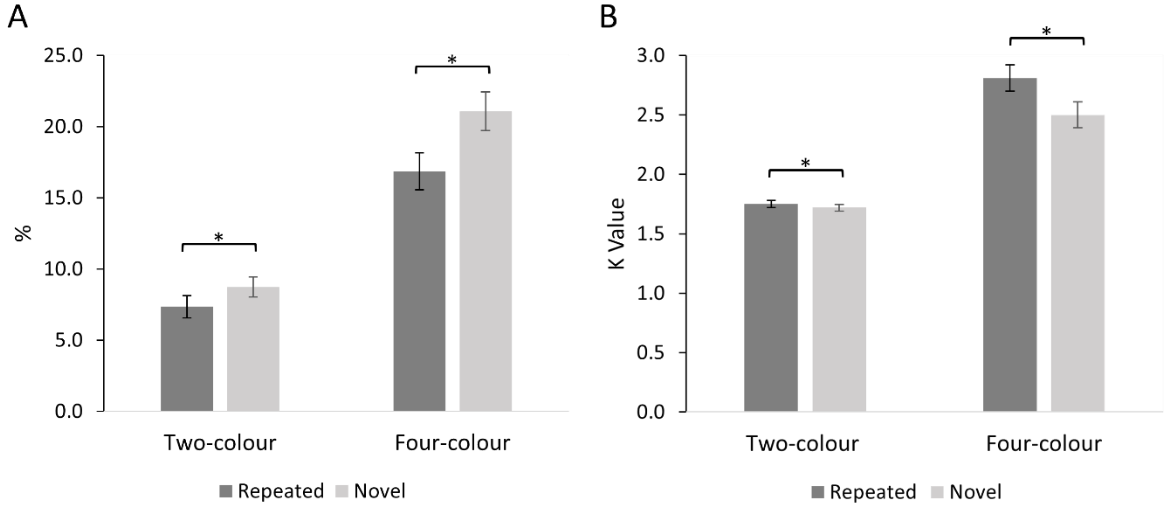


Figure 9 Behavioural performance as a function of task (two-colour, four-colour) and stimulus configuration (repeated, novel). A) Mean error rates (%) for the two-colour and four-colour task for repeated and novel configurations. B) Mean K values for the two-colour and four-colour task for repeated and novel configurations. Error bars represent the standard error of the mean and significant differences are marked by asterisks.

3.2.2.2 ERP results

Figure 10 shows grand-averaged ERPs measured at lateral posterior electrodes PO7/8 contralateral and ipsilateral to the relevant side of the memory display, separately for novel (top panels) and repeated stimulus configurations (middle panels) in the two-colour (left panels) and four-colour tasks (right panels). CDA components were elicited in all task conditions, but they seemed to be attenuated in repeated as compared to novel stimulus configurations. This is further illustrated in the bottom panels of Figure 10, which shows the difference waveforms obtained by subtracting ERPs at ipsilateral electrodes from contralateral ERPs. Statistically, this observation was confirmed by means of an omnibus repeated-measures ANOVA with the factors memory load (two-colour vs four-colour), stimulus configuration (novel vs repeated), and laterality (electrode left vs right of the relevant memory side). The ANOVA produced a main effect of laterality, $F(1,25)=5.8$, $p=.024$, $\eta_p^2= .19$, which confirmed the presence of reliable CDA components. Most importantly, laterality interacted with stimulus configuration, demonstrating that CDA amplitudes were significantly smaller in trials with repeated (-0.4µV) as compared to novel stimulus configurations (-0.9µV), $F(1,25)=12.5$, $p=.002$, $\eta_p^2=.33$. There were no interactions involving the factor memory load, all $F(1,25)<1$, $p>.450$, $\eta_p^2<.02$, suggesting that the learning effects (i.e., smaller CDA amplitudes for repeated than novel stimulus configurations) were statistically comparable in the two- and four-colour task. This was confirmed by two follow-up ANOVAs with the factors stimulus configuration and laterality, conducted separately for the two-colour and four-colour task. Both ANOVAs produced significant interactions, demonstrating that CDA amplitudes were smaller for repeated than novel stimulus configurations both in the two-colour (-0.4µV vs -0.8µV), $F(1,25)=4.3$, $p=.049$, $\eta_p^2=.14$, and the four-colour tasks (-0.4µV vs -1.1µV), $F(1,25)=6.4$, $p=.018$, $\eta_p^2=.20$.

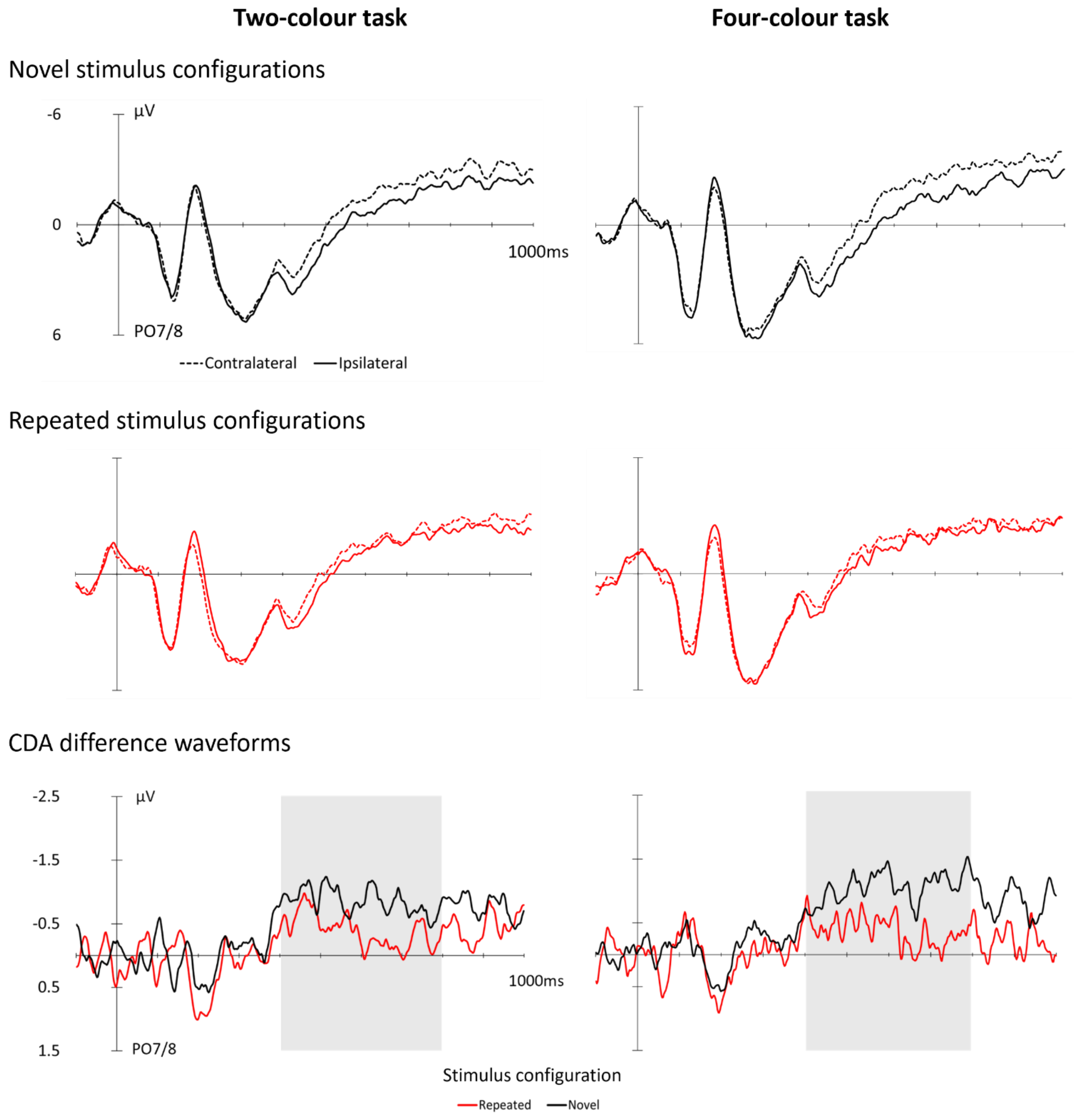


Figure 10 Grand-averaged ERPs measured at lateral posterior electrodes PO7/8 contralateral and ipsilateral to the relevant side of the memory display, separately for novel (top panels) and repeated stimulus configurations (middle panels) in the two-colour (left panels) and four-colour tasks (right panels) of Experiment 1. The corresponding CDA difference waveforms (contralateral minus ipsilateral ERPs) are shown in the bottom panels. Shaded areas mark the CDA time window (400-800ms after memory display onset).

3.2.3 Discussion of Experiment 1

Experiment 1 firstly confirmed our central hypothesis that contextual cueing can occur directly at the level of vWM. Accuracy and vWM capacity (K) was higher for repeated displays than it was for novel displays for both the two-colour and four-colour tasks, suggesting that repeated exposure to consistent specific configurations facilitates more efficient encoding and retrieval from vWM. Simply, repeating displays, made it possible for participants to recognise them implicitly, suggesting that long-term memories of these displays were established (Chun & Jiang, 1998; Vickery et al., 2005). Importantly, this contextual cueing effect emerged across both memory loads, suggesting contextual learning is not constrained by the number of items held in vWM. In fact, the contextual cueing effect was comparable across different memory loads, even though the difficulty of the task increased. Maintaining information in vWM did not impair contextual learning, even at a load of four items, considered to be the upper limit of vWM capacity (Cowan, 2001, 2010; Luck & Vogel, 1997), suggesting that contextual learning seems to operate independently of the storage demands of vWM, irrespective of load. This finding was consistent with that of Vickery et al. (2010) who used varying memory loads and found no effect of vWM load on the contextual cueing effect. However, unlike the dual task approach in Vickery et al. (2010) study, the change detection paradigm used here offered a more direct measure of vWM performance.

At the neural level, as measured by the CDA, no significant difference was observed based on WM load: no difference between two vs four items. However, the CDA amplitude was clearly influenced by context repetition: repeated configurations elicited reduced CDA amplitudes compared to novel configurations. This indicates that contextual cues actively shape memory processes, potentially easing the maintenance of relevant information and reducing mental effort when the context is familiar. One explanation is that in repeated contexts, participants may rely less on active vWM maintenance because the familiar context ‘does the work’ for them. Consequently, even under higher load conditions (e.g., four items), the CDA may not increase as expected. Since the CDA reflects the number of items actively maintained in WM rather than the total number presented (Carlisle et al., 2011), repeated contexts might enable participants to offload some memory demands onto LTM representations. These findings emphasise the role of context consistency in facilitating more efficient WM maintenance, as evidenced by the modulation of CDA by context repetition rather than load.

One limitation of this experiment is that while the CDA provided evidence for reduced memory maintenance on repeated trials, it does not specify which aspects of the context contributed to this benefit. Therefore, perhaps the most crucial explanation, and one which we attempt to address in Experiment 2, is that repeated displays are likely guiding attention to relevant items locations and allowing participants to anticipate what is important. Indeed, the visual system is sensitive to spatial layout and contextual cueing relies on a highly discriminable instance-based memory for spatial configurations (Chun & Jiang, 1998). This implies that individuals automatically encode both identity and location information when processing visual stimuli. Consequently, repeated spatial layouts may serve as a contextual scaffold, allowing participants to anticipate where items will appear, potentially explaining the observed reduction in CDA amplitudes. In contrast, novel displays lack these spatial cues, likely requiring more holistic processing (Jiang & Wagner, 2004).

Experiment 2 was designed to test this by manipulating the spatial stability of repeated displays, specifically, by randomising the positions of repeated items while keeping their identities constant. If the reduction in CDA relies specifically on the repetition of spatial information, then disrupting this location element should eliminate the contextual benefit. Therefore, because both types of trials would now lack spatial predictability, the CDA may be similar for repeated and novel displays. This would help isolate the contribution of spatial context in driving WM efficiency and clarify whether identity repetition alone is sufficient to support the observed neural effects (Brady & Alvarez, 2011).

3.3 Experiment 2

3.3.1 Introduction

Although individuals can obtain perceptual familiarity with repeated spatial configurations (Beesley et al., 2015; Geyer et al., 2010), the effectiveness of these layouts lies predominantly in their predictability of target locations (Olson & Jiang, 2004; Olson et al., 2005). When the spatial layout of distractors is repeated but not paired with a consistent target location, it does not facilitate search (Chun & Jiang, 1998), suggesting there needs to be a consistent association between a target and its layout. Given this, contextual learning can occasionally incur a 'contextual cost' whereby disrupting an association between a repeated context and its target location inhibits search. Indeed, individuals may struggle to find the target in a repeated display if it appears in a location that previously contained a distractor (Makovski & Jiang, 2010).

While changes in object identities do not seem to affect spatial contextual cueing (Chun & Jiang, 1998), altering locations within a repeated context typically results in a pronounced reduction of the contextual cueing effect (Annac et al., 2017; Conci et al., 2011; Conci & Müller, 2012; Makovski & Jiang, 2010; Manginelli & Pollmann, 2009; Zellin et al., 2013). In the context of these findings, Experiment 2 was designed to investigate the role of non-spatial vWM. In Experiment 1, both colour and spatial location were maintained between memory and test displays in repeated configurations. Experiment 2 replicated this procedure in a visuospatial condition, where both colour and spatial location remained constant between memory and test displays, but introduced a new manipulation in a non-spatial condition. In the non-spatial condition, repeated configurations contained the same colours across trials, but the spatial locations of these colours changed. If contextual cueing relies specifically on spatial information, we would not expect any contextual cueing effects in the non-spatial condition.

A critical feature of contextual cueing is that it is driven by implicit memory representations that are acquired incidentally. In typical recognition tests of contextual cueing, repeated displays are presented amongst novel displays and participants are instructed to distinguish between them. However, these recognition tests fail to find evidence of explicit recognition of repeated displays (Goujon et al., 2015). Therefore, while explicit recognition may not always be demonstrated, implicit memory representations of repeated displays support contextual cueing (Vadillo et al., 2020). Based on these findings, an awareness test was added in Experiment 2, in which participants were asked a yes/no question if they were aware of the repeated contexts. This was added to help provide further insights into whether implicit memory (as reflected in the contextual cueing effect) for repeated displays drives the observed effects. Overall, this experiment will allow us to isolate the role of spatial context in contextual cueing, adding clarity to the role of spatial and non-spatial vWM in this process.

3.3.2 Methods

3.3.2.1 Participants

Thirty-one new participants took part in Experiment 2. Participant procedures were identical to Experiment 1. Five participants were excluded due to excessive eye movement activity (>40% trials lost during artifact rejection, a priori criterion). The remaining 26 participants were aged between 18 and 59 years (M_{age} = 25.6, SD_{age} = 10.0). Of those, twenty-one were female and five were male, twenty-five were right-handed, and one was left-handed. All participants had normal or corrected to-normal vision and normal colour vision (as tested with Ishihara, 1972).

3.3.2.2 Stimuli and procedures

Experiment 2 followed the same general procedures as Experiment 1, with the following exceptions. There were two new blocked task conditions in Experiment 2 (Figure 8, bottom panels). In the colour-location task, both the colours and locations of the relevant memory squares were kept constant in repeated stimulus configuration trials (as in Experiment 1). In the colour-only task, only the colours of the relevant memory squares were kept constant in repeated stimulus configuration trials, the locations of these squares were determined independently and randomly in each trial (following the rules described in Experiment 1). The colours and locations of the relevant memory squares in novel stimulus configuration trials were determined randomly as described in Experiment 1. To be able to test Experiment 2 in a single session, we only tested one memory load condition and settled for a set size of three as the happy medium between the two- and four-colour tasks of Experiment 1. Experimental blocks contained 48 trials, i.e., six trials for each combination of stimulus configuration (novel, and repeated configuration), trial type (change, and no-change trial), and relevant memory side (left, and right side). Ten blocks were tested for each task (colour-location, and colour-only), for a total of 960 trials in Experiment 2. Thirteen participants completed the colour-location task first and then the colour-only task, and vice versa for the other 13 participants. Participants practised each task before the experiment proper, and these data were not recorded. After the experiment, we recorded participants' awareness of the repeated stimulus configurations, i.e., they answered yes or no to the following question: “Did you notice that some of the memory displays repeated between trials?”.

3.3.2.3 Data analysis and EEG recordings

All data analyses and EEG procedures were identical to Experiment 1. Artefact rejection resulted in an exclusion of 21.0% of all trials in the colour-location task (*SD*=8.9%; ranging between 2.7% and 36.3% across participants) and 20.4% of all trials in the colour-only task (*SD*=6.6%; ranging between 5.2% and 32.7% across participants). The remaining segments were averaged separately for left- and right-side memory squares in trials with novel and repeated stimulus configurations of the colour-location and colour-only tasks.

3.3.3 Results

3.3.3.1 Behavioural results

Error rates were subjected to an omnibus repeated-measures ANOVA with the factors task (colour-location vs- colour-only) and stimulus configuration (novel vs repeated). There was no main effect of task, $F(1,25)=1.8$, $p=.188$, $\eta_p^2=.07$, but there was a main effect of stimulus configuration, $F(1,25)=48.1$, $p<.001$, $\eta_p^2=.66$, which interacted with task, $F(1, 25)=16.2$, $p<.001$, $\eta_p^2=.39$. Mirroring the results from Experiment 1, follow-up *t*-tests demonstrated that there were reliable learning effects (i.e., lower error rates in repeated than novel stimulus configuration trials) in the colour-location task (11.5% vs 15.3%), $t(25)=8.2$, $p<.001$, $d=1.62$. However, in the colour-only task, error rates were statistically comparable in trials with repeated and novel stimulus configurations (13.7% vs 14.7%), $t(25)=1.9$, $p=.060$, $d=.39$ (Figure 11A). The same statistical pattern was observed for K values, indexing vWM capacity. The ANOVA produced no main effect of task, $F(1,25)=2.2$, $p=.149$,

η_p^2=.08, but there was a main effect of stimulus configuration, *F*(1,25)=51.3, *p*<.001, η_p^2=.67, and a significant interaction, *F*(1,25)=16.2, *p*<.001, η_p^2=.39. K values were substantially larger in repeated than novel stimulus configuration trials in the colour-location task (2.4 vs 2.2), *t*(25)=7.6, *p*<.001, *d*=1.50, but there was no difference in the colour-only task (2.3 vs 2.2), *t*(25)=1.9, *p*=.067, *d*=.38 (Figure 11B).

Awareness test

Twenty-two participants (85%) reported that they were aware of the repeated memory displays in the colour-location task. However, significantly fewer participants (16; 62%) also noticed the repeated colour memory displays in the colour-only condition, *t*(21)=2.8, *p*=.011, *d*=.46.

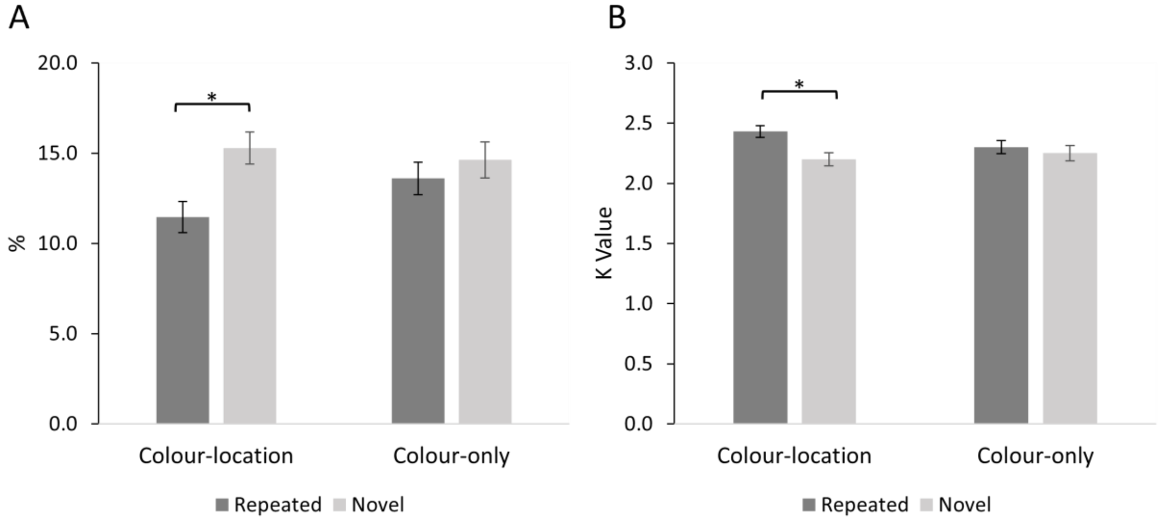


Figure 11 Behavioural performance as a function of task (colour-location, colour-only) and stimulus configuration (repeated, novel). A) Mean error rates (%) for the colour-location and colour-only task for repeated and novel configurations. B) Mean K values for the colour-location and colour-only task for repeated and novel configurations. Error bars represent the standard error of the mean and significant differences are marked by asterisks.

3.3.3.2 ERP results

Figure 12 shows grand-averaged ERPs measured at lateral posterior electrodes PO7/8 contralateral and ipsilateral to the squares on the relevant memory display side, separately in trials with novel (top panels) and repeated stimulus configurations (middle panels) in the colour-location (left panels) and the colour-only task (right panels). CDA components were triggered in all task conditions. From the CDA difference waveforms (Figure 12, bottom panels), CDA components were attenuated in trials with repeated as compared to novel stimulus configurations, but only in the colour-location task. In the colour-only task, CDA components appeared to be identical between repeated and novel stimulus configuration trials. An omnibus repeated-measures ANOVA with the factors task (colour-location vs colour-only), stimulus configuration (novel vs repeated), and laterality (electrode left vs right of the relevant memory side), confirmed these observations. As in Experiment 1, there was a main effect of laterality, $F(1,25)=12.8$, $p=.001$, $\eta_p^2=.34$, which interacted with stimulus configuration, $F(1,25)=9.2$, $p=.005$, $\eta_p^2= .27$, confirming that there were reliable CDAs that differed between trials with novel versus repeated stimulus configurations. The interaction between task and laterality failed to reach significance, $F(1,25)=.3$, $p=.574$, $\eta_p^2=.01$, but there was a reliable three-way interaction, $F(1,25)=8.0$, $p=.009$, $\eta_p^2=.24$, suggesting that the learning effects differed between the two tasks. Follow-up ANOVAs with the factors stimulus configuration and laterality, conducted separately for the colour-location and colour-only task, confirmed the presence of learning effects (i.e., smaller CDA amplitudes for repeated than novel stimulus configurations) in the colour-location task (-0.4µV vs -0.9µV), $F(1,25)=13.4$, $p=.001$, $\eta_p^2=.35$, mirroring the findings from Experiment 1. However, the CDAs triggered in response to repeated and novel stimulus configurations in the colour-only task were virtually identical (-0.7µV), $F(1,25)<1$, $p=.723$, $\eta_p^2=.01$.

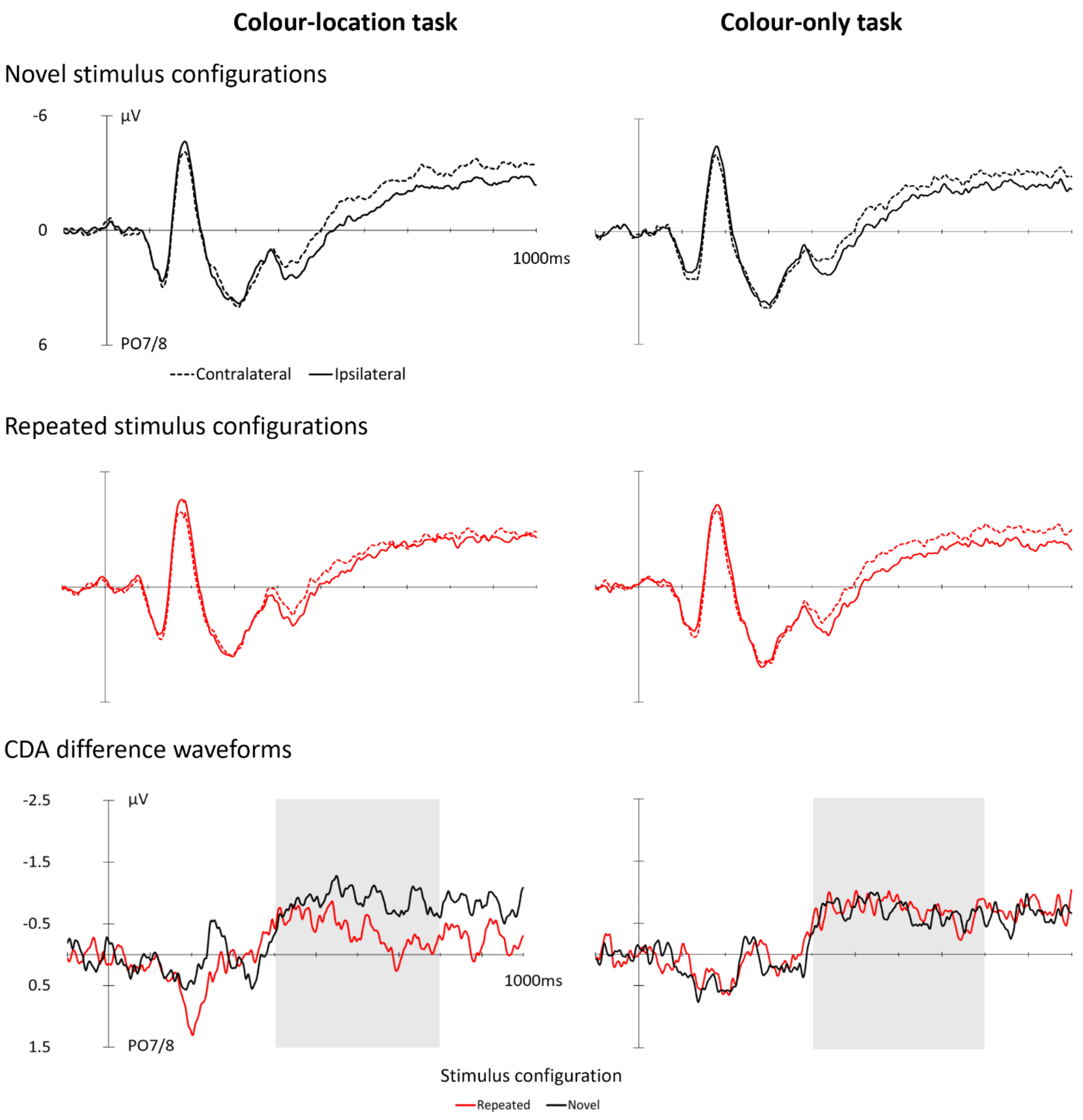


Figure 12 Grand-averaged ERPs measured at lateral posterior electrodes PO7/8 contralateral and ipsilateral to the relevant side of the memory display, separately for novel (top panels) and repeated stimulus configurations (middle panels) in the colour-location (left panels) and colour-only tasks (right panels) of Experiment 2. The corresponding CDA difference waveforms (contralateral minus ipsilateral ERPs) are shown in the bottom panels. Shaded areas mark the CDA time window (400-800ms after memory display onset).

3.3.4 Discussion of Experiment 2

Experiment 2 was designed to isolate the role of spatial information in contextual cueing by comparing a standard visuospatial condition, where both colour and location were repeated (as in Experiment 1), with a non-spatial condition, where only colour information was repeated and locations randomised. The results provide convincing evidence that spatial configuration is a critical component in driving contextual cueing effects within vWM. Performance in the visuospatial condition replicated the findings from Experiment 1 in that repeated displays led to enhanced accuracy and vWM capacity, suggesting participants benefitted from prior exposure to consistent spatial layouts. In contrast, the new condition (non-spatial), where only the colours were repeated, failed to show a contextual cueing effect. These differences suggest that a repeated surface feature like colour, is insufficient to drive contextual cueing in the absence of spatial information. These results are in line with previous findings that altering target locations within repeated displays disrupts contextual learning (Manginelli & Pollmann, 2009; Zellin et al., 2013) and support the view that spatial regularities play a critical role in contextual memory.

At the neural level, like Experiment 1, the reduction in CDA amplitude was observed for repeated displays in the visuospatial condition, indicating that participants were able to offload active vWM storage when spatial information was present. In contrast, no reduction was found in the non-spatial condition, suggesting the repetition of colour alone was insufficient to support contextual cueing at the level of the CDA. Removing spatial consistency eliminated the advantage in repeated displays, resulting in a comparable neural load to that of novel trials. This supports the notion that contextual cueing is primarily driven by spatial information.

The awareness test introduced in this experiment offered insight into the implicit nature of contextual cueing. Participants were more likely to respond “yes” to recognising repeated displays in the visuospatial condition compared to the non-spatial condition, suggesting a greater subjective awareness of repetition when both colour and spatial information were present. However, as the test relied on a simple yes or no response, it remains unclear whether these responses reflect explicit recognition or more general familiarity. Previous research has shown that contextual cueing typically arises from implicit memory, even when participants cannot distinguish repeated from novel displays (Goujon et al., 2015; Vadillo et al., 2020). Thus, while participants responses here do suggest some awareness, it is unclear whether this awareness was explicitly accessible or driven by implicit learning mechanisms. Overall, the results from Experiment 2 clarify the importance of spatial context in contextual cueing within vWM. While repeated exposure to non-spatial features like colour does not appear to produce facilitative effects, a stable spatial layout allows participants to form robust implicit memory traces that can guide attention and reduce cognitive load.

3.4 General discussion

The goal of the present study was to test whether contextual cueing can operate directly within vWM and whether spatial consistency is critical for this effect. Can individuals implicitly use repeated spatial and/or featural stimulus configurations to enhance memory encoding during a change detection task? To answer this question, we conducted two experiments measuring behavioural performance (accuracy and K values) and CDA amplitudes, in a change detection task involving repeated and novel configurations. In Experiment 1, repeated and novel configurations were shown across trials under both low and high memory load. This enabled us to test whether context-related benefits to memory performance can be observed even when many (vs few) items must be maintained (four items in the high load condition). In Experiment 2, we tested whether this effect depends on spatial consistency by comparing a visuospatial condition (context was defined by colour and location) with a non-spatial condition (context was defined by colour only).

Indeed, our data provided evidence for a contextual cueing effect operating directly within vWM, particularly when spatial consistency was maintained. Accuracy, K values and the reduction in CDA amplitude, associated with contextual cueing, was specifically dependant on the repetition of configurations in vWM. These findings contribute to growing evidence that vWM and LTM operate in close interaction rather than in isolation. One possible mechanism is contextual cueing, whereby LTM representations guide attention, thus facilitating more efficient encoding into WM (Cowan, 2001; Oberauer, 2009). In Experiment 1, repeated configurations led to significantly higher accuracy and K values and lower CDA amplitudes compared to novel configurations across both low and high memory load conditions. This suggests that familiar spatial layouts reduce the need for active maintenance within vWM, thus allowing participants to offload onto LTM (Carlisle et al., 2011; Vickery et al., 2010). However, in Experiment 2, this effect was only observed in the visuospatial condition, where both identity and spatial position were repeated. In contrast, when a consistent spatial layout was disrupted in the non-spatial condition, accuracy, K values and CDA amplitudes for repeated displays remained comparable to those of novel displays. This pattern suggests that without a consistent spatial context, repetition does not reduce vWM load. If the identity of the items alone was enough to drive contextual cueing, a CDA should have emerged in both conditions. The absence of such a finding in the non-spatial condition aligns with previous work showing that when spatial layout is disrupted, contextual learning is depleted (Makovski & Jiang, 2010; Manginelli & Pollmann, 2009; Zellin et al., 2013). Moreover, this also suggests that spatial regularities are particularly important for the neural efficiency of repeated contexts. Taken together, both experiments support the notion that contextual cueing effects, specifically in vWM, are not driven by identity repetition alone and that a consistent spatial structure is needed to guide attention to locations during encoding and maintenance (Chun & Jiang, 1998; Jiang & Wagner, 2004).

Despite the clear patterns observed in both experiments, several caveats should be considered when interpreting these findings. Firstly, although the task was designed to investigate implicit learning, the use of repeated configurations across trials may have accidentally introduced conscious awareness of repetition in some participants. Whilst a higher percentage of participants answered “yes” in the visuospatial condition of Experiment 2, which supports the idea that spatial regularities have played a role in contextual cueing, the measure of awareness was vague. It failed to capture how participants recognised the repetition (e.g., visual similarity, prediction, familiarity) or whether they could identify which displays were repeated. A more sensitive recognition test might have provided better insight into the explicit/implicit nature of context learning. For instance, some studies have employed explicit awareness tests by presenting participants with the original repeated configurations from the experiment to assess recognition directly (Wang, 2020).

Secondly, whilst the CDA is a well-established index of vWM load, it traditionally reflects the number of items actively maintained in vWM. However, in the present study, CDA modulation was observed as a function of contextual repetition rather than load. Similar effects showing a decline in CDA with increasing repetition of the same items across trials has also been found in previous studies assessing learning and the transfer to LTM (Carlisle et al., 2011; Grubert et al., 2016; Gunseli et al., 2014a, 2014b; Reinhart & Woodman, 2014). The departure from typical CDA findings here raises the possibility that the amplitude reductions observed here may not just reflect storage but also other mechanisms e.g. increased familiarity or other encoding strategies. Additionally, it is possible that participants may have reached their vWM capacity at 2-3 items, thus increasing the load to 4 would not necessarily elicit further increase in CDA amplitude, consistent with the well-known CDA plateau effect (Vogel & Machizawa, 2004). Finally, the analysis approach used here did not account for how learning developed over time. Segmenting trials into epochs would have allowed us to examine when learning occurred and thus the emerging of contextual cueing effects (Zellin et al., 2011, 2013, 2014; Zinchenko et al., 2018), particularly in the CDA.

As predicted, the results found here are similar to repetition effects observed in visual search. In their studies of contextual cueing, Chun and Jiang (1998) showed that consistent pairing of display configurations and target locations led to significant improvements in search. However, when target locations varied across repetitions, learning was diminished. Similarly, Wolfe et al. (2000) found that repeating the same search displays, even up to 300 times, did not lead to more efficient search if the target location was inconsistent. While these studies tap into attentional guidance during search, the current study examines the influence of contextual learning directly on vWM. In fact, to our knowledge, this is the first study to investigate contextual cueing effects using a change detection paradigm while simultaneously measuring CDA activity. By directly demonstrating that repeated contextual information can modulate the neural correlates of memory load, our approach offers an extension of previous behavioural contextual cueing findings into the electrophysiological domain. Importantly, understanding how contextual regularities support memory encoding has potential applications in educational settings and even cognitive rehabilitation. For example, environments or tasks designed with consistent spatial structures might enhance memory performance in populations with vWM deficits.

By repeatedly presenting the same visual displays in a change detection task, we have clarified the role of learning in enhancing vWM. We suggest that although it is very difficult to increase the overall capacity of vWM through learning, learning can shift attentional priorities, thereby improving vWM performance. Importantly, this study demonstrated that contextual learning is not limited to specific tasks like visual search, it also emerges in change detection tasks, despite this task relying on different spatial attention demands (Rensink, 2002; Wolfe, 1998). The presence of contextual cueing in this task suggests that the visual system supports a spatial context learning mechanism that generalises across tasks with differing attentional requirements. In sum, contextual cueing demonstrates the powerful ability of humans to extract meaningful regularities from complex, noisy environments and to utilise this knowledge to influence fundamental cognitive processes. The present findings provide strong evidence that there is no learning of ‘what’ without ‘where’, and, at least for contextual cueing, both are required.

Chapter 4: Individual working memory capacity predicts search performance in multiple colour search

4.1 Introduction

Visual search, the process of locating a target among distractors, is a fundamental component of human cognition. Central to efficient search performance is the use of attentional templates, which are internal representations of target features actively maintained in vWM (Duncan & Humphreys, 1992; Olivers et al., 2011). These templates serve as top-down signals that bias visual processing, enabling prioritisation of stimuli that match goal-relevant features. They are therefore activated in a preparatory fashion, prior to the appearance of a search display (Chelazzi et al., 1998; Grubert & Eimer, 2018). When activated, these attentional templates enhance the salience of matching items and guide attention towards them in an automatic and spatially global manner (Berggren et al., 2017; Desimone & Duncan, 1995; Eimer, 2014; Grubert & Eimer, 2016a, 2016b; Martinez-Trujillo & Treue, 2004; Wolfe, 2021).

A central question in the attentional literature concerns the number of templates that can be maintained simultaneously during visual search. While early models assume that only a single attentional template can be active at one time (Houtkamp & Roelfsema, 2009; Olivers et al., 2011), more recent evidence suggests that multiple attentional templates can be held and used in parallel allowing individuals to search for more than one target feature simultaneously (Barrett & Zobay, 2014; Berggren & Eimer, 2019, 2020; Christie et al., 2015; Grubert & Eimer, 2016a, 2016b; Irons et al., 2012; Kerzel & Grubert, 2022; Moore & Weissman, 2010; Ort et al., 2019). For example, Irons et al. (2012) showed that when individuals searched for one of two equally likely target colours, spatially uninformative cues matching either target colour, but not distractor colours, captured attention and facilitated faster responses when targets appeared at the cued location. This suggests that search templates for both target colours were activated in parallel.

This ability to engage multiple templates in parallel is thought to be constrained by the capacity of vWM, which is typically limited to around 3-4 items (Cowan, 2001). However, vWM capacity varies across individuals (Awh et al., 2007; Cowan et al., 2005; Fukuda et al., 2010b; Luck & Vogel, 2013) and individuals with higher vWM capacity are generally better at maintaining multiple items, filtering out distractions and flexibly updating representations. For instance, Fukuda and Vogel (2009) demonstrated that although low- and high-capacity individuals could attend to a target equally well, low-capacity individuals were more susceptible to involuntary attentional capture by irrelevant distractors. In a later study, Fukuda and Vogel (2011) found that high-capacity individuals could recover attentional control quicker after capture than low-capacity individuals. Similarly, Vogel et al. (2005) showed that low-capacity individuals store more distractors in vWM resulting in reduced efficiency. Together, these findings indicate that vWM capacity may not only reflect storage quantity, but also the efficiency of attentional selection, particularly during the maintenance of multiple templates. In support of this, Grubert et al. (2016) demonstrated that both working memory (WM), and long-term memory (LTM) can support template based visual search, though their study did not directly examine how these mechanisms vary across individuals.

Given these observations, the critical question is whether individual differences in vWM capacity predict the efficiency of visual search, particularly when multiple target templates must be maintained simultaneously. While vWM capacity has been linked to goal maintenance (Kane & Engle, 2003) and attentional control, including suppression of irrelevant and prioritisation of relevant items (Fukuda & Vogel, 2011; Vogel et al., 2005), few studies have directly examined how individual capacity limitations affect the ability to deploy multiple attentional templates in parallel. The question is critical because real world search often involves searching for several possible targets at once, i.e. looking for multiple ingredients down one aisle in the supermarket. Behavioural evidence suggests that increasing the number of templates leads to a decline in search accuracy and reaction times (Beck et al., 2012; Grubert & Eimer, 2016a, 2016b), but little is known about if and how these performance costs vary across individuals as a function of their vWM capacity.

Some studies have reported inconsistent links between vWM capacity and visual search performance, particularly in the light of individual differences. Kane et al. (2006) found that while high-span individuals were more accurate overall, WM capacity did not influence search slopes or speed. Similarly, Poole and Kane (2009) observed benefits for high-span individuals only under specific conditions, such as when sufficient preparation time was provided and distractors were present. These findings suggest that the link between vWM and search may only emerge when tasks involve the simultaneous maintenance and selection of multiple templates, thus placing heavier demands on vWM and attentional control.

Specifically, it remains unclear whether the capacity to store items in vWM is directly correlated with the capacity to activate and use multiple attentional templates and whether this is reflected in visual search performance. To assess these relationships more precisely, neurophysiological markers, particularly event-related potentials (ERPs), offer valuable insight into the deployment of attention in real time. Two ERPs are particularly relevant in this context: the contralateral delay activity (CDA) and the N2pc. The CDA, a sustained negativity over posterior electrodes contralateral to remembered items, refers to the number of items actively maintained in vWM (Vogel & Machizawa, 2004). The CDA amplitude increases with memory load and plateaus at an individual's capacity limit, making it a useful index of storage.

Feldmann-Wüstefeld (2021) demonstrated that CDA amplitudes correlate with individual differences in vWM capacity as measured by K. They showed that larger CDA amplitudes were associated with higher behavioural capacity estimates, confirming the CDA as a reliable neural index of vWM load. Extending this work to attention, Luria and Vogel (2011) examined correlations between CDA amplitudes and individual differences in search performance. They found that as search difficulty increased, CDA amplitudes also increased. Importantly, they reported that individual differences in vWM capacity correlated with search efficiency across all difficulty levels, suggesting that higher-capacity individuals performed better overall, but relied less on active memory storage (as shown by reduced CDA amplitudes). Further support comes from Carlisle et al. (2011), who found that CDA amplitude was significantly correlated with search accuracy and response times. As participants gained experience with the same target over repeated trials, both CDA amplitude and its correlation with performance declined, suggesting a shift from WM to LTM. These studies highlight that both task demands, and individual capacity independently modulate the engagement of vWM during search.

In the domain of visual search, the N2pc component, a lateralised ERP associated with the deployment of visual attention, reflects the selection of target-matching items based on active attentional templates (Eimer, 1996; Luck & Hillyard, 1994). Importantly, individual differences in N2pc amplitudes have been linked to search efficiency and vWM capacity (Gaspar et al., 2016), suggesting that the strength and precision of attentional selection are directly related to internal vWM resources and again may reflect how many templates can be concurrently activated and deployed during visual search.

To directly assess these preparatory template activation processes, Grubert and Eimer (2018) developed the rapid serial probe presentation (RSPP) paradigm, a method that tracks the real-time activation of attentional templates during search preparation using EEG. In this method, participants search for colour defined targets while irrelevant probe displays are presented rapidly (every 200ms) during the intervals between search displays. Each probe contains a lateralised colour singleton that either matches a currently relevant target or a distractor colour. If a target colour matching probe elicits an N2pc component, this indicates that the corresponding attentional template was active and attracting attention at that time, even though the probe display was irrelevant to the task.

Using the RSPP paradigm, Grubert and Eimer (2018) found that in single template search, probe displays matching the target colour elicited N2pc components, while distractor matching probes did not, providing clear evidence that target templates were transiently activated during the preparation phase between search trials. In a follow up study, the RSPP method was used to examine multiple template activation. In a two-colour search where the relevant target colour alternated between trials, probes matching the upcoming and previous target colours elicited reliable N2pc components (Grubert & Eimer, 2020), suggesting that both colour templates were active, supporting a parallel template activation account.

In a further study, Grubert and Eimer (2023) used the RSPP paradigm in a three-template search task, where participants searched for one of three colour defined targets.

Probes matching all three target colours elicited reliable N2pc components, suggesting all three templates were active in parallel during search preparation. Interestingly, these probe-evoked N2pcs did not decline in amplitude with increasing template load, suggesting that increased template load did not necessarily weaken template activation at the group level. This raises the question of whether the capacity to activate multiple templates is uniform or varies across individuals. Grubert and Eimer speculated that maintaining three templates may approach or exceed vWM for some individuals. However, because their study did not include direct measures of individual vWM capacity, it remains unclear whether the ability to activate multiple attentional templates varies systematically across individuals. Furthermore, previous studies examining correlations between vWM and attention have typically done so within a single task context, limiting the ability to assess how capacity generalises across different domains. Consequently, the degree to which template activation scales with vWM capacity, either behaviourally or neurally, remains an open question.

The current study aimed to address this gap in the literature by investigating whether individual vWM capacity predicts search performance in a visual search task, particularly when multiple attentional templates are required. To do this, we used two separate paradigms, combining behavioural and electrophysiological measures of both vWM capacity (K; CDA) measured during a change detection task, and search efficiency (reaction times, error rates; N2pc) measured during a search task. In the change detection task, participants were presented with one, two, or three colours, which determined the set size to be remembered. In the visual search task, participants were also shown up to three colours and were required to retain all of them, as the target colour could be any of those presented. To capture the real time activation of target templates during search preparation, the RSPP paradigm was embedded into the search task, allowing for direct measurement of whether and when specific templates are active. If vWM capacity constrains the number of attentional templates that can be simultaneously activated during search, then we would expect individuals with lower vWM capacity to show reduced efficiency and weaker probe N2pcs as the number of templates increases. In contrast, individuals with higher vWM capacity may maintain efficiency up to two templates but show costs between two- and three-colour conditions, reflecting the upper limit of their template capacity. Alternatively, those with a vWM capacity exceeding three items may show no such cost.

4.2 Methods

4.2.1 Participants

Forty-two participants took part and either received £10/hour or course credits as compensation for their time. The experiment was approved by the Ethics Committee of the Psychology Department at Durham University and was conducted in accordance with the Declaration of Helsinki. Participants gave informed written consent prior to testing. Seven participants were excluded from analysis due to excessive eye movement artifacts (>40% of trials lost during artifact rejection; a priori criterion). The remaining 35 participants were between 18 and 30 years of age (M_{age} = 22.3, SD_{age} = 2.7). Thirty participants were female and five were male. Three participants were ambidextrous, and the remaining participants were all right-handed. All participants had normal or corrected-to-normal vision and no colour deficiency (as tested with the Ishihara colour vision test; Ishihara, 1972). The sample size of 35 was calculated by means of an a priori power analysis using MorePower 6.0.1 (Campbell & Thompson, 2012). To detect a significant correlation with an assumed alpha of .05, power of .90, and the effect sizes reported by Zhong et al. (2024; r = .51, between the N2pc amplitudes and K value), it was suggested that a minimum of 34 participants were required. We therefore consider this sample size (N=35) to be sufficient to reliably detect the planned correlations.

4.2.2 Stimuli and procedures

Participants were sat in a dimly lit sound attenuated Faraday cage with an approximately 90cm viewing distance from the monitor. Stimuli were presented on a 22-inch MSI Optix G272 LCD monitor with a 100-Hz refresh rate and a resolution of 1920×1080 pixels. Stimulus presentation, timing, and response collection were controlled by PsychoPy (psychophysics software in Python; Peirce et al., 2019) on an LG Pentium PC running under Windows 10. All stimuli were presented on a black background. A central grey fixation point was presented throughout the experimental blocks (CIE x, y colour coordinates: .327/.350; luminance~21.2 cd/m^2, 0.2° x 0.2° of visual angle).

The experiment was tested in two sessions on non-consecutive days: the first session was always the change detection task, and the second session was the visual search task. Before each session, participants practised the respective task until they reached 75% accuracy in the low-load condition. The practice data were not recorded.

4.2.2.1 Change detection task

Each trial started with the presentation of an indicator display for 200ms, which was followed by a 200ms memory display. After a jittered retention interval of 800-900ms, the test display was presented for 200ms (Figure 13A). The inter-trial interval (between the offset of the test display in trial n-1 and the onset of the indicator display in trial n) was 2000ms. Indicator displays contained two grey arrowheads presented at 0.6° under and above the central fixation on the vertical meridian, pointing towards the hemifield to be memorised. There were three equiprobable load conditions presented in an intermixed fashion. Both memory and test displays showed coloured squares in both hemifields at 1° eccentricity (Figure 13B). The locations depended on load but were determined randomly in each trial. For the one-colour condition, each hemifield contained a single square on the horizontal meridian. In the two-colour condition, each hemifield contained two differently coloured squares at the 2, 4, 8, and 10 o'clock positions, and for the three-colour condition, each hemifield contained three differently coloured squares at the 1, 3, 5, 7, 9, and 11 o'clock positions. Squares subtended 0.5 x 0.5 degrees of visual angle.

Participants task was to memorise the colours of the squares on the left or right side (as shown by the indicator) of the memory display and compare them to the colour set shown on the same side of the test display. The relevant memory side was completely randomised between trials. In no-change trials, the colours of the test display were identical to the colours in the memory display. In change trials, one of the squares in the test display had a new colour which was not previously present in the memory display. Participants were instructed to press the up or down arrow key on a standard keyboard if they detected a colour change or felt that the colours in the test display were the same as in the memory display, respectively. The response-to-key mapping (change/no-change response on arrow up/down key) and the hand-to-key mapping (left/right hand on arrow up/down key) was counterbalanced across participants but was kept constant for each participant for the duration of the whole experiment. The colours for the one, two or three squares on the relevant memory side were randomly chosen from a set of six possible memory colours: red (.670/.318), green (.268/.652), blue (.166/.153), yellow (.414/.506), pink (.362/.212) and cyan (.205/.297). The colours for the squares on the irrelevant memory side were then determined from the remaining memory colours plus brown (.494/.378) and grey

(.328/.346), which only ever served as colours for squares on the irrelevant side of the memory display. All colours were equiluminant (~9.0 cd/m^2).

Each experimental block condition contained 36 trials, i.e., six trials for each combination of load (1, 2 and 3 colours) and trial type (change, and no-change trial). Twenty blocks were tested for a total of 720 trials in the change detection task.

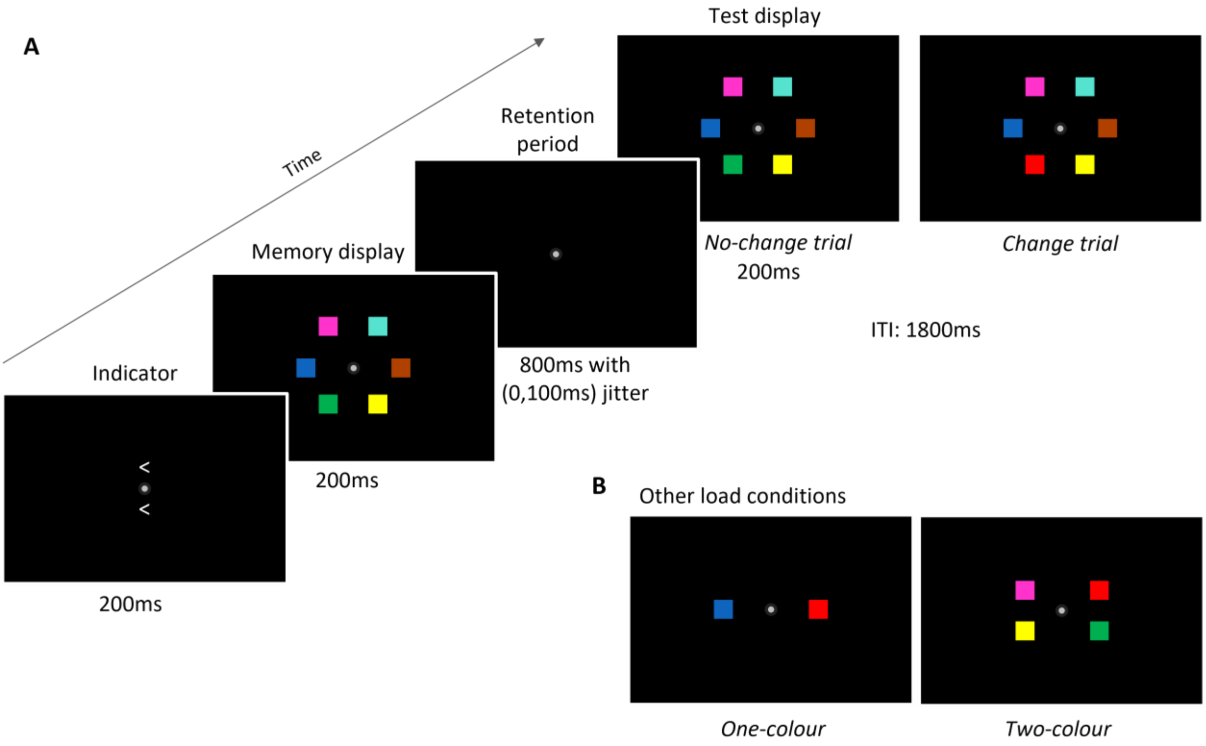


Figure 13 Schematic illustration of the stimuli and temporal trial sequence in the change detection task. Panel A: Each trial began with an indicator display containing two arrowheads pointing either left or right, indicating the hemifield that was task-relevant for that trial. This was followed by a memory display, a retention period, and then a test display. Participants were instructed to remember the colours of the squares presented in the relevant hemifield (left or right) during the memory display, and to report whether any square in the same hemifield changed colour in the test display. Panel B: Example memory displays for one-colour and two-colour load conditions.

4.2.2.2 Visual search task

Each trial started with the presentation of a stimulus display for 50ms, followed by a 150ms blank interval (200ms stimulus onset asynchrony). The first seven stimulus displays in each trial contained a probe display (probes 1 to 7 in Figure 14A), the eighth display contained both a probe and search display, which was only presented for visual consistency.

Search displays were presented at an eccentricity of 1.4° from central fixation and contained six vertically (0.2° × 0.6°) or horizontally (0.6° × 0.2°) oriented bars at the 1, 3, 5, 7, 9, and 11 o'clock positions of an imaginary clock face. The orientations of the six bars were selected independently and randomly for each search display. Each display contained six bars in different colours (selected from the same colours used in the change detection task), exactly one of which matched the target colour; the remaining five were non-target colours. Target colours (either one, two or three colours) were chosen from one of two possible sets: primary colours (red, green, blue) or mixed colours (pink, yellow, cyan). Non-target colours were drawn from the colour set that did not contain the target colours as well as grey and brown (never targets). In each trial, only one of the three target colours was response relevant. The assignment of the six colours to the six bars was completely random as well as the target bar location, with the exception that it appeared equiprobably on either side of the display.

Participants task was to report the orientation (vertical/horizontal) of the target colour bar in each trial by pressing the up/down arrow keys on a standard keyboard. Participants were presented with a display indicating the one, two or three target colours for the upcoming block. The response-to-key mapping (vertical/horizontal response on arrow up/down key) and the hand-to-key mapping (left/right hand on arrow up/down key) was counterbalanced across participants but was kept constant for each participant for the duration of the whole experiment.

Probe displays contained six items composed of four closely aligned dots, two on the vertical, and two on the horizontal axis (0.1° × 0.1° for each dot, 0.25° × 0.25° for each four-dot probe item). The probe items were presented at the same positions as the search bars (1, 3, 5, 7, 9, and 11 o'clock) but closer to fixation (at an eccentricity of 0.5°). Five of the six probe items were uniformly grey; the sixth item was either one of the target colours or a random non-target colour. In the two- and three-colour conditions, target-colour probes were equally likely to match any of that blocks assigned target colours. The probe singleton locations were selected randomly and independently in each probe display, with the following two restrictions: successive singleton probes were equally likely to appear on the same or opposite display sides, but immediate repetitions of the exact probe location (on the imaginary clock face) were not allowed. As a result, each probe display was equally likely to be preceded and followed by probe displays that contained a colour singleton on the same or the opposite side. This was done to ensure that lateralised responses to any particular probe singleton would remain unaffected by any lateralised response triggered by singletons in temporally adjacent probe displays. Participants were informed that probe displays were task-irrelevant and could be ignored.

There were 120 blocks of 12 trials. Each trial consisted of eight stimulus displays presented in a rapid serial presentation, shown in Figure 14A. The search load (one, two or three colours) was determined randomly but equiprobably at the start of each block (40 blocks for each search load). Blocks were kept as short as possible and participants were instructed not to blink during the blocks, if possible. The twelfth search display in each block was followed by seven additional probe displays to keep stimulus conditions during the post-target response interval identical across all 12 trials in a block. Thus, each block contained 12 search displays and 91 probe displays (7 probes preceding each of the 12 searches, plus 7 additional probes following the final search).

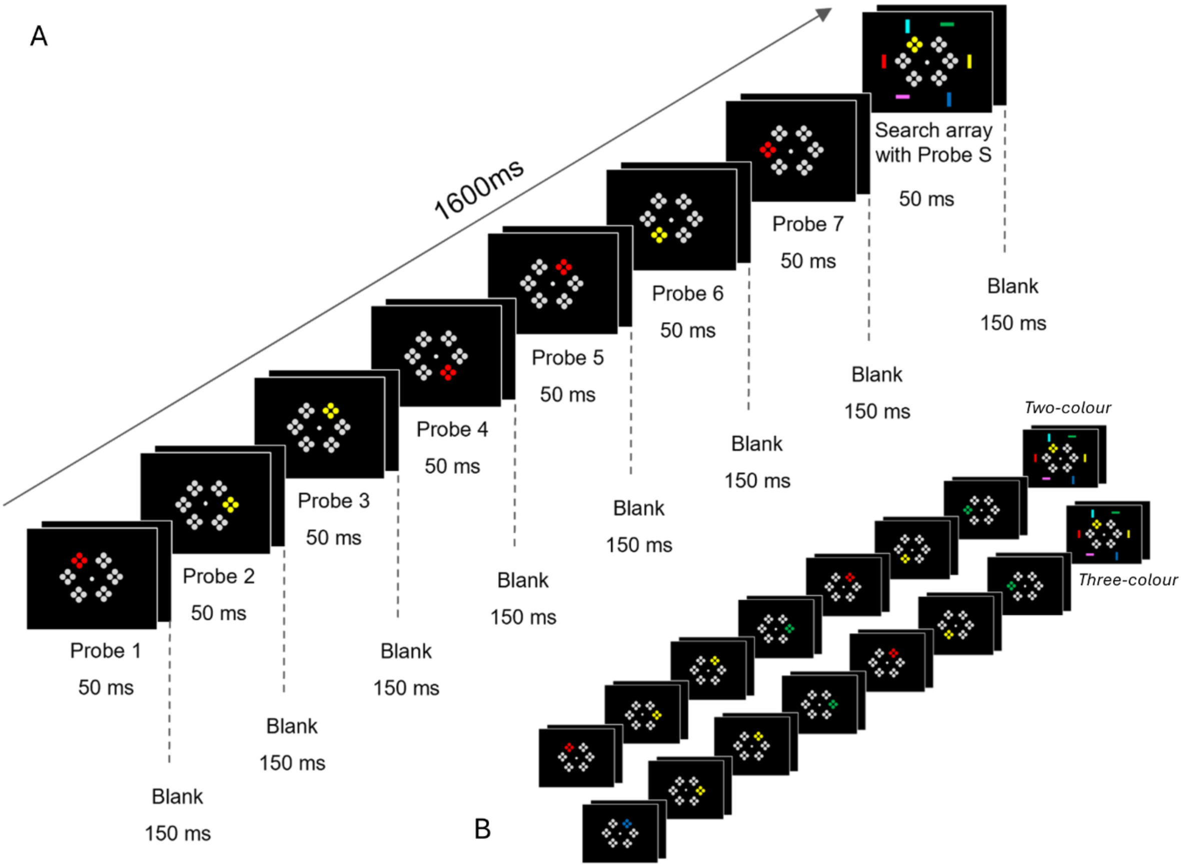


Figure 14 Schematic illustration of the stimuli and presentation times in the visual search task (Panel A). Search displays contained six differently oriented (horizontal or vertical) bars. One bar matched the target colour (e.g., red); the other five were non-target colours. Probe displays were presented every 200ms in the intervals between search displays (probes 1-7) and simultaneously with each search display. Each probe display contained a colour singleton that either matched the target or a non-target colour. Panel B shows the trial sequence for the two and three colour search conditions. The target colour for each trial was randomly selected from the two or three cued colours indicated at the beginning of each block.

4.2.3 Data analyses and EEG recordings

EEG was DC-recorded from 23 scalp electrodes (EasyCap, Brain Products) at standard positions of the extended 10/20 system, and two HEOG electrodes placed at the outer canthi of the eyes. EEG data were recorded using the Brain Vision Recorder (BrainAmp DC amplifier, Brain Products GmbH, Gilching, Germany) at a sampling rate of 500Hz with a 40Hz low-pass filter. No other filters were applied after data acquisition. Impedances were kept below 5kΩ. The left earlobe served as online reference during data acquisition, but all channels were re-referenced offline to linked earlobes. All EEG preprocessing was conducted with the BrainVision Analyzer software (Brain Products GmbH, Gilching, Germany). Data from trials with anticipatory (<200ms) or slow responses (>1500ms), were excluded from the analysis. Trials with eye movements (exceeding ±30µV in the bipolar HEOG channel), blinks (exceeding ±60µV at Fpz), and muscular movements (exceeding ±80µV in all other channels) were also excluded.

4.2.3.1 Change detection task

For the change detection task, error rates were used to calculate K values separately for one, two and three colour conditions. Following the suggestions of Rouder et al. (2011), K was calculated according to Pashler's equation (Pashler, 1988):

$$K = N\left(\frac{h-f}{1-f}\right)$$

where N is the number of memory items (set size), and *h* and *f* are the measured hit and false alarm rates, respectively. Hit rates were calculated as

$$h = \frac{hits}{hits + misses}$$

and false alarm rates were calculated as

$$f = \frac{false\ alarms}{false\ alarms + correct\ rejections}$$

The EEG was segmented into 1100ms time windows including a 100ms pre-stimulus baseline and a 1000ms ERP time window following the onset of the memory displays. Data from incorrect responses were excluded from ERP analyses. Artifact rejection resulted in an exclusion of 16.2% of all trials (SD=12.2%, ranging between 1.5% and 33.8% across participants). The remaining segments were averaged separately for left- and right-hemifield memory squares in trials with one, two and three colours. CDA mean amplitudes were quantified at lateral posterior electrodes PO7 and PO8, contralateral and ipsilateral to the memory side, in the 350-950ms interval after memory display onset.

4.2.3.2 Visual search task

For the visual search task, EEG epochs (500ms) were locked to the onsets of the probes (probes 1 to 7) and the search displays and included a 100ms pre-stimulus baseline and a 400ms post-stimulus ERP time window. Data from the first and last seven probe displays in each block were excluded from analysis as well as incorrect or missing responses. Artifact rejection resulted in an exclusion of 10.2% of all epochs (SD=8.6%, ranging between 0.7% and 29.3% across participants). The remaining epochs were averaged separately for each probe number (probes 1-7), search conditions (one, two and three colours) and probe hemifield (left vs right). Separate averages were also computed for search displays with a target in the left or right hemifield for each of the search conditions. Only correct trials were included in the target N2pc analysis. N2pc components to probes were quantified based on ERP mean amplitudes obtained at lateral posterior electrodes PO7 and PO8, contralateral and ipsilateral to the side of a probe, in the 210-290ms interval after each respective probe display onset. Target N2pcs in the search displays were computed with a 230-330ms interval after each search display onset. The time windows were determined objectively by the collapsed-localiser method (Luck & Gaspelin, 2017), with ±40ms around the peak amplitude within a 150-350ms time window in a localiser waveform averaged using difference waveforms over PO7/8 across all conditions and all participants.

Effect sizes are reported as Cohen's d (Cohen, 1988) with a CI of 95% for *t*-tests, and as partial eta squared (η_p^2) for *F*-tests. When necessary, Greenhouse-Geisser corrections were applied to *F*-tests, and Bonferroni corrections to *t*-tests. All *t*-tests were two-tailed. Correlations were assessed using Pearson's *r*; when the assumption of normality was violated (Shapiro-Wilk test, $p<.05$), a non-parametric approach (Kendall's tau) was applied. All statistical analyses were conducted with JASP statistical software (version 0.95.0.0).

4.3 Results

4.3.1 Working memory – Change detection task

4.3.1.1 Behavioural results

4.3.1.1.1 Mean K values

vWM capacity (K) values were entered into a repeated-measures ANOVA with the within-subjects factor load (1, 2 and 3 colours). There was a significant main effect of load, $F(2,68)=387.0$, $p<.001$, $\eta_p^2=.92$, reflecting an increase in capacity across levels of load. Follow-up *t*-tests showed a significant increase in capacity (K) from load 1 to load 2 (.95 vs 1.8), $t(34)=40.1$, $p<.001$, $d=6.78$, and from load 2 to load 3 (1.8 vs 2.3), $t(34)=10.0$, $p<.001$, $d=1.68$. See Figure 15 for means.

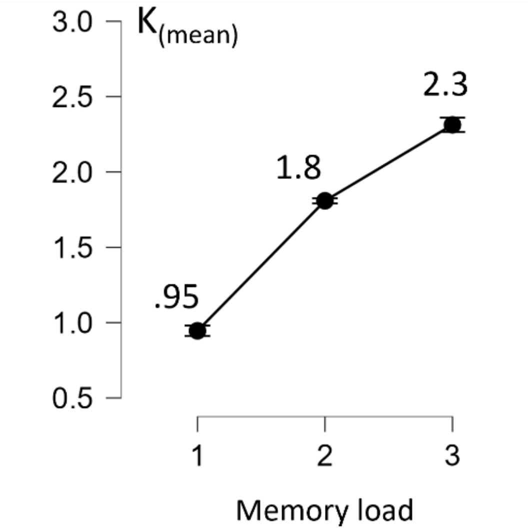


Figure 15 Mean vWM capacity (K) as a function of load (1, 2 and 3 colours). Error bars represent ± 1 SEM.

4.3.1.1.2 Mean K values in High- and Low-K groups

Participants were divided via a median split at K=1.7 into High-K (N=20) and Low-K (N=15) groups. A 2 (K group: High vs Low) x 3 (Load: 1, 2 and 3 colours) mixed-design ANOVA was conducted on behavioural K values. Significant main effects of load, $F(2,66)=910.3$, $p<.001$, $\eta_p^2=.97$ and K group, $F(1,33)=53.0$, $p<.001$, $\eta_p^2=.62$ were found. As well as a load × K group interaction, $F(2,66)=53.9$, $p<.001$, $\eta_p^2=.62$. In both groups, K values increased with load. Performance was highest at load 3 (High-K: $M=2.6$, $SD=.18$; Low-K: $M=1.9$, $SD=.33$), followed by load 2 (High-K: $M=1.9$, $SD=.07$; Low-K: $M=1.7$, $SD=.16$) and lowest at load 1 (High-K: $M=1.0$; Low-K: $M=.9$). For both K groups, paired samples *t*-tests confirmed significant K increases from load 1 to load 2 (all $t>21.7$, all $p<.001$) and load 2 to load 3 (all $t>3.9$, all $p<.002$). The High-K group showed significantly higher K values at load 1, 2 and 3 ($M=1.0$, 1.9, 2.6) than the Low-K group ($M=.9$, 1.7, 1.9); all $t>3.1$, all $p<.004$. See Figure 16A for mean K values and 16B and 16C for the distribution of K by group.

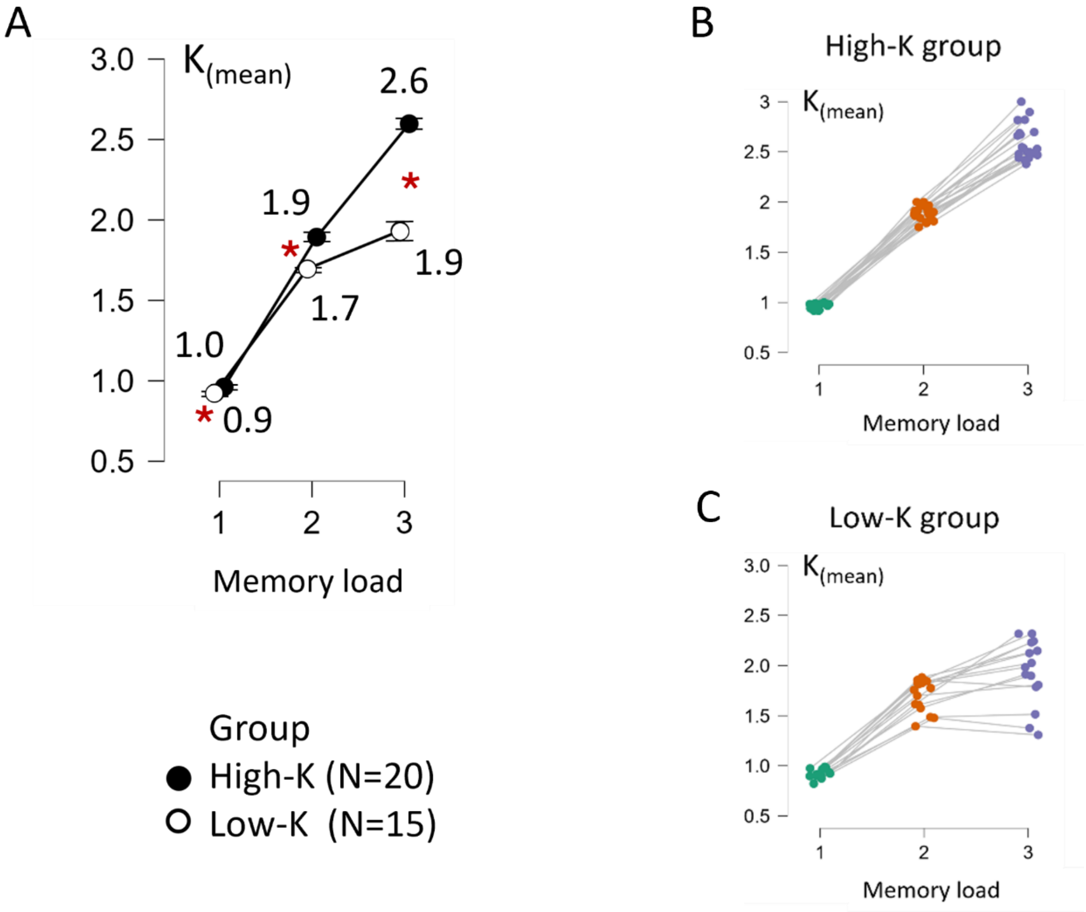


Figure 16 Mean vWM capacity (K) as a function of load (1, 2, and 3 colours), shown separately for High-K (filled circles) and Low-K (open circles) groups. Panel A displays group means with red asterisks indicating significant differences between groups at each load level. Error bars represent ±1 SEM. Panel B shows individual participant data for the High-K group, and panel C shows individual data for the Low-K group. Lines connect individual K values across loads, demonstrating the within-subject increase in K.

4.3.1.2 ERP Results

4.3.1.2.1 Mean CDA amplitudes

CDA amplitudes were submitted to a repeated-measures ANOVA with load (1, 2 and 3 colours) as a within-subjects factor. There was a significant main effect of load, $F(2,68)=30.7$, $p<.001$, $\eta_p^2=.48$. Descriptive statistics showed that mean CDA amplitudes ($\mathrm{CDA}_{(mean)}$) became larger with increasing load: load 1 ($M=-.01$, $SD=.70$), load 2 ($M=-.68$, $SD=.73$), and load 3 ($M=-.98$, $SD=.75$). Paired samples *t*-tests confirmed significant increases in CDA negativity from load 1 to load 2, *t*=5.6, *p*<.001, and from load 2 to load 3, *t*=2.9, *p*=.006. Paired-samples *t*-tests were conducted comparing activity at PO7 and PO8 electrodes for each load. No significant difference was found at load 1, *p*=.912, but significant lateralisation emerged at load 2, *t*=5.5, *p*<.001, and load 3, *t*=7.7, *p*<.001. See Figure 17A for difference waves and 17B for means.

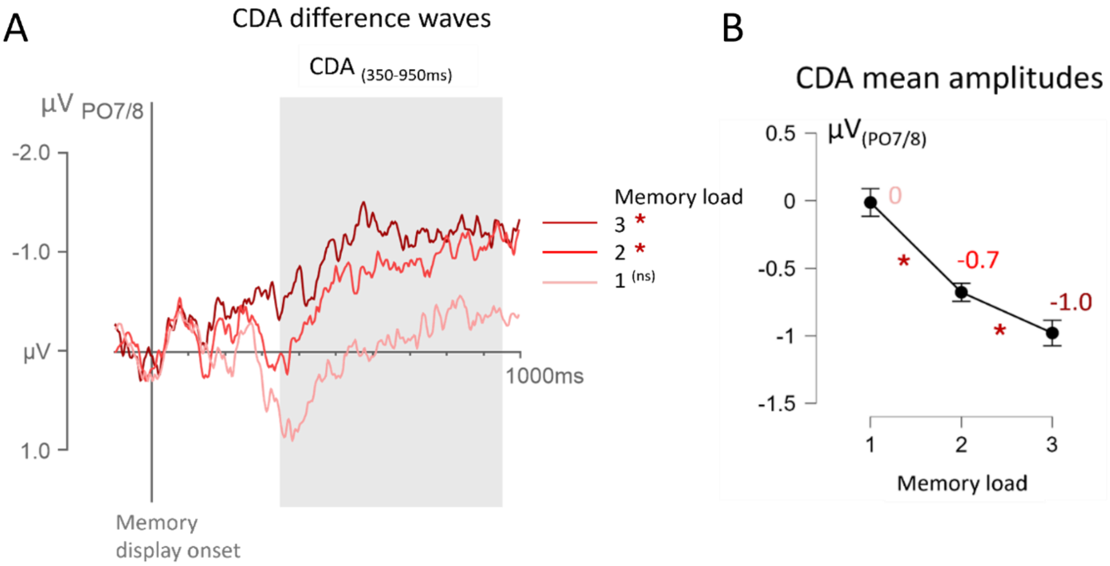


Figure 17 Contralateral delay activity (CDA) as a function of load. Panel A shows CDA difference waveforms at PO7/8 for loads of 1, 2, and 3 colours, with increased negativity corresponding to higher loads. The shaded region indicates the CDA time window (350–950ms post-stimulus). Red asterisks denote significant CDA's. Panel B shows mean CDA amplitudes as a function of load, with values shown above each point and red asterisks indicating statistically significant differences between adjacent loads. Error bars represent ±1 SEM.

4.3.1.2.2 Correlations between mean behavioural K and mean CDA amplitudes

To examine the relationship between behavioural performance (K values) and neural activity (CDA amplitudes), Pearson correlations were calculated at each load level. There was no significant correlation found between the mean value of K ($K_{(mean)}$) and the mean CDA ($CDA_{(mean)}$) across three loads, *r*(33)=–.24, *p*=.166 (Figure 18A). There was also no significant correlation found between $K_{(load1)}$ and $CDA_{(load1)}$, *r*(33)=.27, *p*=.121. The scatterplot (Figure 18B) suggests a ceiling effect as many participants clustered near the maximum possible K value of 1. A significant negative Pearson correlation was found between $K_{(load2)}$ and $CDA_{(load2)}$, *r*(33)=–.37, *p*=.029 (Figure 18C). A significant negative Pearson correlation also emerged between $K_{(load3)}$ and $CDA_{(load3)}$, *r*(33)=–.35, *p*=.037 (Figure 18D).

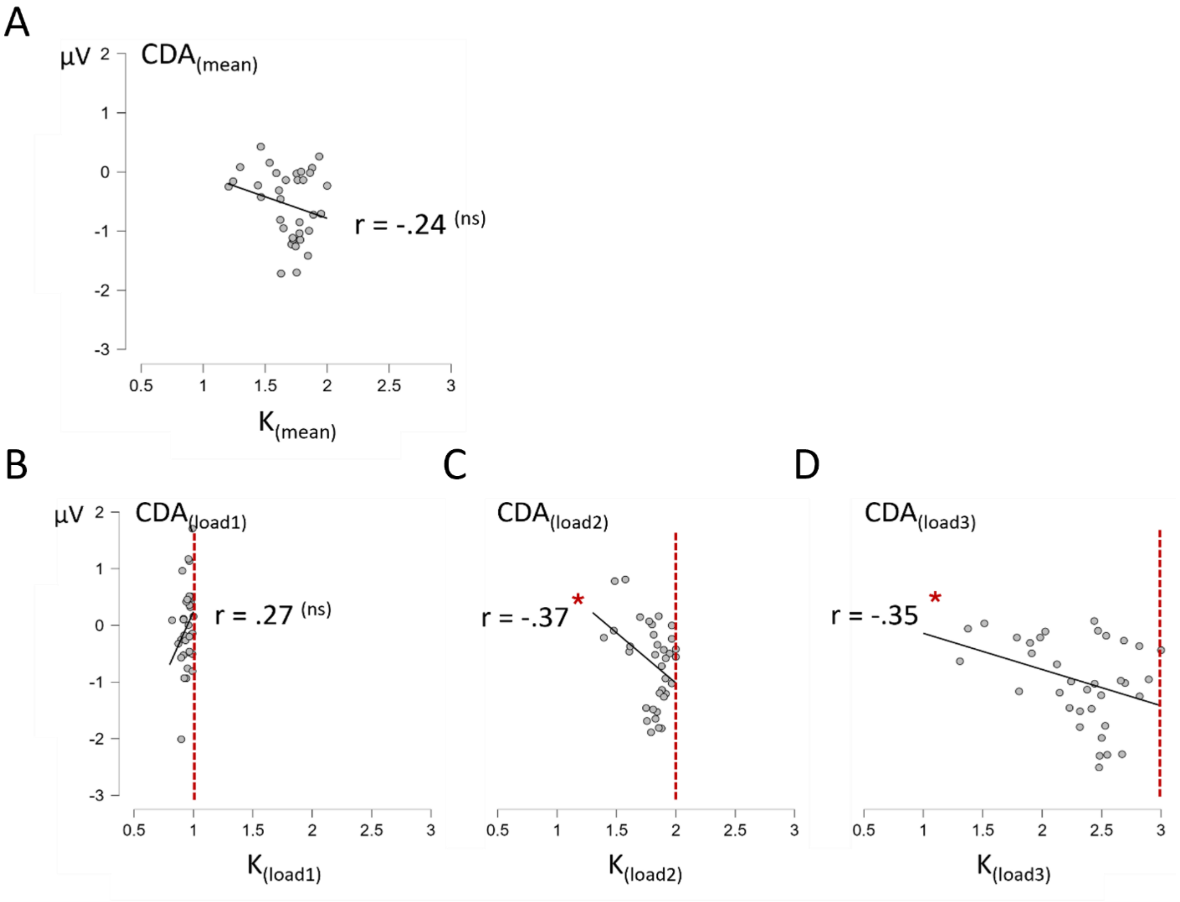


Figure 18 Scatterplots showing the relationship between contralateral delay activity (CDA) and vWM capacity (K values) across mean values (Panel A) and individual loads (Panels B, C, D). Each dot represents one participant. Panel A shows the correlation between mean CDA amplitude and mean K across all loads. Bottom panels show correlations for load 1 (B), load 2 (C), and load 3 (D), with red dashed lines denoting ceiling values for K at each load level. Pearson correlation coefficients (*r*) are displayed in each plot. Asterisks indicate significant correlations (*p*<.05).

4.3.1.2.3 Mean CDA amplitudes in High- and Low-K groups

A 2 (K group: High vs Low) x 3 (Load: 1, 2 and 3 colours) mixed-design ANOVA was conducted on CDA amplitudes. Significant main effects of load, $F(2,66)=32.8$, $p<.001$, $\eta_p^2=.50$, and a significant load × K group interaction, $F(2,66)=9.0$, $p<.001$, $\eta_p^2=.22$, were observed. The main effect of K group was not significant, $F(1,33)=1.9$, $p=.179$, $\eta_p^2=.05$. In both groups CDA amplitudes became more negative with increasing load. CDA amplitude was largest at load 3 (High-K: $M=-1.17$, $SD=.81$; Low-K: $M=-.72$, $SD=.59$), followed by load 2 (High-K: $M=-.95$, $SD=.60$; Low-K: $M=-.31$, $SD=.73$) and weakest at load 1 (High-K: $M=.11$; Low-K: $M=-.17$). In the High-K group, paired samples *t*-tests confirmed significant CDA amplitude increases from load 1 to load 2, $t(19)=9.5$, $p<.001$, indicating a strong neural response to increased vWM demand. However, the change from load 2 to load 3 was not statistically significant, $t(19)=1.6$, $p=.129$. In contrast, Low-K individuals did not show a significant change in CDA amplitude from load 1 to load 2, $t(14)=.9$, $p=.371$, but did exhibit a significant increase from load 2 to load 3, $t(14)=2.6$, $p=.021$, suggesting a delayed neural response to increasing memory demands with CDA enhancement emerging only at the highest load. See Figure 19B for CDA differences/load costs. Significant lateralisation emerged at all loads for the Low-K group (all $t>2.2$, all $p<.042$). However, significant lateralisation was only observed at load 2 and load 3 for the High-K group (all $t>.6$, all $p<.001$) and not at load 1 ($p=.508$). The High-K group showed significantly larger CDA amplitudes than the Low-K group at load 2, $t(33)=2.8$, $p=.008$, and load 3 (marginal significance), $t(33)=1.8$, $p=.076$. No significant group difference was found at load 1 ($p=.247$). See Figure 19A for CDA mean amplitudes and 19C for difference waves for each group.

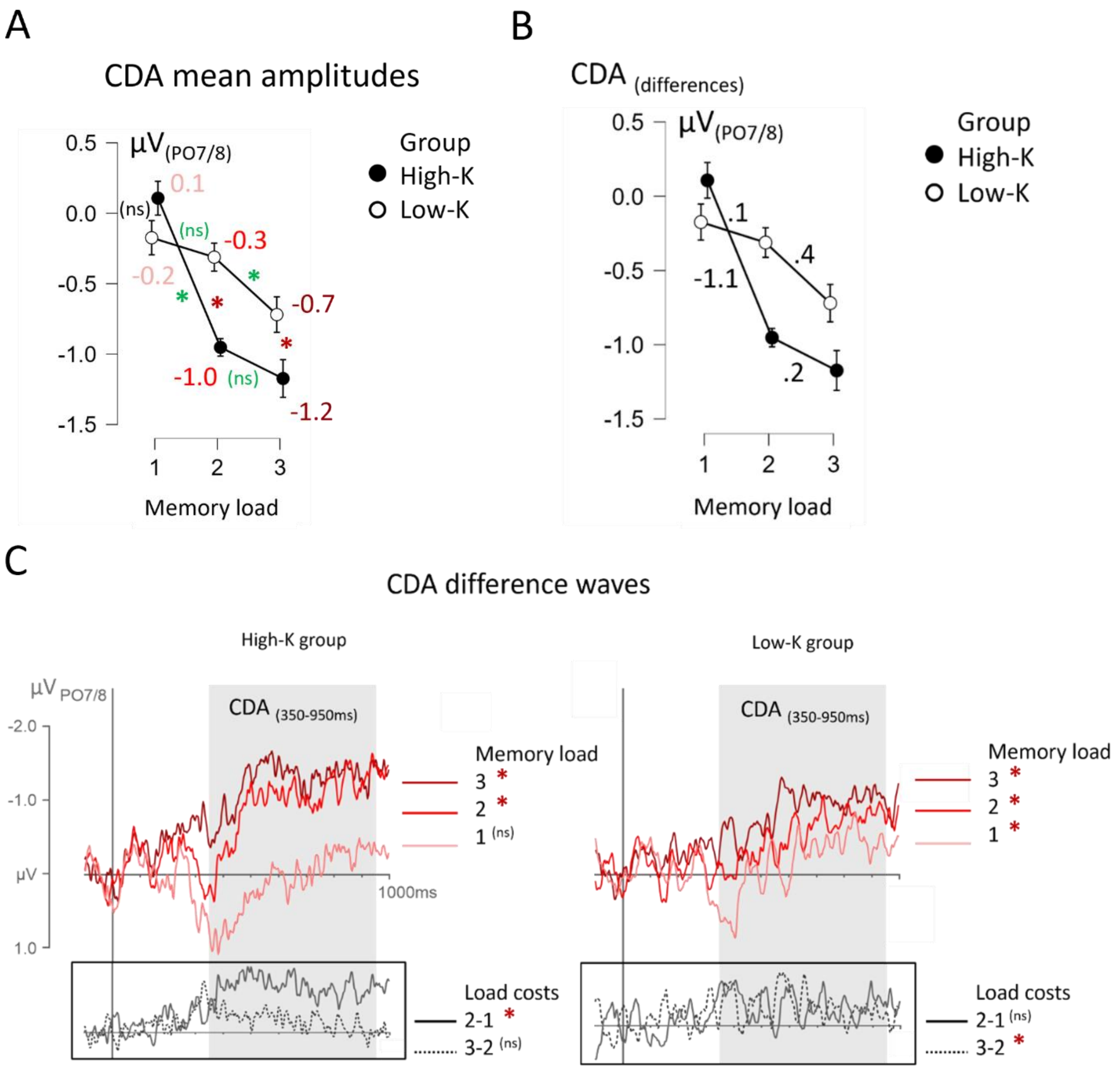


Figure 19 Contralateral delay activity (CDA) as a function of load and vWM capacity (K) group. A: mean CDA amplitudes at PO7/8 from 350–950ms post-stimulus, plotted against load (1–3 colours), separately for High- (filled circles) and Low-K (open circles) groups. Red asterisks indicate significant differences between groups and green asterisks represent significant differences within groups. B: CDA load costs, calculated as difference scores between successive loads (e.g., load 2 – load 1; load 3 – load 2), based on CDA difference waveforms. Error bars represent ±1 SEM. Panel C displays CDA difference waveforms at PO7/8 for each load condition, shown separately for High-K (left) and Low-K (right) groups. The shaded region indicates the CDA measurement window (350–950ms). The bottom waveforms in each panel reflect CDA load costs. Red asterisks indicate significant differences (*p*<.05).

4.3.2 Attention – Visual search task

4.3.2.1 Behavioural results

4.3.2.1.1 Mean error rates and reaction times

Error rates (ERs) were entered into a repeated-measures ANOVA with the within-subjects factor load (1, 2 and 3 colours). There was a significant main effect of load, *F*(2,68)=280.4, *p*<.001, η_p^2=.90. Descriptive statistics showed a substantial rise in errors across loads: load 1 (*M*=7.8%, *SD*=4.20), load 2 (*M*=19.0%, *SD*=6.64), and load 3 (*M*=25.8%, *SD*=7.30). Paired-samples *t*-tests confirmed that each increase was significant: load 1 vs load 2, *t*(34)=14.2, *p*<.001 and load 2 vs load 3, *t*(34)=11.1, *p*<.001. Cohen's *d* values indicated small to moderate effect sizes (*d*=.25 and .15, respectively). See Figure 20A for mean ERs.

Reaction times (RTs) were submitted to a repeated-measures ANOVA with load (1, 2 and 3 colours) as a within-subjects factor. There was a significant main effect of load on RTs, *F*(2,68)=219.8, *p*<.001, η_p^2=.87. RTs increased across loads: load 1 (*M*=616ms, *SD*=55.56), load 2 (*M*=707ms, *SD*=66.74), and load 3 (*M*=752ms, *SD*=76.44). Paired-samples *t*-tests confirmed these differences were significant: load 1 vs load 2, *t*(34)=6.0, *p*<.001, and load 2 vs load 3, *t*(34)=9.6, *p*<.001. As with error rates, Cohen's *d* values were small (*d*=.19 and .10). See Figure 20B for mean RTs.

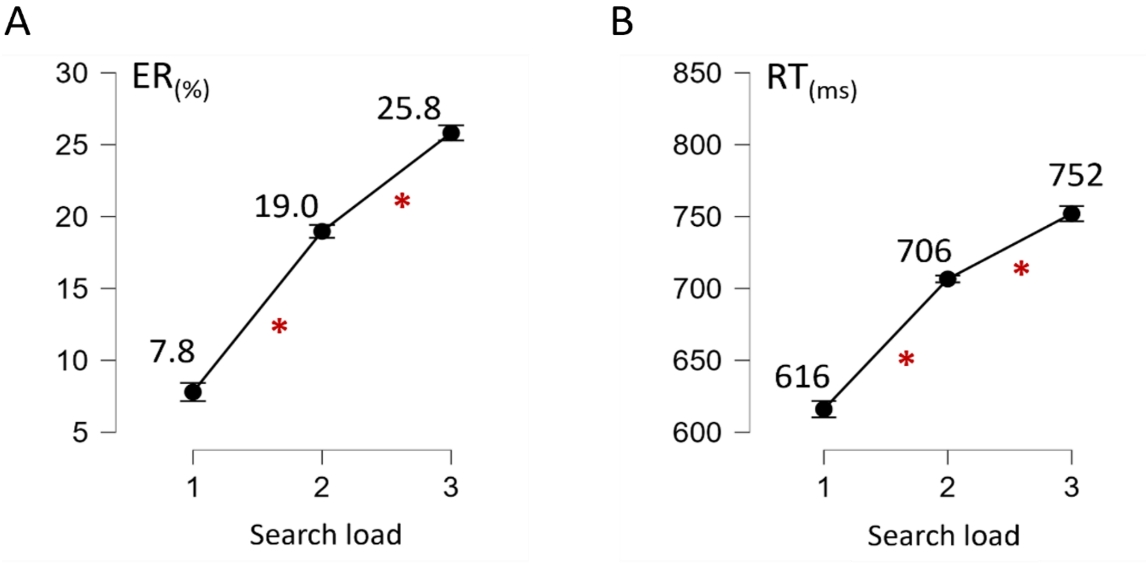


Figure 20 Search performance as a function of load. Panel A shows mean error rates (ER, %) and Panel B shows mean reaction times (RT, ms) across three levels of load (1, 2, and 3 colours). Error bars represent ±1 SEM. Red asterisks denote statistically significant differences between adjacent loads ($p<.05$).

4.3.2.2 ERP Results

4.3.2.2.1 Mean target N2pc amplitudes

N2pc amplitudes were submitted to a repeated-measures ANOVA with load (1, 2, and 3 colours) as a within-subjects factor. A significant main effect of load was found, $F(2,68)=47.6$, $p<.001$, $\eta_p^2=.58$. Descriptive statistics showed that N2pc amplitudes became smaller with increasing load: load 1 ($M=–1.30$, $SD=.67$), load 2 ($M=–.74$, $SD=.53$), and load 3 ($M=–.44$, $SD=.40$). Paired-samples t-tests confirmed significant reductions in N2pc amplitude between load 1 and load 2, $t(34)=7.3$, $p<.001$, and between load 2 and load 3, $t(34)=3.8$, $p<.001$. Paired-samples t-tests were conducted comparing activity at PO7 and PO8 electrodes for each load. Significant lateralisation was observed at all load levels, all $t>6.6$, all $p<.001$. See Figure 21A for difference waves and 21B for means.

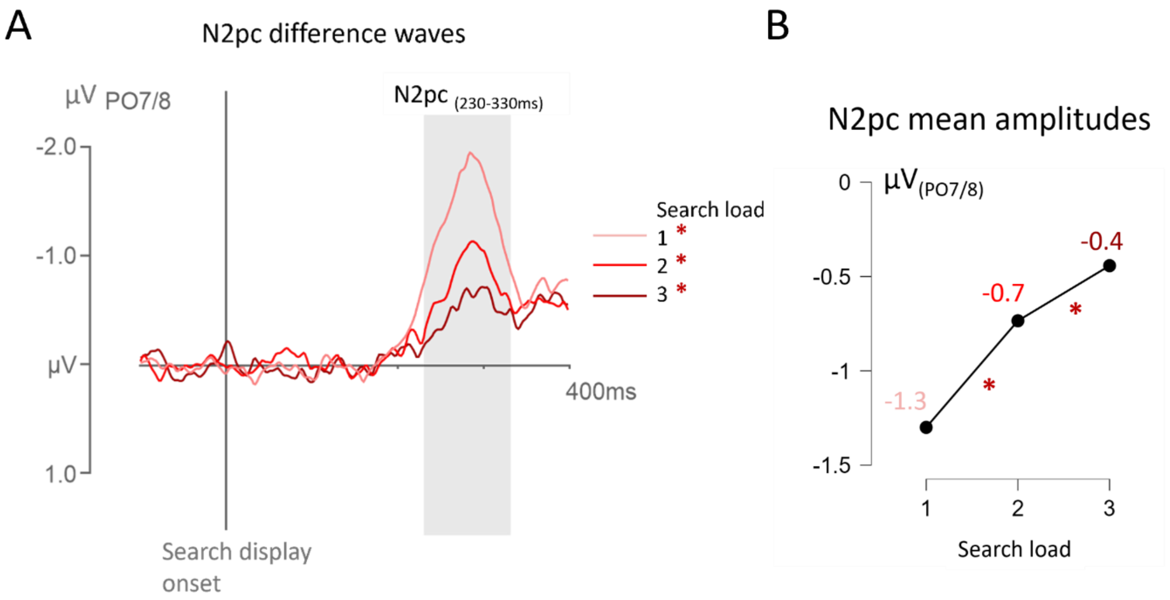


Figure 21 Target N2pc as a function of load. Panel A shows target N2pc difference waveforms at PO7/8 for loads of 1, 2, and 3 colours, with increased negativity corresponding to lower loads. The shaded region indicates the target N2pc time window (230–330ms post-stimulus). Red asterisks denote significant target N2pcs. Panel B shows mean target N2pc amplitudes as a function of load, with values shown above each point and red asterisks indicating statistically significant differences between adjacent loads. Error bars represent ±1 SEM.

4.3.2.2.2 Correlations between mean behavioural K and mean target N2pc amplitudes

To examine the relationship between behavioural performance (K values) and neural activity (N2pc amplitudes), Pearson correlations were calculated at each load level. No significant correlation was found between $K_{(load1)}$ and $N2pc_{(load1)}$, *r*(33)=–.05, *p*=.783 (Figure 22B). At load 2, the correlation approached significance, *r*(33)=–.32, *p*=.058 (Figure 22C). At load 3, the correlation was significant, (τ=–.24, *p*=.046; Figure 22D). Finally, a significant negative correlation was found between $K_{(mean)}$ and $N2pc_{(mean)}$, *r*(33)=–.35, *p*=.048 (Figure 22A). These findings indicate that individuals with greater K values tend to exhibit stronger (i.e., more negative) N2pc amplitudes, particularly under higher load.

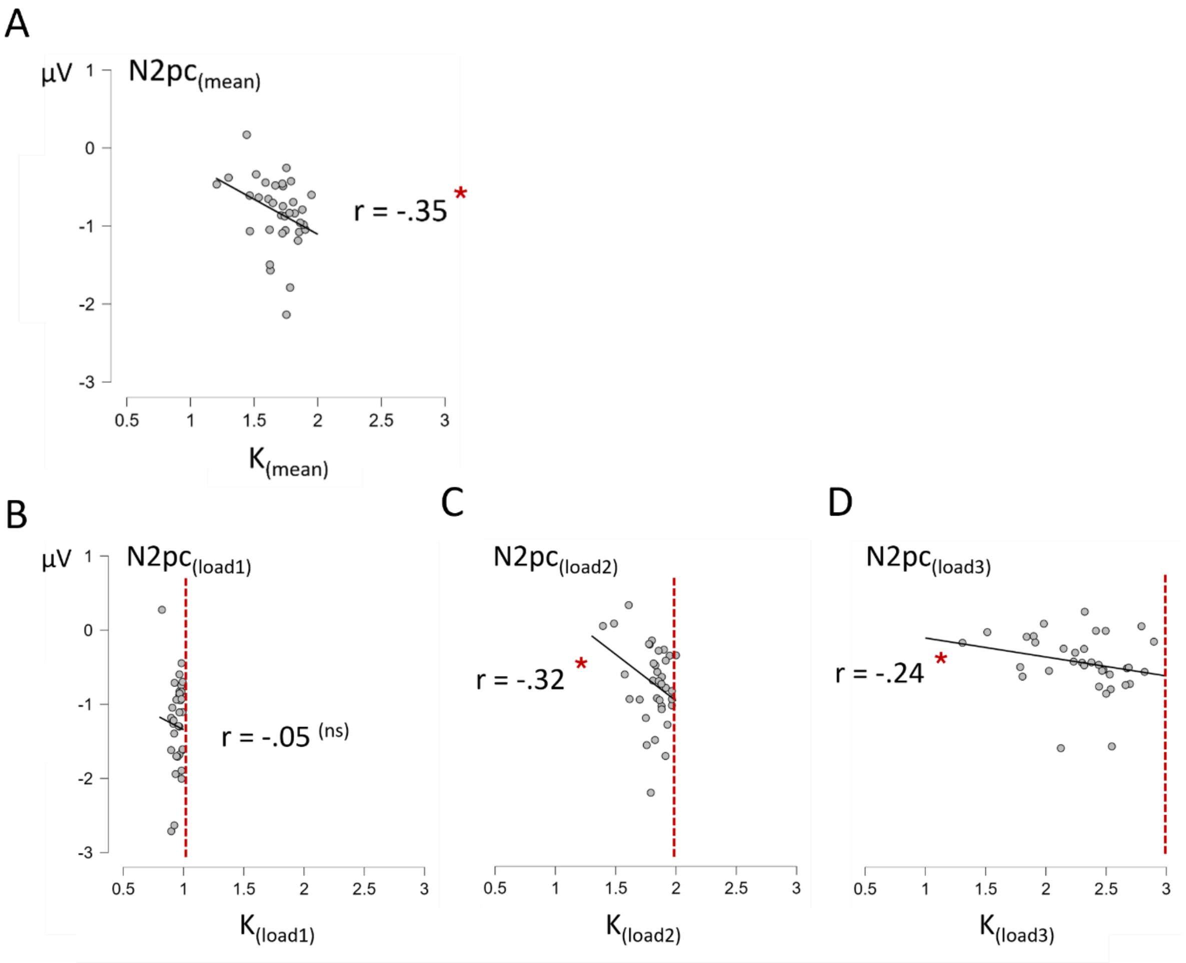


Figure 22 Scatterplots showing the relationship between target N2pc and vWM capacity (K values) across mean values (A) and individual loads (B, C, D). Each dot represents one participant. Panel A shows the correlation between mean N2pc amplitude and mean K across all loads. Bottom panels show correlations for load 1 (B), load 2 (C), and load 3 (D), with red dashed lines denoting approximate ceiling values for K at each load level. Pearson correlation coefficients (r) are displayed in each plot. Red asterisks indicate significant correlations (p<.05).

4.3.2.2.3 Mean target N2pc amplitudes in High- and Low-K groups

A 2 (K group: High vs Low) x 3 (Load: 1, 2 and 3 colours) mixed-design ANOVA was conducted on N2pc amplitudes. A significant main effect of load was found, $F(2,66)=46.8$, $p<.001$, $\eta_p^2=.59$, indicating that N2pc amplitude varied systematically across load levels. However, the main effect of K group was not significant ($p=.174$), nor was the load × K group interaction (*p*=.384). These results suggest that group-level N2pc differences were less pronounced than for CDA. Despite the lack of interaction, follow-up comparisons were conducted to explore trends between High-K and Low-K individuals. In both groups N2pc amplitudes were attenuated with increasing load. N2pc amplitudes were largest at load 1 (High-K: *M*=–1.36, *SD*=.69; Low-K: *M*=–1.22, *SD*=.65), followed by load 2 (High-K: *M*=–.89, *SD*=.49; Low-K: *M*=–.53, *SD*=.54), and weakest at load 3 (High-K: *M*=–.50, *SD*=.40; Low-K: *M*=–.37, *SD*=.40). In the High-K group, paired samples *t*-tests confirmed significant N2pc amplitude decreases from load 1 to load 2, *t*(19)=4.6, *p*<.001, and from load 2 to load 3, *t*(19)=3.6, *p*=.002. For participants in the Low-K group, N2pc amplitudes also decreased significantly from load 1 to load 2, *t*(14)=6.1, *p*<.001. However, the change from load 2 to load 3 was not significant (*p*=.137). This suggests that while attentional selection sharply declined with the initial increase in memory demands, it plateaued thereafter, indicating a possible upper limit in flexible resource allocation within this group. See Figure 23B for N2pc differences/load costs. Significant lateralisation emerged at all loads for both groups (High-K: all *t*>5.6, all *p*<.001; Low-K: all *t*>3.6, all *p*<.003). At each load, the High-K group showed larger N2pc amplitudes than the Low-K group. However, no significant differences were found at load 1, *t*(33)=.6, *p*=.530, or load 3, *t*(33)=1.0, *p*=.327. At Load 2, group differences approached significance, *t*(33)=2.0, *p*=.052. See Figure 23A for N2pc mean amplitudes and 23C for difference waves for each group.

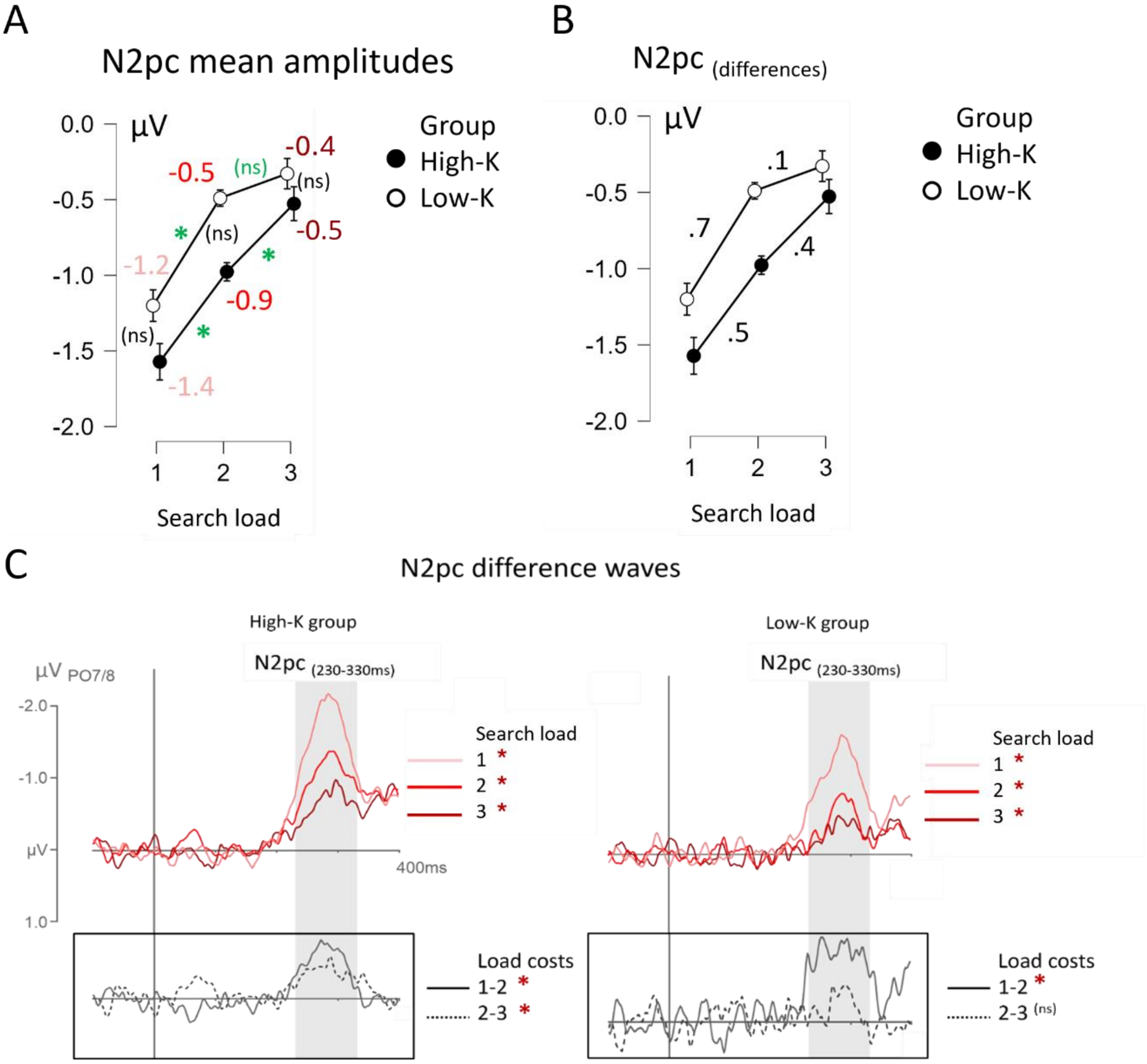


Figure 23 Target N2pc activity as a function of load and vWM capacity (K) group. A: Mean target N2pc amplitudes at PO7/8 from 230-330ms post-stimulus, plotted against load (1–3 colours), separately for High- (filled circles) and Low-K (open circles) groups. Red asterisks indicate significant differences between groups and green asterisks represent significant differences within groups. B: Target N2pc load costs, calculated as difference scores between successive loads (e.g., load 2 – load 1; load 3 – load 2), based on target N2pc difference waveforms. Error bars represent ±1 SEM. Panel C displays target N2pc difference waves at PO7/8 for each load condition, shown separately for High-K (left) and Low-K (right) groups. The shaded region indicates the target N2pc measurement window (230–330ms). The bottom waveforms in each panel reflect target N2pc load costs. Red asterisks indicate significant differences (*p*<.05).

4.3.3 Cross task correlations (change detection and visual search task)

4.3.3.1 K vs N2pc load costs between load 1 and load 2

To explore the relationship between behavioural K improvement and N2pc efficiency, a Pearson correlation was conducted between $K_{(2-1)}$ (change in K value from load 1 to load 2) and $N2pc_{(1-2)}$ (change in N2pc amplitude from load 1 to load 2). A significant correlation was observed, *r*(33)=.34, *p*=.048, suggesting larger differences in K were associated with smaller N2pc load costs. However, the assumption of bivariate normality was violated (Shapiro–Wilk, *W*=.854, *p*<.001) and Kendall's tau did not reach significance (τ=.148, *p*=.211), indicating the result should be interpreted cautiously (Figure 24A).

4.3.3.2 K vs N2pc load costs between load 2 and load 3

Pearson correlations were computed between $K_{(3-2)}$ (change in K value from load 2 to load 3) and $N2pc_{(2-3)}$ (change in N2pc amplitude from load 2 to load 3). A significant correlation was observed, *r*(33)=–.43, *p*=.009, indicating that individuals who showed larger differences in K between load 2 and load 3 exhibited larger N2pc load costs (Figure 24B).

4.3.3.3 CDA vs N2pc load costs between load 1 and load 2

To examine whether changes in neural activity associated with attentional selection (N2pc) were related to changes in vWM maintenance (CDA), a Pearson correlation was conducted between $N2pc_{(1-2)}$ (change in N2pc amplitude from load 1 to load 2) and $CDA_{(2-1)}$ (change in CDA amplitude from load 1 to load 2). A significant correlation was observed, *r*(33)=–.33, *p*=.050, indicating that smaller differences in CDA amplitude were associated with larger N2pc load costs (Figure 24C).

4.3.3.4 CDA vs N2pc load costs between load 2 and load 3

A Pearson correlation was conducted between $N2pc_{(2-3)}$ (change in N2pc amplitude from load 2 to load 3) and $CDA_{(3-2)}$ (change in CDA amplitude from load 2 to load 3). A significant correlation was found, *r*(33)=.34, *p*=.047, indicating that individuals who exhibited larger differences in CDA amplitude from load 2 to load 3 also showed larger N2pc load costs (Figure 24D).

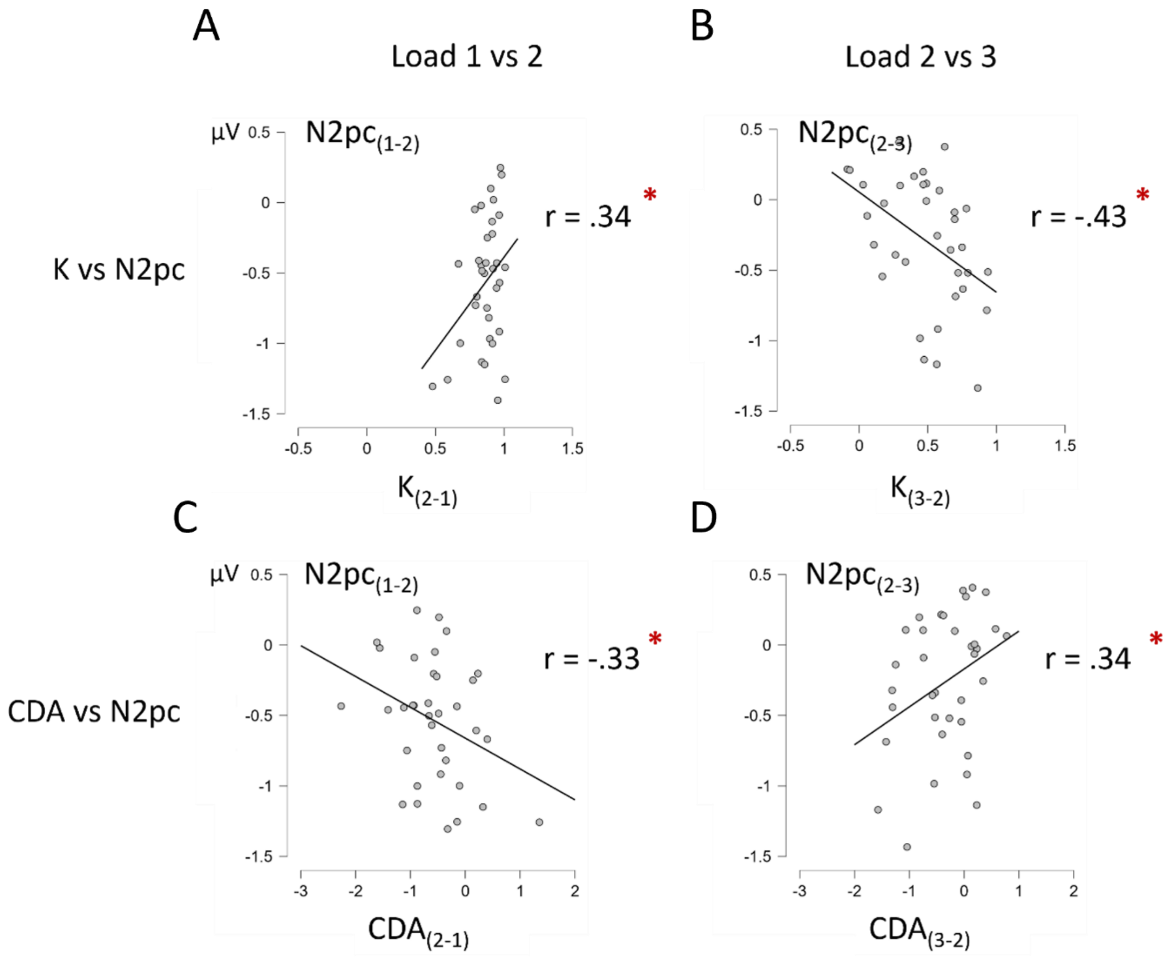


Figure 24 Scatterplots illustrating the relationships between changes in N2pc amplitude and changes in behavioural vWM capacity (K) and CDA amplitude across loads. A: Correlation between N2pc amplitude difference from load 1 to load 2 ($\text{N2pc}_{(1-2)}$) and K difference ($\text{K}_{(2-1)}$). B: Correlation between $\text{N2pc}_{(2-3)}$ and $\text{K}_{(3-2)}$. C: Correlation between $\text{N2pc}_{(1-2)}$ and CDA difference ($\text{CDA}_{(2-1)}$). D: Correlation between $\text{N2pc}_{(2-3)}$ and $\text{CDA}_{(3-2)}$. Red asterisks indicate significant correlations (*p*<.05).

4.3.4 Rapid serial probe presentation (RSPP)

To determine the time course of template activation in preparation for search, N2pc components triggered in each of the seven successive probes (probes 1–7) were measured by computing ERPs at posterior sites PO7/8, contralateral and ipsilateral to the side of a probe, separately for probes that matched the upcoming target-colour (target-colour probes) and probes that did not match the upcoming target-colour (distractor-colour probes). Probe N2pc difference waves, obtained by subtracting ipsi from contralateral ERPs at PO7/8 for each individual probe, are shown in Figures 25-28. Figure 25-27 displays target-colour probe N2pcs and Figure 28 displays distractor-colour probe N2pcs. All figures show probes 1–7 in a temporally continuous fashion, separately for High-K (top panel of the figures) and Low-K (bottom panel of the figures) groups. Note that N2pc components were extracted individually for each probe (probes 1–7) and that Figures 25-28 simply illustrate these probe N2pcs in a successive format. As probes appeared every 200ms, each individual probe was therefore presented within the N2pc time interval of its immediately preceding probe. The visualisations begin with probe 1, which was the first probe presented after a preceding search display and is shown from 100ms before onset to 350ms post-onset. For probes 2–7, each is shown in a 200ms window from 150 to 350ms post-onset, with temporally interpolated data points between adjacent intervals. The onset of each probe is marked by a vertical line, and the N2pc measurement window (210-290ms post-stimulus) is shaded in grey in all panels.

4.3.4.1 N2pc target-colour probes

To assess modulation of the N2pc across the preparation interval, ERP mean amplitudes (210-290ms post-probe onset) at PO7/8 were entered into repeated-measures omnibus ANOVA. For target-colour probes, separate analyses were conducted for High- and Low-K groups. Each analysis included the within-subject factors of load (1, 2 and 3 colours), probe number (probes 1–7), and laterality (contralateral vs ipsilateral). For the High-K group, there was a significant main effect of laterality, *F*(1,19)=22.0, *p*<.001, η_p^2=.54, confirming the presence of an N2pc. This effect was qualified by a significant laterality × probe number interaction, *F*(6,114)=13.1, *p*<.001, η_p^2=.41. There was no interaction with load. Follow-up ANOVAs with the factors probe number (probes 1–7) and laterality (contralateral vs ipsilateral) revealed significant N2pcs at probe 5, *F(*1,19)=16.0, *p*<.001, probe 6, *F*(1,19)=23.1, *p*<.001, and probe 7, *F*(1,19)=39.8, *p*<.001. Paired-samples *t*-tests were conducted to compare ERP amplitudes at contralateral (PO8) and ipsilateral (PO7) electrodes for each probe and load condition. At load 1, the earliest probe to produce a reliable N2pc was probe 5 (*t*(19)=3.7, *p*=.005), continuing at probe 6 (*t*=3.4, *p*=.003) and probe 7 (*t*=5.0, *p*<.001; Figure 25). At load 2, lateralised activity was significant at probe 6 (*t*=2.1, *p*=.048) and again at probe 7 (*t*=4.5, *p*<.001; Figure 26). At load 3, significant N2pcs emerged at probe 5 (*t*=3.0, *p*=.008) probe 6 (*t*=2.2, *p*=.039) and probe 7 (*t*=4.5, *p*<.001; Figure 27). These results confirm robust N2pc components in the later stages of the preparation period.

For the Low-K group, there was no significant main effect of laterality, *F*(1,14)=1.4, *p*=.260, η_p^2=.09. Notably, there was a significant laterality × probe number, *F*(6,84)=6.7, *p*<.001, η_p^2=.32 and laterality x load interaction, *F*(2,28)=3.8, *p*=.035, η_p^2=.21. To investigate load-specific lateralisation, paired samples t-tests were conducted separately for each load. At load 1, significant lateralised effects were observed at probe 6 (*t*=3.7, *p*=.005) and probe 7 (*t*=3.4, *p*=.003; Figure 25). At load 2, probe 7 also elicited a significant N2pc (*t*(19)=2.1, *p*=.048; Figure 26). However, load 3 probes did not elicit any significant N2pcs (Figure 27). At probe 7, a significant effect of laterality was found, *F*(1,14)=14.3, *p*=.002, as well as a significant laterality x load interaction (*F*(2,28)=3.7, *p*=.038), indicating that the presence of an N2pc at this later time point depended on the level of load.

4.3.4.2 N2pc distractor-colour probes

No reliable N2pcs were observed for distractor-colour probes, in either the High-K or Low-K group. For the High-K group, there was no main effect of laterality, $F(1,19)=.68$, $p=.420$, $\eta_p^2=.04$, but there was a significant interaction between laterality and load, $F(2,38)=7.7$, $p=.002$ and laterality and probe number, $F(6,114)=2.5$, $p=.048$. For the Low-K group, the omnibus ANOVA yielded no significant main effects but there was a significant interaction between laterality and probe number, $F(6,84)=2.5$, $p=.029$. Only isolated *t*-tests reached marginal significance (e.g., probe 6, $t=2.1$, $p=.050$), but these were not consistent across the preparation window and did not reflect reliable N2pc components (Figure 28). Overall, the pattern confirms that only target-colour probes elicited reliable N2pcs, and these emerged most robustly for the High-K group in later stages of the preparation period. Distractor colour probes failed to produce reliable N2pc components, suggesting that they were not selected during search preparation.

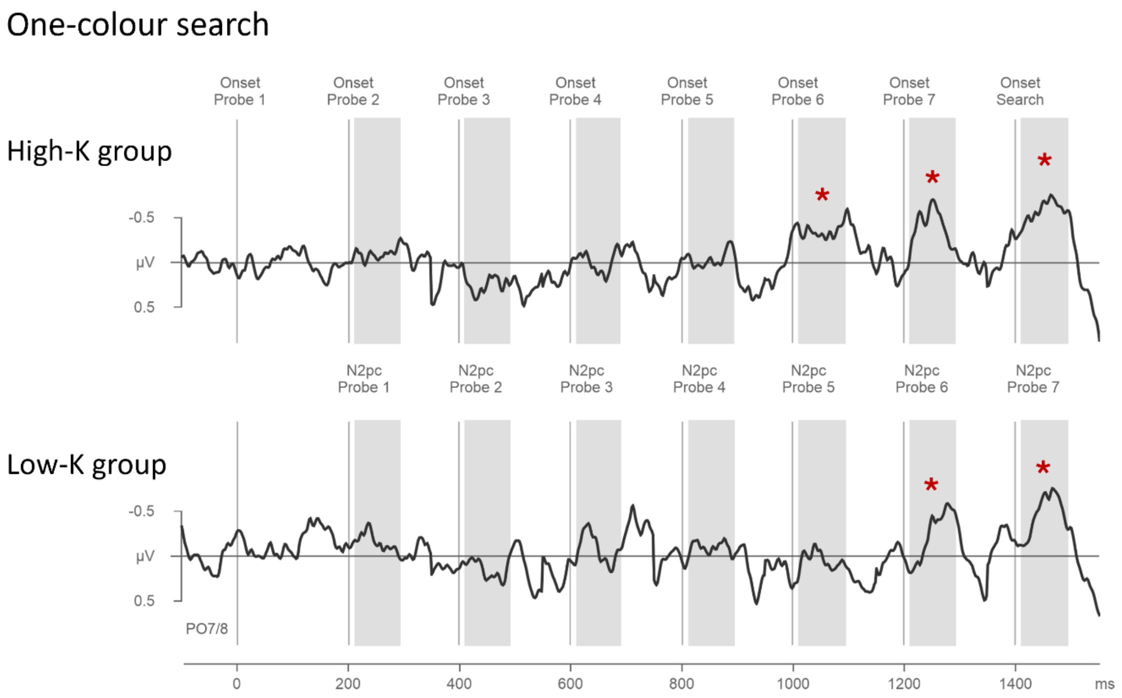


Figure 25 N2pc difference waveforms obtained by subtracting ipsilateral from contralateral ERPs triggered by the target-colour probes in the one-colour search, High-K group (top panel) and Low-K group (bottom panel). Here, difference waves for the seven probes (probes 1–7) are illustrated in a temporally continuous fashion, but the seven individual probe N2pc were extracted independently of each other from the raw signal. Probe onsets are indicated by vertical lines, and probe N2pc time windows by shaded areas (210–290ms after the onset of each individual probe). Statistically reliable probe N2pcs are marked by asterisks.

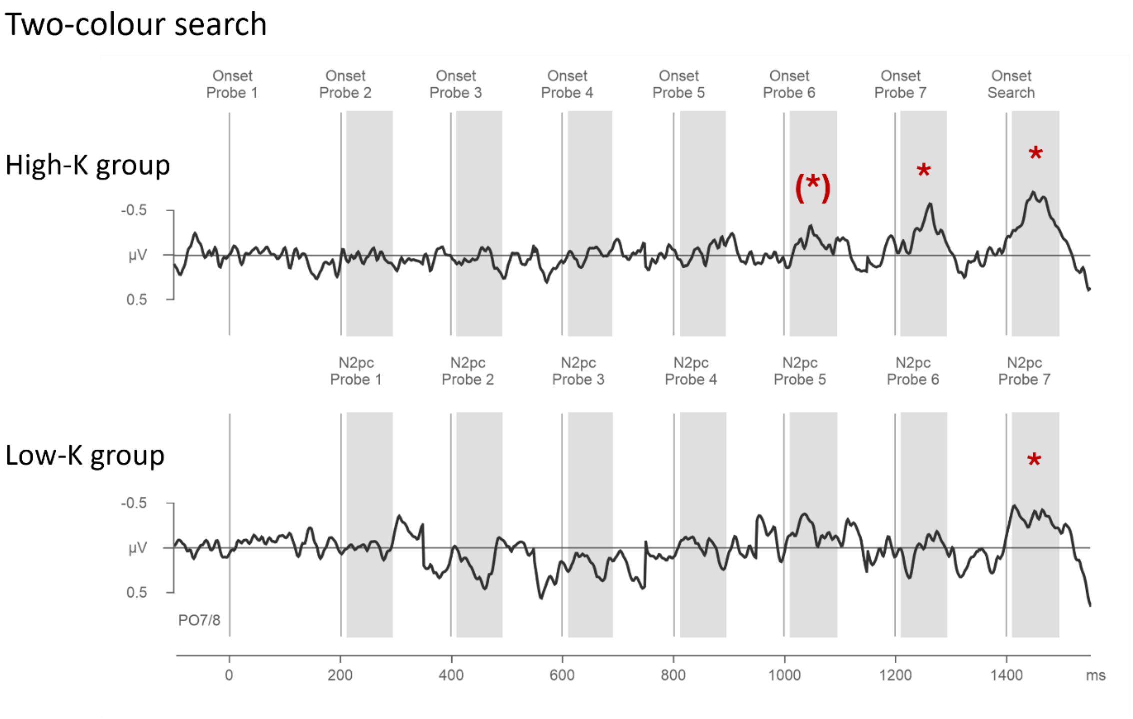


Figure 26 N2pc difference waveforms obtained by subtracting ipsilateral from contralateral ERPs triggered by the target-colour probes in the two-colour search, High-K group (top panel) and Low-K group (bottom panel). Difference waves triggered by individual probes are shown in the same continuous fashion as in Figure 14. Probe onsets are indicated by vertical lines, and probe N2pc time windows by shaded areas (210–290ms after the onset of each individual probe). Statistically reliable probe N2pcs are marked by asterisks.

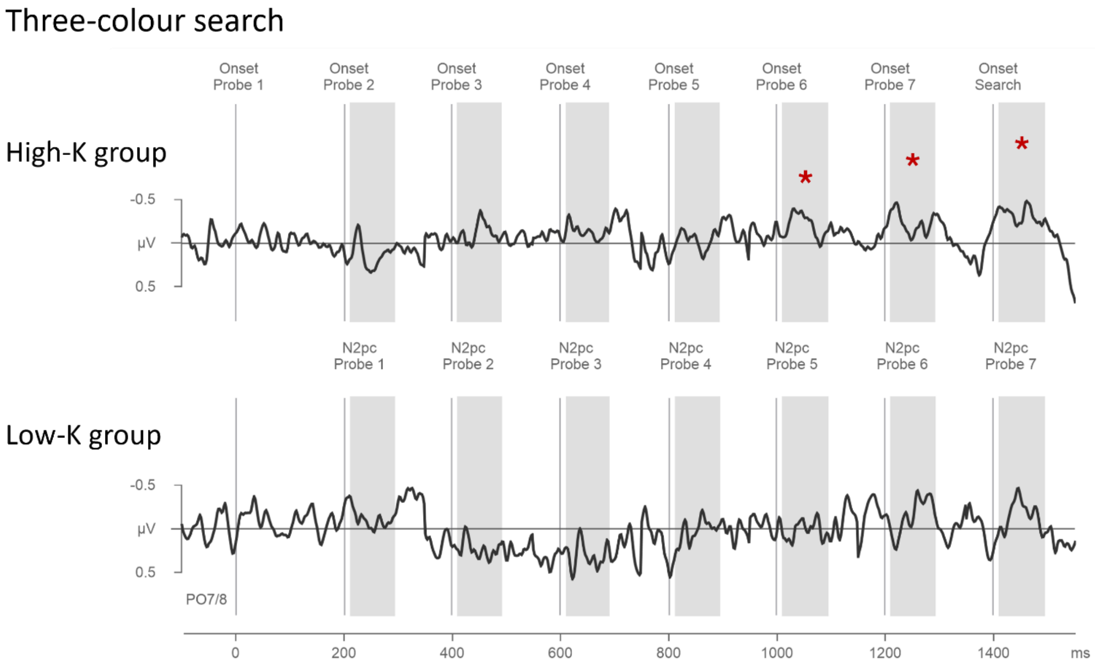
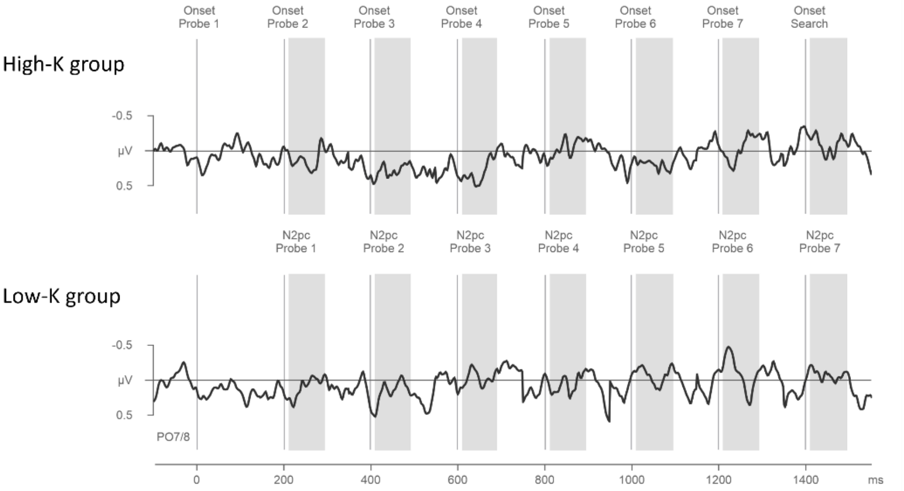
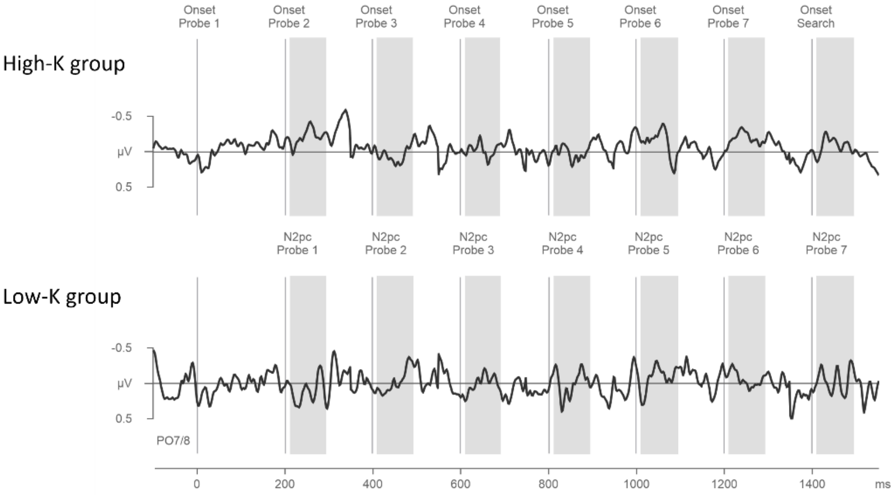


Figure 27 N2pc difference waveforms obtained by subtracting ipsilateral from contralateral ERPs triggered by the target-colour probes in the three-colour search, High-K group (top panel) and Low-K group (bottom panel). Difference waves triggered by individual probes are shown in the same continuous fashion as in Figure 14. Probe onsets are indicated by vertical lines, and probe N2pc time windows by shaded areas (210–290ms after the onset of each individual probe). Statistically reliable probe N2pcs are marked by asterisks.

One-colour search



Two-colour search



Three-colour search

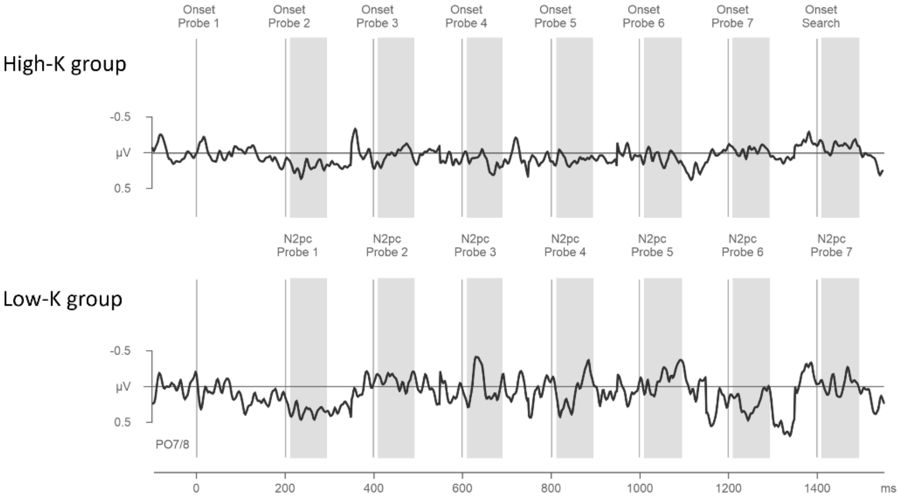


Figure 28 N2pc difference waveforms obtained by subtracting ipsilateral from contralateral ERPs triggered by the distractor-colour probes in top panel: One-colour search, middle panel: Two-colour search, and bottom panel: Three-colour search. Waveforms are plotted separately for High- and Low-K groups. Difference waves triggered by individual probes are shown in the same continuous fashion as in Figure 14. Probe onsets are indicated by vertical lines, and probe N2pc time windows by shaded areas (210–290ms after the onset of each individual probe.

4.4 Discussion

The goal of the present study was to examine whether individual differences in vWM capacity predict the ability to activate and use multiple attentional templates during visual search. To our knowledge, no prior work has examined this relationship across multiple task contexts; here we tested it using two distinct paradigms: a change detection task to index vWM capacity via K and CDA and an RSPP search task that assessed search capacity and real time template activation through error rates, reaction times and target- and probe-evoked N2pc components, respectively.

The results from the change detection task confirmed the effectiveness of both behavioural (K) and electrophysiological (CDA) measures in indexing vWM capacity. Both K values and CDA amplitudes increased with load, and this pattern was observed across both high- and low-K groups. Importantly, K and CDA amplitudes were significantly correlated: participants with higher K values showed larger CDAs, whereas those with lower K values showed smaller CDAs, supporting the validity of the CDA as a marker of vWM capacity. More detailed analyses of CDA load effects revealed a capacity dependent pattern. High-K individuals showed no reliable CDA at load 1, but significant amplitude differences between load 1 and load 2, suggesting they only began to recruit sustained maintenance at load 2. The additional increase from load 2 to load 3 was smaller, consistent with a capacity near 2-3 items (Mean K=2.6 at load 3 in this group). In contrast, for low-K individuals, the opposite pattern emerged. They already showed a substantial CDA at load 1, leaving less room for further loads. The load 1 to load 2 increase was small, with minimal further change from load 2 to load 3, consistent with an average capacity below 2 items (Mean K=1.9 at load 3 in this group). Thus, this 'double difference' arises because high-K individuals begin to express CDA at higher loads, plateauing higher, whereas low-K individuals show CDA from load 1 and thus plateau at a lower level. These findings suggest that individuals not only differ in overall capacity, but also in how they distribute memory resources across increasing loads.

In the visual search task, behavioural performance declined as the number of potential target templates increased: with both error rates and reaction times increasing across one, two and three search loads. While these load-related costs are consistent with prior work (Ort et al., 2019), they likely reflect post-selection processes rather than limitations in template activation itself. At the neural level, across all participants, N2pc amplitudes decreased with increasing search load. This reduction most likely reflects a greater proportion of trials in the load 2 and load 3 colour search in which the target was not selected, minimising the averaged N2pc relative to the load 1 condition. Follow up analyses showed that this attenuation was driven primarily by a subset of low-K participants. Importantly, vWM capacity (K) predicted these N2pc responses. High-K participants tended to produce larger N2pc components overall. One plausible interpretation is that this reflects a higher likelihood of successful target selection across trials, rather than a stronger single trial neural response, although this cannot be determined definitively from the present data. They also showed significant reductions from load 1 to 2 and again from load 2 to 3, reflecting sensitivity to increasing template competition. This pattern aligns with accounts of mutual inhibition between simultaneously activated templates (Kerzel & Grubert, 2022; Ort & Olivers, 2020) and suggests high-K individuals can engage in more flexible template prioritisation. In contrast, low-K participants also showed reduced N2pc amplitudes at higher search loads, but significant load costs were only observed between load 1 and 2, not from 2-3. By load 2 the N2pc was already very small, leaving little room for further reduction at load 3, potentially indicating their inability to activate more than two (often fewer than two) templates (Olivers et al., 2011). Relevantly, as ERPs are averages, amplitude differences are also influenced by the number of target selected trials. Low-K participants contributed more late selection trials at higher loads, which attenuated their N2pc.

The correlational analyses between tasks provided key insights into how vWM capacity, measured behaviourally (K) and neurally (CDA) in the change detection task, relates to the efficiency of attentional template activation in visual search. Participants with higher K values and larger CDA amplitudes showed increased N2pc responses across increasing search load, suggesting that individuals with greater vWM capacity were better able to activate and maintain multiple attentional templates. Importantly, these correlation patterns varied for the different load conditions. For example, smaller differences in K and CDA amplitudes between load 1 to 2, as produced by low-K individuals, were associated with larger N2pc reductions from load 1 to 2. This suggests that low-K individuals experienced more attentional costs when required to activate more than one template. In contrast, larger K and CDA differences between load 2 and 3, as produced by high-K individuals, were associated with larger N2pc reductions from load 2 to 3. This suggests that high-K individuals maintained attentional efficiency even when managing three potential templates. Overall, these findings reinforce the fact that vWM varies across individuals (Awh et al., 2007; Cowan et al., 2005; Fukuda et al., 2010a; Luck & Vogel, 2013), and so does the capacity to activate and use multiple attentional templates. Low-K individuals showed larger costs between one and two-colour search, while high-K individuals showed larger costs between two- and three-colour search, consistent with a capacity threshold that differs across individuals.

To track target template activation in real time, N2pc components were recorded in response to brief probe displays containing a colour singleton that either matched the upcoming target or a random distractor colour. These displays appeared in rapid succession between search displays. Similar to previous studies that employed analogous RSPP procedures (Grubert & Eimer, 2018, 2020, 2023), probes that matched the fully predictable upcoming target-colour in the one-colour task triggered N2pc components when they appeared during the 800ms interval prior to the arrival of the next search display. This shows that these probes attracted attention, and that a corresponding colour-specific target template was active at the moment when they were presented. However, in the current study, participants were required to maintain up to three potential target colours, each of which was likely to appear on any given trial. Thus, template activation could not be driven by certainty about the upcoming target colour but instead reflect the ability to maintain multiple target templates in parallel. Importantly, these effects varied as a function of individual vWM capacity. High-K participants demonstrated robust N2pc components to target-colour probes, particularly during the later stages of the preparation interval (probes 5-7), across all load conditions. This suggests that individuals with higher vWM capacity were able to activate multiple target templates simultaneously, allowing them to selectively respond to matching probes even for up to three colours. In contrast, low-K participants showed delayed and attenuated N2pc probe responses. While some activity emerged for the one- and two-colour conditions at probes 6 and 7, no reliable N2pc was observed in the three-colour condition, suggesting individuals were unable to activate three templates simultaneously. These differences suggest that low-K individuals may rely on sequential rather than parallel template activation, which is in line with accounts that propose limitations in simultaneous template use (Olivers et al., 2011).

While the presence of reliable probe-evoked N2pcs in the three-colour condition, particularly for the high-K group, suggests parallel template activation, this finding seems to contradict the behavioural and target-evoked N2pc costs observed under the same three-colour condition. This disassociation mirrors findings observed in Grubert et al. (2025), who found that probe-evoked N2pcs remained robust during three-colour search, even as search performance declined. This suggests that while templates may be activated in parallel, increasing the number of active templates reduces the efficiency of attentional guidance during the actual search task.

One possible explanation, derived from Grubert et al. (2025), is that probe-evoked N2pcs reflect not the absolute strength of template activation, but rather the interactive effect between a visual probe and an activated template. Specifically, when a probe matches one of the currently held target colours, it may transiently boost the activation of that template, thus eliciting a reliable N2pc even if the overall activation is reduced under high-load conditions. This may account for the presence of similar probe N2pc amplitudes across all load conditions despite increased behavioural costs in the higher load searches. Indeed, the present findings are consistent with this interpretation. High-K individuals showed sustained probe-evoked N2pcs across all loads and low-K individuals exhibited delayed and weaker N2pc responses, especially in the three-colour condition.

These results show that the RSPP method provides a unique window into template dynamics during search preparation. However, probe-evoked N2pc components should be interpreted not as direct measures of template strength, but as reflections of attention being captured by stimuli that match active templates. The combination of probe-evoked N2pc amplitudes, target-evoked N2pc amplitudes, and behavioural performance suggests that while individuals can activate multiple templates in advance, the efficiency of using these templates for attentional selection declines as capacity limits are approached.

While the current study provides strong evidence that individual differences in vWM capacity limit the ability to activate and use multiple attentional templates during visual search, several caveats should be acknowledged. Firstly, although both K and CDA amplitudes provided converging indices of vWM capacity, it cannot definitively be determined whether observed search limitations reflect a hard storage limit or simply the difficulty to allocate or prioritise templates under competitive load. Factors such as strategic control, motivation, or long-term memory contributions (Grubert et al., 2016) may also modulate template activation. Secondly, although stronger N2pc responses were linked to higher capacity individuals, it remains unclear whether this reflects more efficient selection, stronger template activation, or both. Finally, while the current design isolated capacity using separate tasks for vWM and visual search, real-world contexts often integrate memory, attention, and learning over time. Thus, this may introduce context effects that are not present in natural search environments.

Taken together, the present findings provide direct support for the hypothesis that vWM capacity constrains the number of attentional templates that can be concurrently activated during visual search. Consistent with parallel template accounts (Beck et al., 2012; Berggren & Eimer, 2019; Christie et al., 2015; Grubert et al., 2016; Grubert & Eimer, 2016a, 2016b, 2023; Irons et al., 2012; Kerzel & Grubert, 2022, Kristjánsson & Kristjánsson, 2018; Moore & Weissman, 2010; Ort et al., 2019), the current results show that multiple templates can be activated at once. However, critically, the extent to which this is possible varies across individuals. High-capacity individuals can flexibly maintain and prioritise multiple active templates, while low-capacity individuals show inability to simultaneously activate more than one or two templates. These findings help to explain previous group level results showing stable template activation across increasing load (Grubert & Eimer, 2023) with the possibility that such effects are not the same across individuals. More broadly, they show that vWM not only stores visual information but can be employed to guide top-down attentional selection.

In summary, the current study provided evidence that individual differences in vWM measured via CDA and K in a change detection task, predict the efficiency of attentional guidance during visual search involving multiple target templates. By employing two distinct paradigms and tracking real-time template activation using the RSPP method, we were able to assess how capacity-related neural markers generalise across cognitive domains. The findings suggest that vWM capacity constrains not only how much information can be maintained, but also how flexibly that information can be used to guide attention when multiple templates are active. This cross-task relationship highlights the functional integration of memory and attention systems and underscores the importance of considering individual capacity limits when examining variability in search performance.

Chapter 5: General Discussion

5.1 Summary of key findings across studies

This thesis set out to examine vWM capacity through a detailed investigation of K. Specifically, the overall aim of this thesis was to move beyond a fixed conception of vWM capacity (K) and examine a more functional, flexible and context-dependant approach. Across three chapters, I examined how K is shaped by perceptual parameters (Chapter 2) and learned context (Chapter 3) and how it relates to individual differences in attention (Chapter 4). This work not only contributes to the bigger picture that vWM capacity is not merely a fixed limit but also offers new insights into the functional nature of vWM, expressing K as a dynamic construct shaped by both internal and external demands. In this final chapter (Chapter 5), I firstly go over the key findings across the three studies, then discuss their broader theoretical and methodological implications, consider limitations and future directions and conclude with the contribution this thesis makes to our understand of vWM capacity.

While K is often interpreted as a fixed indicator of how many items an individual can store, the studies presented here challenged this view by suggesting that K reflects both individual differences and flexible aspects of cognitive functioning, shaped by task demands and environmental regularities.

Study 1 (Chapter 2) examined whether K remains stable when perceptual parameters of the memory display are systematically varied. Across five behavioural experiments, various perceptual manipulations (stimulus density, stimulus eccentricity, stimulus organisation, stimulus type and spatial variability of stimuli) were altered in a change detection task. Results demonstrated that while individual K values were reliable across conditions (i.e. participants with a high K in one condition tended to score high in another), absolute K values varied significantly depending on the perceptual structure of the display. All experiments, apart from Experiment 3, showed significant differences across conditions, suggesting that K is not robust against low level visual properties. Nevertheless, the consistency within participants demonstrates that K still shows meaningful trait-like differences.

Study 2 (Chapter 3) extended this by asking whether vWM capacity can be enhanced through contextual learning. The study embedded repeated spatial configurations into a change detection task to see if environmental regularities could support vWM performance and enhance capacity. Two experiments explored this under different conditions. Experiment 1 tested whether repeated configurations improve vWM performance across two memory loads (two and four). Participants showed significantly higher K values in repeated configurations as opposed to novel ones. These behavioural gains were also shown by changes in the CDA. These findings suggest that learned contexts not only improve performance but also modulate the neural efficiency of memory maintenance. Experiment 2 investigated more specifically whether this effect depends on the spatial relevance of the display. Participants performed the change detection task under two conditions: visuospatial (context was defined by colour and location) and non-spatial (context was defined by colour only). The benefits of the repeated configurations were only found in the visuospatial condition, suggesting spatial regularities are key for contextual learning effects in vWM. This study provided strong evidence that contextual learning can directly support vWM capacity, but only when the learned context is spatially relevant.

Study 3 (Chapter 4) investigated whether individual differences in K predict the number of attentional templates that can be activated and used during visual search. This study combined a change detection task (to measure K and CDA amplitudes) with an RSPP visual search task to assess attentional selection via N2pc components. Across three memory loads (1, 2, and 3), individuals with higher K showed higher behavioural accuracy and more robust probe- and target-evoked N2pcs in the search task. These results suggest that those with higher vWM capacity can maintain and use more templates simultaneously. These results reinforce the idea that capacity is not only about how much information can be stored, but how efficiently that information can guide attention.

5.2 Theoretical implications for models of vWM capacity

The findings across all three studies have important implications for how vWM is conceptualised within existing frameworks. Rather than supporting a single framework, the results here draw to a bigger picture of hybrid and integrative models, which incorporate core aspects from discrete-slot, fixed resource and interference models of vWM capacity.

Discrete-slot models propose a fixed number of items (three to four) that can be held in vWM, regardless of their features (Cowan, 2001; Luck & Vogel, 1997). The findings from Study 1 provide mixed support for this model. For example, in Experiment 1a, K peaked at five colours but declined when increasing to seven colours, suggesting a capacity limit of around four items, consistent with previous research also supporting a discrete-slot model (Luck & Vogel, 1997; Rouder et al., 2008; Vogel et al., 2001). However, in Experiment 1b, many distractors impaired capacity relative to trials without distractors, despite the target set size remaining fixed. This experiment indicates that distractor interference can diminish capacity even before a slot limit is reached. This finding is not easily explained by the discrete-slot model which is based purely on item count. Similarly, the discrete-slot model struggles to account for the differences observed across conditions in Experiment 2. In

Experiment 2a and 2b, K declined as eccentricity increased from near to far conditions, suggesting that eccentricity can modulate capacity even within the same set size. As neither item complexity or set size changed, these effects are difficult to explain with both slot and resource models. Instead, these results support interference-based accounts (Oberauer & Kliegl, 2001), which propose that it is not only the number of items stored but how well they can be differentiated in space.

Perceptual grouping in Experiment 3 had no effects on capacity estimates, suggesting that Gestalt principles may not consistently enhance encoding under these conditions, inconsistent with previous research which finds strong differences in gestalt principles (Jiang et al., 2000; Peterson & Berryhill, 2013; Woodman et al., 2003; Xu, 2002, 2006; Xu & Chun, 2007). In Experiment 4, K varied largely by stimulus type, supporting previous research (Alvarez & Cavanagh, 2004). Results showed higher K values for letters than for shapes, demonstrating that item complexity influences capacity, consistent with flexible resource-based predictions. Zhang and Luck's (2008) hybrid model suggests that multiple slots can be assigned to a single item to increase representational precision when fewer items are encoded. In other words, when the total number of items are below an individual's maximum capacity, the 'unused' slots do not remain inactive but can instead be devoted to strengthening the representation of existing items, meaning each item benefits from more memory resources. Thus, this could help explain the high K values observed in the letter trials, where fewer complex features may have allowed for precision enhancing slot allocation. In Experiment 5, K declined in the randomised locations condition, consistent with previous research (Jiang et al., 2000; Logie et al., 2011) which aligns with interference-based models (Oberauer & Kliegl, 2001, 2006). This model states that when spatial or feature bindings vary, items compete for representation. Thus, capacity appears reduced, not because fewer items can be stored, but because their representations are more vulnerable to interference.

Study 1 highlights the importance that no single capacity model fully captures the dynamics of vWM capacity. Collectively, the results of Study 1 are best explained by hybrid (Zhang & Luck, 2008) or interference-based models (Oberauer & Kliegl, 2001, 2006) which incorporate aspects of slot-based approaches with dynamic encoding constraints. These models suggest that interference, spatial crowding, or perceptual competition can reduce capacity, even without increased set size or complexity of items (Davelaar et al., 2005; Oberauer & Kliegl, 2006). Overall, Study 1 shows that vWM capacity is not purely restricted by item limits but is shaped by how visual information is organised and structured.

While Study 1 specifically emphasised how perceptual structure modulates vWM performance, Study 2 shifted focus to the influence of contextual regularities on vWM capacity. The findings from Study 2 provide strong support for the view that vWM capacity is not strictly determined by fixed item limits but can be flexibly modulated by learned regularities in the environment. Thus, the observed capacity improvements under repeated spatial configurations in this study are difficult to settle with traditional slot-based models (Cowan, 2001; Luck & Vogel, 1997) which assume capacity is fixed regardless of context. Instead, and like Study 1, the results in Study 2 are better accounted for by hybrid models of vWM capacity, such as Zhang and Luck's (2008) proposal. In repeated configurations, with stable spatial regularities, this allocation may become more efficient. These results are further supported by the interference-based model (Oberauer & Kliegl, 2006) which suggests that vWM is not only limited by quantity but also item interference. For instance, repeated configurations, compared to novel configurations, may reduce the overlap between stored features, minimising interference, thus allowing for more efficient storage.

Neural evidence from the CDA in Study 2 also presents further evidence to these interpretations. As such, the CDA is understood to reflect the number of items held in memory (Vogel & Machizawa, 2004) but it may also reflect the quality of the representation (Luria et al., 2016). CDA amplitudes in Study 2 were found to be modulated by contextual regularities in the display. For example, in Experiment 2, the CDA was reduced in the repeated (consistent) configurations compared to the novel ones, despite identical set sizes. This reduction suggests that fewer resources were needed to maintain the same number of items when repeated configurations could guide encoding. Rather than reflecting a drop in stored items in the repeated configuration condition, the reduced CDA may indicate less effortful maintenance due to reduced competition. The findings from Study 2 also have important implications for research that emphasises the context-sensitive nature of vWM. For instance, studies on statistical learning suggest that vWM can encode structured environments better by exploiting spatial or temporal regularities (Brady & Oliva, 2008; Brady et al., 2009). Alike Study 2, this process is often implicit, thus participants are unaware of repeated configurations yet still show improved behavioural and neural performance.

Together the findings of Study 2, like those of Study 1, move away from slot-based approaches and support more hybrid (Zhang & Luck, 2008) and interference-based models (Oberauer & Kliegl, 2001, 2006). Capacity here appears to be shaped by environmental structure and prior learning. Thus, when input is more predictable, vWM capacity improves and fewer resources are needed to maintain the information, as shown by the reduced CDA amplitude. Overall, the findings of Study 2 stress that capacity is sensitive to regularities in the environment, supporting the idea that vWM operates as a context-sensitive and adaptive system.

Shifting from an external focus to an internal focus, Study 3 provided important theoretical insight into the role of attentional control in individual differences in vWM capacity. The shift from external contextual factors in Study 2 to internal control processes in Study 3 allowed us to understand a more comprehensive view of vWM as shaped not only by the structure of the information processed but by the cognitive systems ability to manage that information.

The findings from Study 3 are consistent with Unsworth and Engle's (2007) dual-store model. This model argues that vWM capacity reflects both the ability to maintain information in primary memory (a capacity limited attentional store) and the efficient search and retrieval of information from secondary memory. Both abilities play a vital role in active maintenance and retrieval of goal-related information, and it is suggested that individual differences in vWM capacity mainly stem from fluctuations in these two abilities. Thus, high-K individuals outperform low-K individuals not because they store more items, but because they are more effective at maintaining goal relevant information and resisting interference from irrelevant representations. The results from Study 3 match this account: high-K participants showed stronger neural indices of attentional deployment (N2pc) and WM maintenance (CDA). This supports the idea that attentional control, not just storage, underlies effective vWM performance. Furthermore, as shown by Engle and colleagues (Ilkowska & Engle, 2010; Kane & Engle, 2003), individuals with higher vWM capacity are better at executing top-down control, particularly in complex tasks. Thus, in the light of

Study 3’s findings, this may have translated into a better ability to maintain three target templates in mind.

Following the same pattern as Study 1 and Study 2, the findings from Study 3 are also compatible with hybrid models of capacity (Zhang & Luck, 2008). Participants who demonstrate larger attentional control may be able to flexibly allocate resources where needed, resulting in stronger encoding and more efficient search performance. Likewise, interference-based models (Oberauer & Kliegl, 2006) can also offer insight into these findings. Perhaps high-K individuals may be more efficient at resolving competition between overlapping representations, particularly when target features must be maintained and then retrieved amongst distractor items.

Importantly, the findings from Study 3 help to reassess the concept of capacity in the broader attentional system. Rather than being a fixed limit, vWM capacity appears to reflect successful coordination between attention and memory, specifically regarding attentional templates, supported by the converging evidence from behavioural measures (K) and neural indices (CDA and N2pc components). Overall, these findings underpin the need for theoretical models of vWM to account for both storage and selection processes.

Altogether, the findings from this thesis support a multi-factorial view of vWM capacity. Whilst discrete-slot models capture a limit of around 3-4 items, they are insufficient to explain variability produced by perceptual parameters, learned regularities or the guidance of attention. Therefore, the present thesis adopts hybrid models as a more comprehensive account, capable of accommodating the range of findings observed across all three studies (Davelaar et al., 2005; Oberauer & Kliegl, 2001, 2006; Zhang & Luck, 2008). Hybrid models are well suited to capture this complexity in that they incorporate multiple mechanisms beyond just a fixed item limit. These include slot-based precision for a limited number of high-quality representations, flexible resource allocation based on item relevance, interference dynamics that constrain encoding and retrieval, and attentional control mechanisms. As opposed to conflicting one model over another, this thesis emphasises the importance of integrative approaches that account for both structural and functional capacity. K serves as an effective behavioural index, but its interpretation is enhanced by combining it with neural data and experimental manipulations that probe encoding strategies, attentional selection, and learning effects.

5.3 Limitations and future directions

While the findings of this thesis offer valuable insights into the flexible and dynamic nature of vWM capacity, several caveats must be acknowledged as well as potential avenues for future research.

While K is widely adopted for estimating the number of items held in vWM, it is built on the assumption of slot-based storage and thus struggles to account for fluctuations in encoding precision, implicit learning or attentional lapses, as presented within the three studies in the current thesis. Alternative methods of estimating capacity, such as the 'fixed-capacity + attention model' (Rouder et al., 2008), add an attentional lapse parameter to estimates of vWM capacity. Adding an attentional lapse parameter accounts for trials in which participants are inattentive to the task. Participants often make errors on trials that should be well within their capacity limits (e.g., set size 1) and this addition of a lapse parameter can help explain these dips in performance. Unlike the method of K, where a K value is computed for each set size, this model uses a log-likelihood estimation technique that estimates a single capacity parameter by considering performance across all set sizes and change probability conditions. Importantly, this model assumes that data are obtained for at least one sub-capacity set size and errors made on this set size would reflect an attentional lapse. If the model is fit to data that do not contain at least one sub-capacity low set size (e.g., set size 1 or 2), then the model will provide invalid parameter estimates. For instance, Van Snellenberg et al. (2014) applied this model to calculate capacity for a change detection task and reported low reliability in the estimated capacity parameter, which also failed to correlate with other WM tasks. This outcome was possibly due to their high set sizes (4 and 8), resulting in poor model fit. In contrast, Study 3 of the present thesis did include lower set sizes (1, 2 and 3), making it especially suitable for applying such models. Applying the 'fixed capacity + attention' framework here could have helped distinguish true capacity limitations from attentional disengagement. This would have been particularly valuable given the studies focus on the relationship between vWM capacity and attentional control.

Another important consideration for this thesis is the role of experience, through either explicit practice or implicit learning, in shaping vWM capacity. Research shows that practice can free up WM resources by making responses more automatic, thus leaving more resources for complex processing (Beilock & DeCaro, 2007; Beilock et al., 2007; Chein & Schneider, 2005). Particularly under high load, practice can reduce susceptibility to distraction, a benefit for individuals specifically more prone to interference (Forster & Lavie, 2007). This is particularly relevant considering the ongoing popularity of whether training affects WM capacity (Melby-Lervag & Hulme, 2013; Shipstead et al., 2012). This interpretation is supported by Xu et al. (2018), who looked at how consistent estimates of capacity are from day to day. They collected thirty-one sessions of change detection performance spanning 60 days and found that while participants' relative K values remained consistent across sessions, absolute performance slightly improved with practice. These findings suggest that while absolute performance can improve with repeated task exposure, the measurement of individual differences in capacity remains robust over time. More specifically, the results highlight the role of experience in optimising vWM performance, even without explicit strategy.

Similar to this broader view, Study 2 provides a complementary example of how prior exposure through implicit learning can enhance vWM. Although participants in Study 2 did not specifically undergo extended training, repeated exposure to spatial configurations led to improved performance for those displays. However, one specific caveat of Study 2 was the inability to assess when implicit learning of repeated configurations occurred. Repeated and novel trials were analysed across the entire task without dividing performance into epochs, making it difficult to determine at what point participants began to benefit from the regularities. Incorporating epoch-based analysis would have allowed a more specific examination of performance over time and determine how quickly these repeated configurations were learnt. Together the findings from long-term practice studies and Study 2 stress the importance that vWM capacity is adaptively modulated by familiarity and contextual predictability.

5.4 Conclusion

This thesis set out to clarify the nature of vWM capacity as both a stable, trait-like individual difference and a flexible, state-like system shaped by environmental and cognitive context. Whilst K remains a widely used and robust measure of how many items can be held in vWM capacity, findings across all three studies demonstrate that vWM capacity is not fixed. Instead, it is shaped by perceptual parameters (Study 1), implicit learning (Study 2), and it is related to attentional control (Study 3). These results extend beyond traditional slot-based models and support more hybrid assumptions which reflect the dynamic nature of vWM capacity that affects how visual information is selected, structured, and prioritised. Thus, future models should move beyond viewing capacity as fixed and instead look at how performance reflects both underlying constraints and the flexible use of cognitive resources.

References

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, *106*(1), 20–29. https://doi.org/10.1016/j.jecp.2009.11.003

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

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences*, *108*(25), 10367–10371. https://doi.org/10.1073/pnas.1104047108

Annac, E., Conci, M., Müller, H. J., & Geyer, T. (2017). Local item density modulates adaptation of learned contextual cues. *Visual Cognition*, *25*(1–3), 262–277. https://doi.org/10.1080/13506285.2017.1339158

Annac, E., Manginelli, A. A., Pollmann, S., Shi, Z., Muller, H. J., & Geyer, T. (2013). Memory under pressure: Secondary-task effects on contextual cueing of visual search. *Journal of Vision*, *13*(13), 6–6. https://doi.org/10.1167/13.13.6

Aspanani, A., Sadeqhi, H., & Omid, A. (2023). The relationship between visual memory and spatial intelligence with students' academic achievement in anatomy. *BMC Medical Education*, *23*(1), 336. https://doi.org/10.1186/s12909-023-04327-9

Atkinson, R. C., & Shiffrin, R. M. (1968). Human Memory: A Proposed System and its Control Processes1. In K. W. Spence & J. T. Spence (Eds), *Psychology of Learning and Motivation* (Vol. 2, pp. 89–195). Academic Press. https://doi.org/10.1016/S0079-7421(08)60422-3

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual Working Memory Represents a Fixed Number of Items Regardless of Complexity. *Psychological Science*, *18*(7), 622–628. https://doi.org/10.1111/j.1467-9280.2007.01949.x

Awh, E., Vogel, E. K., & Oh, S.-H. (2006). Interactions between attention and working memory. *Neuroscience*, *139*(1), 201–208. https://doi.org/10.1016/j.neuroscience.2005.08.023

Baddeley, A. (1986). *Working memory* (pp. xi, 289). Clarendon Press/Oxford University Press.

Baddeley, A. (1996). The fractionation of working memory. *Proceedings of the National Academy of Sciences*, *93*(24), 13468–13472. https://doi.org/10.1073/pnas.93.24.13468

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423. https://doi.org/10.1016/S1364-6613(00)01538-2

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*(3), 189–208. https://doi.org/10.1016/S0021-9924(03)00019-4

Baddeley, A. (2012). Working Memory: Theories, Models, and Controversies. *Annual Review of Psychology*, *63*(Volume 63, 2012), 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

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

Balaban, H., Fukuda, K., & Luria, R. (2019). What can half a million change detection trials tell us about visual working memory? *Cognition*, *191*, 103984. https://doi.org/10.1016/j.cognition.2019.05.021

Balaban, H., & Luria, R. (2016). Integration of Distinct Objects in Visual Working Memory Depends on Strong Objecthood Cues Even for Different-Dimension Conjunctions. *Cerebral Cortex*, *26*(5), 2093–2104. https://doi.org/10.1093/cercor/bhv038

Barlow, H. B. (1972). Single Units and Sensation: A Neuron Doctrine for Perceptual Psychology? *Perception*, *1*(4), 371–394. https://doi.org/10.1068/p010371

Barrett, D. J. K., & Zobay, O. (2014). Attentional Control via Parallel Target-Templates in Dual-Target Search. *PLOS ONE*, *9*(1), e86848. https://doi.org/10.1371/journal.pone.0086848

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

Beck, D. M., & Lavie, N. (2005). Look Here but Ignore What You See: Effects of Distractors at Fixation. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(3), 592–607. https://doi.org/10.1037/0096-1523.31.3.592

Beck, V. M., Hollingworth, A., & Luck, S. J. (2012). Simultaneous Control of Attention by Multiple Working Memory Representations. *Psychological Science*, *23*(8), 887–898. https://doi.org/10.1177/0956797612439068

Beesley, T., Vadillo, M. A., Pearson, D., & Shanks, D. R. (2015). Pre-exposure of repeated search configurations facilitates subsequent contextual cuing of visual search. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(2), 348–362. https://doi.org/10.1037/xlm0000033

Beilock, S. L., & DeCaro, M. S. (2007). From poor performance to success under stress: Working memory, strategy selection, and mathematical problem solving under pressure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(6), 983–998. https://doi.org/10.1037/0278-7393.33.6.983

Beilock, S. L., Rydell, R. J., & McConnell, A. R. (2007). Stereotype threat and working memory: Mechanisms, alleviation, and spillover. *Journal of Experimental Psychology: General*, *136*(2), 256–276. https://doi.org/10.1037/0096-3445.136.2.256

Berggren, N., & Eimer, M. (2018). Object-based target templates guide attention during visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(9), 1368–1382. https://doi.org/10.1037/xhp0000541

Berggren, N., & Eimer, M. (2019). The roles of relevance and expectation for the control of attention in visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(9), 1191–1205. https://doi.org/10.1037/xhp0000666

Berggren, N., & Eimer, M. (2020). Attentional Access to Multiple Target Objects in Visual Search. *Journal of Cognitive Neuroscience*, *32*(2), 283–300. https://doi.org/10.1162/jocn_a_01476

Berggren, N., Jenkins, M., McCants, C. W., & Eimer, M. (2017). The spatially global control of attentional target selection in visual search. *Visual Cognition*, *25*(1–3), 196–214. https://doi.org/10.1080/13506285.2017.1287806

Brady, T. F., & Alvarez, G. A. (2011). Hierarchical Encoding in Visual Working Memory: Ensemble Statistics Bias Memory for Individual Items. *Psychological Science*, *22*(3), 384–392. https://doi.org/10.1177/0956797610397956

Brady, T. F., Konkle, T., & Alvarez, G. A. (2009). Compression in visual working memory: Using statistical regularities to form more efficient memory representations. *Journal of Experimental Psychology: General*, *138*(4), 487–502. https://doi.org/10.1037/a0016797

Brady, T. F., & Oliva, A. (2008). Statistical Learning Using Real-World Scenes: Extracting Categorical Regularities Without Conscious Intent. *Psychological Science*, *19*(7), 678–685. https://doi.org/10.1111/j.1467-9280.2008.02142.x

Broadbent, D. E. (1958). *Perception and communication* (pp. v, 340). Pergamon Press. https://doi.org/10.1037/10037-000

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, *97*(4), 523–547. https://doi.org/10.1037/0033-295X.97.4.523

Campbell, J. I. D., & Thompson, V. A. (2012). MorePower 6.0 for ANOVA with relational confidence intervals and Bayesian analysis. *Behavior Research Methods*, *44*(4), 1255–1265. https://doi.org/10.3758/s13428-012-0186-0

Capon, A., Handley, S., & Dennis, I. (2003). Working memory and reasoning: An individual differences perspective. *Thinking & Reasoning*, *9*(3), 203–244. https://doi.org/10.1080/13546781343000222

Carlisle, N. B., Arita, J. T., Pardo, D., & Woodman, G. F. (2011). Attentional templates in visual working memory. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(25), 9315–9322. https://doi.org/10.1523/JNEUROSCI.1097-11.2011

Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences*, *9*(7), 349–354. https://doi.org/10.1016/j.tics.2005.05.009

Chein, J. M., & Schneider, W. (2005). Neuroimaging studies of practice-related change: fMRI and meta-analytic evidence of a domain-general control network for learning. *Cognitive Brain Research*, *25*(3), 607–623. https://doi.org/10.1016/j.cogbrainres.2005.08.013

Chelazzi, L., Duncan, J., Miller, E. K., & Desimone, R. (1998). Responses of Neurons in Inferior Temporal Cortex During Memory-Guided Visual Search. *Journal of Neurophysiology*, *80*(6), 2918–2940. https://doi.org/10.1152/jn.1998.80.6.2918

Chicherov, V., Plomp, G., & Herzog, M. H. (2014). Neural correlates of visual crowding. *NeuroImage*, *93*, 23–31. https://doi.org/10.1016/j.neuroimage.2014.02.021

Chong, S. C., & Treisman, A. (2005a). Attentional spread in the statistical processing of visual displays. *Perception & Psychophysics*, *67*(1), 1–13. https://doi.org/10.3758/BF03195009

Chong, S. C., & Treisman, A. (2005b). Statistical processing: Computing the average size in perceptual groups. *Vision Research*, *45*(7), 891–900. https://doi.org/10.1016/j.visres.2004.10.004

Christie, G. J., Livingstone, A. C., & McDonald, J. J. (2015). Searching for Inefficiency in Visual Search. *Journal of Cognitive Neuroscience*, *27*(1), 46–56. https://doi.org/10.1162/jocn_a_00716

Chua, K.-P., & Chun, M. M. (2003). Implicit scene learning is viewpoint dependent. *Perception & Psychophysics*, *65*(1), 72–80. https://doi.org/10.3758/BF03194784

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences*, *4*(5), 170–178. https://doi.org/10.1016/S1364-6613(00)01476-5

Chun, M. M., & Jiang, Y. (1998). Contextual Cueing: Implicit Learning and Memory of Visual Context Guides Spatial Attention. *Cognitive Psychology*, *36*(1), 28–71. https://doi.org/10.1006/cogp.1998.0681

Chun, M. M., & Jiang, Y. (1999). Top-Down Attentional Guidance Based on Implicit Learning of Visual Covariation. *Psychological Science*, *10*(4), 360–365. https://doi.org/10.1111/1467-9280.00168

Chun, M. M., & Turk-Browne, N. B. (2008). Associative Learning Mechanisms in Vision. In S. J. Luck & A. Hollingworth (Eds), *Visual Memory* (pp. 209–246). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195305487.003.0007

Chung, S. T. L. (2004). Reading Speed Benefits from Increased Vertical Word Spacing in Normal Peripheral Vision. *Optometry and Vision Science : Official Publication of the American Academy of Optometry*, *81*(7), 525–535. https://doi.org/10.1097/00006324-200407000-00014

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed). Earlbaum.

Colagiuri, B., & Livesey, E. J. (2016). Contextual cuing as a form of nonconscious learning: Theoretical and empirical analysis in large and very large samples. *Psychonomic Bulletin & Review*, *23*(6), 1996–2009. https://doi.org/10.3758/s13423-016-1063-0

Colom, R., Abad, F. J., Quiroga, M. Á., Shih, P. C., & Flores-Mendoza, C. (2008). Working memory and intelligence are highly related constructs, but why? *Intelligence*, *36*(6), 584–606. https://doi.org/10.1016/j.intell.2008.01.002

Conci, M., & Müller, H. J. (2012). Contextual learning of multiple target locations in visual search. *Visual Cognition*, *20*(7), 746–770. https://doi.org/10.1080/13506285.2012.694376

Conci, M., Sun, L., & Müller, H. J. (2011). Contextual remapping in visual search after predictable target-location changes. *Psychological Research*, *75*(4), 279–289. https://doi.org/10.1007/s00426-010-0306-3

Connolly, M., & Van Essen, D. (1984). The representation of the visual field in parvicellular and magnocellular layers of the lateral geniculate nucleus in the macaque monkey. *Journal of Comparative Neurology*, *226*(4), 544–564. https://doi.org/10.1002/cne.902260408

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, *104*(2), 163–191. https://doi.org/10.1037/0033-2909.104.2.163

Cowan, N. (1999). An Embedded-Processes Model of Working Memory. In A. Miyake & P. Shah (Eds), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 62–101). Cambridge University Press. https://doi.org/10.1017/CBO9781139174909.006

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–114. https://doi.org/10.1017/S0140525X01003922

Cowan, N. (2008). Chapter 20 What are the differences between long-term, short-term, and working memory? In W. S. Sossin, J.-C. Lacaille, V. F. Castellucci, & S. Belleville (Eds), *Progress in Brain Research* (Vol. 169, pp. 323–338). Elsevier. https://doi.org/10.1016/S0079-6123(07)00020-9

Cowan, N. (2010). The Magical Mystery Four: How Is Working Memory Capacity Limited, and Why? *Current Directions in Psychological Science*, *19*(1), 51–57. https://doi.org/10.1177/0963721409359277

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review*, *24*(4), 1158–1170. https://doi.org/10.3758/s13423-016-1191-6

Cowan, N., Elliott, E. M., Scott Saults, J., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, *51*(1), 42–100. https://doi.org/10.1016/j.cogpsych.2004.12.001

Cunningham, C. A., & Wolfe, J. M. (2014). The role of object categories in hybrid visual and memory search. *Journal of Experimental Psychology: General*, *143*(4), 1585–1599. https://doi.org/10.1037/a0036313

Curcio, C. A., & Allen, K. A. (1990). Topography of ganglion cells in human retina. *Journal of Comparative Neurology*, *300*(1), 5–25. https://doi.org/10.1002/cne.903000103

Curcio, C. A., Sloan, K. R., Packer, O., Hendrickson, A. E., & Kalina, R. E. (1987). Distribution of Cones in Human and Monkey Retina: Individual Variability and Radial Asymmetry. *Science*, *236*(4801), 579–582. https://doi.org/10.1126/science.3576186

Dai, M., Li, Y., Gan, S., & Du, F. (2019). The reliability of estimating visual working memory capacity. *Scientific Reports*, *9*(1), 1155. https://doi.org/10.1038/s41598-019-39044-1

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466. https://doi.org/10.1016/S0022-5371(80)90312-6

Davelaar, E. J., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H. J., & Usher, M. (2005). The Demise of Short-Term Memory Revisited: Empirical and Computational Investigations of Recency Effects. *Psychological Review*, *112*(1), 3–42. https://doi.org/10.1037/0033-295X.112.1.3

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222. https://doi.org/10.1146/annurev.ne.18.030195.001205

Drew, T., Horowitz, T. S., Wolfe, J. M., & Vogel, E. K. (2011). Delineating the Neural Signatures of Tracking Spatial Position and Working Memory during Attentive Tracking. *The Journal of Neuroscience*, *31*(2), 659–668. https://doi.org/10.1523/JNEUROSCI.1339-10.2011

Duncan, J., & Humphreys, G. (1992). Beyond the search surface: Visual search and attentional engagement. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(2), 578–588. https://doi.org/10.1037/0096-1523.18.2.578

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*(3), 433–458. https://doi.org/10.1037/0033-295X.96.3.433

Duncan, R. O., & Boynton, G. M. (2003). Cortical Magnification within Human Primary Visual Cortex Correlates with Acuity Thresholds. *Neuron*, *38*(4), 659–671. https://doi.org/10.1016/S0896-6273(03)00265-4

Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalography and Clinical Neurophysiology*, *99*(3), 225–234. https://doi.org/10.1016/0013-4694(96)95711-9

Eimer, M. (2014). The neural basis of attentional control in visual search. *Trends in Cognitive Sciences*, *18*(10), 526–535. https://doi.org/10.1016/j.tics.2014.05.005

Eimer, M., & Grubert, A. (2014). The gradual emergence of spatially selective target processing in visual search: From feature-specific to object-based attentional control. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(5), 1819–1831. https://doi.org/10.1037/a0037387

Emrich, S. M., Riggall, A. C., LaRocque, J. J., & Postle, B. R. (2013). Distributed Patterns of Activity in Sensory Cortex Reflect the Precision of Multiple Items Maintained in Visual Short-Term Memory. *The Journal of Neuroscience*, *33*(15), 6516–6523. https://doi.org/10.1523/JNEUROSCI.5732-12.2013

Eng, H. Y., Chen, D., & Jiang, Y. (2005). Visual working memory for simple and complex visual stimuli. *Psychonomic Bulletin & Review*, *12*(6), 1127–1133. https://doi.org/10.3758/BF03206454

Engel de Abreu, P. M. J., Conway, A. R. A., & Gathercole, S. E. (2010). Working memory and fluid intelligence in young children. *Intelligence*, *38*(6), 552–561. https://doi.org/10.1016/j.intell.2010.07.003

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*(3), 309–331. https://doi.org/10.1037/0096-3445.128.3.309

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/bf03193146

Feldmann-Wüstefeld, T. (2021). Neural measures of working memory in a bilateral change detection task. *Psychophysiology*, *58*(1), e13683. https://doi.org/10.1111/psyp.13683

Forster, S., & Lavie, N. (2007). High Perceptual Load Makes Everybody Equal. *Psychological Science*, *18*(5), 377–381. https://doi.org/10.1111/j.1467-9280.2007.01908.x

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

Fukuda, K., & Vogel, E. K. (2009). Human Variation in Overriding Attentional Capture. *Journal of Neuroscience*, *29*(27), 8726–8733. https://doi.org/10.1523/JNEUROSCI.2145-09.2009

Fukuda, K., & Vogel, E. K. (2011). Individual Differences in Recovery Time From Attentional Capture. *Psychological Science*, *22*(3), 361–368. https://doi.org/10.1177/0956797611398493

Fukuda, K., Vogel, E., Mayr, U., & Awh, E. (2010b). Quantity, not quality: The relationship between fluid intelligence and working memory capacity. *Psychonomic Bulletin & Review*, *17*(5), 673–679. https://doi.org/10.3758/17.5.673

Fukuda, K., Woodman, G. F., & Vogel, E. K. (2015). Individual Differences in Visual Working Memory Capacity. In *Mechanisms of Sensory Working Memory* (pp. 105–119). Elsevier. https://doi.org/10.1016/B978-0-12-801371-7.00009-0

Gaspar, J. M., Christie, G. J., Prime, D. J., Jolicœur, P., & McDonald, J. J. (2016). Inability to suppress salient distractors predicts low visual working memory capacity. *Proceedings of the National Academy of Sciences*, *113*(13), 3693–3698. https://doi.org/10.1073/pnas.1523471113

Geyer, T., Zehetleitner, M., & Müller, H. J. (2010). Contextual cueing of pop-out visual search: When context guides the deployment of attention. *Journal of Vision*, *10*(5), 20. https://doi.org/10.1167/10.5.20

Gong, M., Liu, T., Liu, X., Huangfu, B., & Geng, F. (2023). Attention relieves visual crowding: Dissociable effects of peripheral and central cues. *Journal of Vision*, *23*(5), 9. https://doi.org/10.1167/jov.23.5.9

Goujon, A., Didierjean, A., & Thorpe, S. (2015). Investigating implicit statistical learning mechanisms through contextual cueing. *Trends in Cognitive Sciences*, *19*(9), 524–533. https://doi.org/10.1016/j.tics.2015.07.009

Green, D. G. (1970). Regional variations in the visual acuity for interference fringes on the retina. *The Journal of Physiology*, *207*(2), 351–356. https://doi.org/10.1113/jphysiol.1970.sp009065

Greene, A. J., Gross, W. L., Elsinger, C. L., & Rao, S. M. (2007). Hippocampal differentiation without recognition: An fMRI analysis of the contextual cueing task. *Learning & Memory*, *14*(8), 548–553. https://doi.org/10.1101/lm.609807

Greenwood, J. A., & Parsons, M. J. (2020). Dissociable effects of visual crowding on the perception of color and motion. *Proceedings of the National Academy of Sciences*, *117*(14), 8196–8202. https://doi.org/10.1073/pnas.1909011117

Grubert, A., Carlisle, N. B., & Eimer, M. (2016). The Control of Single-color and Multiple-color Visual Search by Attentional Templates in Working Memory and in Long-term Memory. *Journal of Cognitive Neuroscience*, *28*(12), 1947–1963. https://doi.org/10.1162/jocn_a_01020

Grubert, A., & Eimer, M. (2016a). All set, indeed! N2pc components reveal simultaneous attentional control settings for multiple target colors. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(8), 1215–1230. https://doi.org/10.1037/xhp0000221

Grubert, A., & Eimer, M. (2016b). Rapid attentional selection processes operate independently and in parallel for multiple targets. *Biological Psychology*, *121*, 99–108. https://doi.org/10.1016/j.biopsycho.2016.10.012

Grubert, A., & Eimer, M. (2018). The Time Course of Target Template Activation Processes during Preparation for Visual Search. *Journal of Neuroscience*, *38*(44), 9527–9538. https://doi.org/10.1523/JNEUROSCI.0409-18.2018

Grubert, A., & Eimer, M. (2020). Preparatory Template Activation during Search for Alternating Targets. *Journal of Cognitive Neuroscience*, *32*(8), 1525–1535. https://doi.org/10.1162/jocn_a_01565

Grubert, A., & Eimer, M. (2023). Do We Prepare for What We Predict? How Target Expectations Affect Preparatory Attentional Templates and Target Selection in Visual Search. *Journal of Cognitive Neuroscience*, *35*(12), 1919–1935. https://doi.org/10.1162/jocn_a_02054

Grubert, A., Eimer, M., Wang, Z., & Jimenez, M. (In press). The guidance of attentional selectivity in visual search is always feature-based: Behavioural and electrophysiological evidence from feature and conjunction search tasks.

Grubert, A., Wang, Z., Williams, E., Jimenez, M., Remington, R., & Eimer, M. (2025). The capacity limitations of multiple-template visual search during task preparation and target selection. *Psychophysiology*, *62*(1), e14720. https://doi.org/10.1111/psyp.14720

Gunseli, E., Meeter, M., & Olivers, C. N. L. (2014). Is a search template an ordinary working memory? Comparing electrophysiological markers of working memory maintenance for visual search and recognition. *Neuropsychologia*, *60*, 29–38. https://doi.org/10.1016/j.neuropsychologia.2014.05.012

Gunseli, E., Olivers, C. N. L., & Meeter, M. (2014). Effects of Search Difficulty on the Selection, Maintenance, and Learning of Attentional Templates. *Journal of Cognitive Neuroscience*, *26*(9), 2042–2054. https://doi.org/10.1162/jocn_a_00600

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

Herzog, M. H., & Manassi, M. (2015). Uncorking the bottleneck of crowding: A fresh look at object recognition. *Current Opinion in Behavioral Sciences*, *1*, 86–93. https://doi.org/10.1016/j.cobeha.2014.10.006

Houtkamp, R., & Roelfsema, P. R. (2009). Matching of visual input to only one item at any one time. *Psychological Research PRPF*, *73*(3), 317–326. https://doi.org/10.1007/s00426-008-0157-3

Huang, L., & Pashler, H. (2007). A Boolean map theory of visual attention. *Psychological Review*, *114*(3), 599–631. https://doi.org/10.1037/0033-295X.114.3.599

Ikkai, A., McCollough, A. W., & Vogel, E. K. (2010). Contralateral Delay Activity Provides a Neural Measure of the Number of Representations in Visual Working Memory. *Journal of Neurophysiology*, *103*(4), 1963–1968. https://doi.org/10.1152/jn.00978.2009

Ilkowska, M., & Engle, R. W. (2010). Trait and State Differences in Working Memory Capacity. In A. Gruszka, G. Matthews, & B. Szymura (Eds), *Handbook of Individual Differences in Cognition: Attention, Memory, and Executive Control* (pp. 295–320). Springer. https://doi.org/10.1007/978-1-4419-1210-7_18

Irons, J. L., Folk, C. L., & Remington, R. W. (2012). All set! Evidence of simultaneous attentional control settings for multiple target colors. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(3), 758–775. https://doi.org/10.1037/a0026578

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, *2*(3), 194–203. https://doi.org/10.1038/35058500

Jackson, M. C., Linden, D. E. J., Roberts, M. V., Kriegeskorte, N., & Haenschel, C. (2015). Similarity, not complexity, determines visual working memory performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(6), 1884–1892. https://doi.org/10.1037/xlm0000125

Janacsek, K., & Nemeth, D. (2013). Implicit sequence learning and working memory: Correlated or complicated? *Cortex*, *49*(8), 2001–2006. https://doi.org/10.1016/j.cortex.2013.02.012

Jiang, Y., & Chun, M. M. (2001). Selective attention modulates implicit learning. *The Quarterly Journal of Experimental Psychology Section A*, *54*(4), 1105–1124. https://doi.org/10.1080/713756001

Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(3), 683–702. https://doi.org/10.1037/0278-7393.26.3.683

Jiang, Y., & Song, J.-H. (2005). Spatial context learning in visual search and change detection. *Perception & Psychophysics*, *67*(7), 1128–1139. https://doi.org/10.3758/BF03193546

Jiang, Y., & Wagner, L. C. (2004). What is learned in spatial contextual cuing—Configuration or individual locations? *Perception & Psychophysics*, *66*(3), 454–463. https://doi.org/10.3758/BF03194893

Jimenez, M., Wang, Z., & Grubert, A. (2024). Attentional templates for target features versus locations. *Scientific Reports*, *14*(1), 22306. https://doi.org/10.1038/s41598-024-73656-6

Johnson, J. S., Woodman, G. F., Braun, E., & Luck, S. J. (2007). Implicit memory influences the allocation of attention in visual cortex. *Psychonomic Bulletin & Review*, *14*(5), 834–839. https://doi.org/10.3758/BF03194108

Kałamała, P., Sadowska, A., Ordziniak, W., & Chuderski, A. (2017). Gestalt Effects in Visual Working Memory. *Experimental Psychology*, *64*(1), 5–13. https://doi.org/10.1027/1618-3169/a000346

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, *132*(1), 47–70. https://doi.org/10.1037/0096-3445.132.1.47

Kane, M. J., Poole, B. J., Tuholski, S. W., & Engle, R. W. (2006). Working memory capacity and the top-down control of visual search: Exploring the boundaries of 'executive attention'. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(4), 749–777. https://doi.org/10.1037/0278-7393.32.4.749

Kerzel, D., & Grubert, A. (2022). Capacity limitations in template-guided multiple color search. *Psychonomic Bulletin & Review*, *29*(3), 901–909. https://doi.org/10.3758/s13423-021-02040-6

Kojouharova, P., Nagy, B., Czigler, I., & Gaál, Z. A. (2023). Mechanisms of spatial contextual cueing in younger and older adults. *Psychophysiology*, *60*(11), e14361. https://doi.org/10.1111/psyp.14361

Kristjánsson, T., & Kristjánsson, Á. (2018). Foraging through multiple target categories reveals the flexibility of visual working memory. *Acta Psychologica*, *183*, 108–115. https://doi.org/10.1016/j.actpsy.2017.12.005

Kyttälä, M., & Lehto, J. E. (2008). Some factors underlying mathematical performance: The role of visuospatial working memory and non-verbal intelligence. *European Journal of Psychology of Education*, *23*(1), 77–94. https://doi.org/10.1007/BF03173141

Lafer-Sousa, R., Conway, B. R., & Kanwisher, N. G. (2016). Color-Biased Regions of the Ventral Visual Pathway Lie between Face- and Place-Selective Regions in Humans, as in Macaques. *Journal of Neuroscience*, *36*(5), 1682–1697. https://doi.org/10.1523/JNEUROSCI.3164-15.2016

Larson, A. M., & Loschky, L. C. (2009). The contributions of central versus peripheral vision to scene gist recognition. *Journal of Vision*, *9*(10), 6. https://doi.org/10.1167/9.10.6

Lee, J., Leonard, C. J., Luck, S. J., & Geng, J. J. (2018). Dynamics of Feature-based Attentional Selection during Color–Shape Conjunction Search. *Journal of Cognitive Neuroscience*, *30*(12), 1773–1787. https://doi.org/10.1162/jocn_a_01318

Levi, D. M., Hariharan, S., & Klein, S. A. (2002). Suppressive and facilitatory spatial interactions in peripheral vision: Peripheral crowding is neither size invariant nor simple contrast masking. *Journal of Vision*, *2*(2), 3. https://doi.org/10.1167/2.2.3

Li, J., Qian, J., & Liang, F. (2018). Evidence for the beneficial effect of perceptual grouping on visual working memory: An empirical study on illusory contour and a meta-analytic study. *Scientific Reports*, *8*(1), 13864. https://doi.org/10.1038/s41598-018-32039-4

Li, Q., & Saiki, J. (2015). Different effects of color-based and location-based selection on visual working memory. *Attention, Perception, & Psychophysics*, *77*(2), 450–463. https://doi.org/10.3758/s13414-014-0775-3

Logie, R. H., Brockmole, J. R., & Jaswal, S. (2011). Feature binding in visual short-term memory is unaffected by task-irrelevant changes of location, shape, and color. *Memory & Cognition*, *39*(1), 24–36. https://doi.org/10.3758/s13421-010-0001-z

Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology*, *54*(1), 146–157. https://doi.org/10.1111/psyp.12639

Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, *31*(3), 291–308. https://doi.org/10.1111/j.1469-8986.1994.tb02218.x

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

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

Luria, R., Balaban, H., Awh, E., & Vogel, E. K. (2016). The contralateral delay activity as a neural measure of visual working memory. *Neuroscience & Biobehavioral Reviews*, *62*, 100–108. https://doi.org/10.1016/j.neubiorev.2016.01.003

Luria, R., & Vogel, E. K. (2011). Visual Search Demands Dictate Reliance on Working Memory Storage. *Journal of Neuroscience*, *31*(16), 6199–6207. https://doi.org/10.1523/JNEUROSCI.6453-10.2011

Luria, R., & Vogel, E. K. (2014). Come Together, Right Now: Dynamic Overwriting of an Object's History through Common Fate. *Journal of Cognitive Neuroscience*, *26*(8), 1819–1828. https://doi.org/10.1162/jocn_a_00584

Makovski, T., & Jiang, Y. V. (2010). Contextual cost: When a visual-search target is not where it should be. *Quarterly Journal of Experimental Psychology*, *63*(2), 216–225. https://doi.org/10.1080/17470210903281590

Mance, I., & Vogel, E. K. (2013). Visual working memory. *WIREs Cognitive Science*, *4*(2), 179–190. https://doi.org/10.1002/wcs.1219

Manginelli, A. A., Geringswald, F., & Pollmann, S. (2012). Visual Search Facilitation in Repeated Displays Depends on Visuospatial Working Memory. *Experimental Psychology*, *59*(1), 47–54. https://doi.org/10.1027/1618-3169/a000125

Manginelli, A. A., Langer, N., Klose, D., & Pollmann, S. (2013). Contextual cueing under working memory load: Selective interference of visuospatial load with expression of learning. *Attention, Perception, & Psychophysics*, *75*(6), 1103–1117. https://doi.org/10.3758/s13414-013-0466-5

Manginelli, A. A., & Pollmann, S. (2009). Misleading contextual cues: How do they affect visual search? *Psychological Research*, *73*(2), 212–221. https://doi.org/10.1007/s00426-008-0211-1

Maor, R., Paz-Baruch, N., Grinshpan, N., Milman, A., Mevarech, Z., Levi, R., Shlomo, S., & Zion, M. (2023). Relationships between metacognition, creativity, and critical thinking in self-reported teaching performances in project-based learning settings. *Thinking Skills and Creativity*, *50*, 101425. https://doi.org/10.1016/j.tsc.2023.101425

Martelli, M., Majaj, N. J., & Pelli, D. G. (2005). Are faces processed like words? A diagnostic test for recognition by parts. *Journal of Vision*, *5*(1), 6. https://doi.org/10.1167/5.1.6

Martinez-Trujillo, J. C., & Treue, S. (2004). Feature-Based Attention Increases the Selectivity of Population Responses in Primate Visual Cortex. *Current Biology*, *14*(9), 744–751. https://doi.org/10.1016/j.cub.2004.04.028

Martini, M., Furtner, M. R., & Sachse, P. (2013). Working Memory and Its Relation to Deterministic Sequence Learning. *PLOS ONE*, *8*(2), e56166. https://doi.org/10.1371/journal.pone.0056166

McCollough, A. W., Machizawa, M. G., & Vogel, E. K. (2007). Electrophysiological Measures of Maintaining Representations in Visual Working Memory. *Cortex*, *43*(1), 77–94. https://doi.org/10.1016/S0010-9452(08)70447-7

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, *49*(2), 270–291. https://doi.org/10.1037/a0028228

Metha, A. B., Vingrys, A. J., & Badcock, D. R. (1994). Detection and discrimination of moving stimuli: The effects of color, luminance, and eccentricity. *JOSA A*, *11*(6), 1697–1709. https://doi.org/10.1364/JOSAA.11.001697

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

Moore, K. S., & Weissman, D. H. (2010). Involuntary transfer of a top-down attentional set into the focus of attention: Evidence from a contingent attentional capture paradigm. *Attention, Perception, & Psychophysics*, *72*(6), 1495–1509. https://doi.org/10.3758/APP.72.6.1495

Morey, C. C. (2019). Perceptual grouping boosts visual working memory capacity and reduces effort during retention. *British Journal of Psychology*, *110*(2), 306–327. https://doi.org/10.1111/bjop.12355

Moutoussis, K., & Zeki, S. (1997). Functional segregation and temporal hierarchy of the visual perceptive systems. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *264*(1387), 1407–1414. https://doi.org/10.1098/rspb.1997.0196

Müller, H. J., & Krummenacher, J. (2006). Locus of dimension weighting: Preattentive or postselective? *Visual Cognition*, *14*(4–8), 490–513. https://doi.org/10.1080/13506280500194154

Nassar, M. R., Helmers, J. C., & Frank, M. J. (2018). Chunking as a rational strategy for lossy data compression in visual working memory. *Psychological Review*, *125*(4), 486–511. https://doi.org/10.1037/rev0000101

Oberauer, K. (2009). Chapter 2 Design for a Working Memory. In *Psychology of Learning and Motivation* (Vol. 51, pp. 45–100). Academic Press. https://doi.org/10.1016/S0079-7421(09)51002-X

Oberauer, K., & Eichenberger, S. (2013). Visual working memory declines when more features must be remembered for each object. *Memory & Cognition*, *41*(8), 1212–1227. https://doi.org/10.3758/s13421-013-0333-6

Oberauer, K., & Kliegl, R. (2001). Beyond resources: Formal models of complexity effects and age differences in working memory. *European Journal of Cognitive Psychology*, *13*(1–2), 187–215. https://doi.org/10.1080/09541440042000278

Oberauer, K., & Kliegl, R. (2006). A formal model of capacity limits in working memory. *Journal of Memory and Language*, *55*(4), 601–626. https://doi.org/10.1016/j.jml.2006.08.009

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences*, *11*(12), 520–527. https://doi.org/10.1016/j.tics.2007.09.009

Olivers, C. N. L., Peters, J., Houtkamp, R., & Roelfsema, P. R. (2011). Different states in visual working memory: When it guides attention and when it does not. *Trends in Cognitive Sciences*, *15*(7), 327–334. https://doi.org/10.1016/j.tics.2011.05.004

Olson, I. R., & Chun, M. M. (2002). Perceptual constraints on implicit learning of spatial context. *Visual Cognition*, *9*(3), 273–302. https://doi.org/10.1080/13506280042000162

Olson, I. R., Chun, M. M., & Allison, T. (2001). Contextual guidance of attention: Human intracranial event-related potential evidence for feedback modulation in anatomically early temporally late stages of visual processing. *Brain*, *124*(7), 1417–1425. https://doi.org/10.1093/brain/124.7.1417

Olson, I. R., & Jiang, Y. (2004). Visual short-term memory is not improved by training. *Memory & Cognition*, *32*(8), 1326–1332. https://doi.org/10.3758/BF03206323

Olson, I. R., Jiang, Y., & Moore, K. S. (2005). Associative learning improves visual working memory performance. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(5), 889–900. https://doi.org/10.1037/0096-1523.31.5.889

Ort, E., Fahrenfort, J. J., ten Cate, T., Eimer, M., & Olivers, C. N. (2019). Humans can efficiently look for but not select multiple visual objects. *eLife*, *8*, e49130. https://doi.org/10.7554/eLife.49130

Ort, E., & Olivers, C. N. L. (2020). The capacity of multiple-target search. *Visual Cognition*, *28*(5–8), 330–355. https://doi.org/10.1080/13506285.2020.1772430

Pailian, H., & Halberda, J. (2015). The reliability and internal consistency of one-shot and flicker change detection for measuring individual differences in visual working memory capacity. *Memory & Cognition*, *43*(3), 397–420. https://doi.org/10.3758/s13421-014-0492-0

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

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

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Peterson, D. J., & Berryhill, M. E. (2013). The Gestalt principle of similarity benefits visual working memory. *Psychonomic Bulletin & Review*, *20*(6), 1282–1289. https://doi.org/10.3758/s13423-013-0460-x

Peterson, D. J., Gözenman, F., Arciniega, H., & Berryhill, M. E. (2015). Contralateral delay activity tracks the influence of Gestalt grouping principles on active visual working memory representations. *Attention, Perception, & Psychophysics*, *77*(7), 2270–2283. https://doi.org/10.3758/s13414-015-0929-y

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

Põder, E., & Wagemans, J. (2007). Crowding with conjunctions of simple features. *Journal of Vision*, *7*(2), 23. https://doi.org/10.1167/7.2.23

Pollmann, S. (2019). Working memory dependence of spatial contextual cueing for visual search. *British Journal of Psychology*, *110*(2), 372–380. https://doi.org/10.1111/bjop.12311

Poole, B. J., & Kane, M. J. (2009). Working-memory capacity predicts the executive control of visual search among distractors: The influences of sustained and selective attention. *Quarterly Journal of Experimental Psychology*, *62*(7), 1430–1454. https://doi.org/10.1080/17470210802479329

Prieto, A., Peinado, V., & Mayas, J. (2022). Does perceptual grouping improve visuospatial working memory? Optimized processing or encoding bias. *Psychological Research*, *86*(4), 1297–1309. https://doi.org/10.1007/s00426-021-01555-w

Reinhart, R. M. G., & Woodman, G. F. (2014). Causal Control of Medial–Frontal Cortex Governs Electrophysiological and Behavioral Indices of Performance Monitoring and Learning. *Journal of Neuroscience*, *34*(12), 4214–4227. https://doi.org/10.1523/JNEUROSCI.5421-13.2014

Rensink, R. A. (2002). Change Detection. *Annual Review of Psychology*, *53*(Volume 53, 2002), 245–277. https://doi.org/10.1146/annurev.psych.53.100901.135125

RepovŠ, G., & Baddeley, A. (2006). The multi-component model of working memory: Explorations in experimental cognitive psychology. *Neuroscience*, *139*(1), 5–21. https://doi.org/10.1016/j.neuroscience.2005.12.061

Ringer, R. V., Coy, A. M., Larson, A. M., & Loschky, L. C. (2021). Investigating Visual Crowding of Objects in Complex Real-World Scenes. *I-Perception*, *12*(2), 2041669521994150. https://doi.org/10.1177/2041669521994150

Rouder, J. N., Morey, R. D., Cowan, N., Zwilling, C. E., Morey, C. C., & Pratte, M. S. (2008). An assessment of fixed-capacity models of visual working memory. *Proceedings of the National Academy of Sciences*, *105*(16), 5975–5979. https://doi.org/10.1073/pnas.0711295105

Rouder, J. N., Morey, R. D., Morey, C. C., & Cowan, N. (2011). How to measure working memory capacity in the change detection paradigm. *Psychonomic Bulletin & Review*, *18*(2), 324–330. https://doi.org/10.3758/s13423-011-0055-3

Rovamo, J., & Virsu, V. (1979). An estimation and application of the human cortical magnification factor. *Experimental Brain Research*, *37*(3), 495–510. https://doi.org/10.1007/BF00236819

Roy, Y., & Faubert, J. (2023). Is the Contralateral Delay Activity (CDA) a robust neural correlate for Visual Working Memory (VWM) tasks? A reproducibility study. *Psychophysiology*, *60*(2), e14180. https://doi.org/10.1111/psyp.14180

Schankin, A., & Schubö, A. (2009). Cognitive processes facilitated by contextual cueing: Evidence from event-related brain potentials. *Psychophysiology*, *46*(3), 668–679. https://doi.org/10.1111/j.1469-8986.2009.00807.x

Scolari, M., Kohnen, A., Barton, B., & Awh, E. (2007). Spatial attention, preview, and popout: Which factors influence critical spacing in crowded displays? *Journal of Vision*, *7*(2), 7. https://doi.org/10.1167/7.2.7

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

Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-Specific Delay Activity in Human Primary Visual Cortex. *Psychological Science*, *20*(2), 207–214. https://doi.org/10.1111/j.1467-9280.2009.02276.x

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin*, *138*(4), 628–654. https://doi.org/10.1037/a0027473

Soto, D., Rotshtein, P., & Kanai, R. (2014). Parietal structure and function explain human variation in working memory biases of visual attention. *NeuroImage*, *89*, 289–296. https://doi.org/10.1016/j.neuroimage.2013.11.036

Sreenivasan, K. K., Curtis, C. E., & D'Esposito, M. (2014). Revisiting the role of persistent neural activity during working memory. *Trends in Cognitive Sciences*, *18*(2), 82–89. https://doi.org/10.1016/j.tics.2013.12.001

Strasburger, H., Harvey, L. O., & Rentschler, I. (1991). Contrast thresholds for identification of numeric characters in direct and eccentric view. *Perception & Psychophysics*, *49*(6), 495–508. https://doi.org/10.3758/BF03212183

Tamber-Rosenau, B. J., Fintzi, A. R., & Marois, R. (2015). Crowding in Visual Working Memory Reveals Its Spatial Resolution and the Nature of Its Representations. *Psychological Science*, *26*(9), 1511–1521. https://doi.org/10.1177/0956797615592394

Travis, S. L., Mattingley, J. B., & Dux, P. E. (2013). On the role of working memory in spatial contextual cueing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(1), 208–219. https://doi.org/10.1037/a0028644

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Treisman, A., & Zhang, W. (2006). Location and binding in visual working memory. *Memory & Cognition*, *34*(8), 1704–1719. https://doi.org/10.3758/BF03195932

Tsubomi, H., & Watanabe, K. (2017). Development of visual working memory and distractor resistance in relation to academic performance. *Journal of Experimental Child Psychology*, *154*, 98–112. https://doi.org/10.1016/j.jecp.2016.10.005

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, *114*(1), 104–132. https://doi.org/10.1037/0033-295X.114.1.104

Vadillo, M. A., Giménez-Fernández, T., Aivar, M. P., & Cubillas, C. P. (2020). Ignored visual context does not induce latent learning. *Psychonomic Bulletin & Review*, *27*(3), 512–519. https://doi.org/10.3758/s13423-020-01722-x

Van Snellenberg, J. X., Conway, A. R. A., Spicer, J., Read, C., & Smith, E. E. (2014). Capacity estimates in working memory: Reliability and interrelationships among tasks. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(1), 106–116. https://doi.org/10.3758/s13415-013-0235-x

Vater, C., Kredel, R., & Hossner, E.-J. (2017). Detecting target changes in multiple object tracking with peripheral vision: More pronounced eccentricity effects for changes in form than in motion. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(5), 903–913. https://doi.org/10.1037/xhp0000376

Velisavljević, L., & Elder, J. H. (2008). Visual short-term memory for natural scenes: Effects of eccentricity. *Journal of Vision*, *8*(4), 28. https://doi.org/10.1167/8.4.28

Vickery, T. J., King, L.-W., & Jiang, Y. (2005). Setting up the target template in visual search. *Journal of Vision*, *5*(1), 8. https://doi.org/10.1167/5.1.8

Vickery, T. J., Sussman, R. S., & Jiang, Y. V. (2010). Spatial context learning survives interference from working memory load. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6), 1358–1371. https://doi.org/10.1037/a0020558

Vogel, E. K., & Awh, E. (2008). How to Exploit Diversity for Scientific Gain: Using Individual Differences to Constrain Cognitive Theory. *Current Directions in Psychological Science*, *17*(2), 171–176. https://doi.org/10.1111/j.1467-8721.2008.00569.x

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

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

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

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychological Bulletin*, *138*(6), 1172–1217. https://doi.org/10.1037/a0029333

Wagemans, J., Feldman, J., Gepshtein, S., Kimchi, R., Pomerantz, J. R., van der Helm, P. A., & van Leeuwen, C. (2012). A century of Gestalt psychology in visual perception: II. Conceptual and theoretical foundations. *Psychological Bulletin*, *138*(6), 1218–1252. https://doi.org/10.1037/a0029334

Wallace, J. M., & Tjan, B. S. (2011). Object crowding. *Journal of Vision*, *11*(6), 19–19. https://doi.org/10.1167/11.6.19

Wang, F. H. (2020). Explicit and implicit memory representations in cross-situational word learning. *Cognition*, *205*, 104444. https://doi.org/10.1016/j.cognition.2020.104444

Whitney, D., & Levi, D. M. (2011). Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences*, *15*(4), 160–168. https://doi.org/10.1016/j.tics.2011.02.005

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

Wolfe, J. M. (1998). What Can 1 Million Trials Tell Us About Visual Search? *Psychological Science*, *9*(1), 33–39. https://doi.org/10.1111/1467-9280.00006

Wolfe, J. M. (2021). Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review*, *28*(4), 1060–1092. https://doi.org/10.3758/s13423-020-01859-9

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, *5*(6), 495–501. https://doi.org/10.1038/nrn1411

Wolfe, J. M., Klempen, N., & Dahlen, K. (2000). Postattentive vision. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 693–716. https://doi.org/10.1037/0096-1523.26.2.693

Woodman, G. F., & Luck, S. J. (2004). Visual search is slowed when visuospatial working memory is occupied. *Psychonomic Bulletin & Review*, *11*(2), 269–274. https://doi.org/10.3758/BF03196569

Woodman, G. F., Vecera, S. P., & Luck, S. J. (2003). Perceptual organization influences visual working memory. *Psychonomic Bulletin & Review*, *10*(1), 80–87. https://doi.org/10.3758/BF03196470

Woodman, G. F., Vogel, E. K., & Luck, S. J. (2012). Flexibility in visual working memory: Accurate change detection in the face of irrelevant variations in position. *Visual Cognition*, *20*(1), 1–28. https://doi.org/10.1080/13506285.2011.630694

Xu, Y. (2002). Encoding color and shape from different parts of an object in visual short-term memory. *Perception & Psychophysics*, *64*(8), 1260–1280. https://doi.org/10.3758/BF03194770

Xu, Y. (2006). Understanding the object benefit in visual short-term memory: The roles of feature proximity and connectedness. *Perception & Psychophysics*, *68*(5), 815–828. https://doi.org/10.3758/BF03193704

Xu, Y., & Chun, M. M. (2007). Visual grouping in human parietal cortex. *Proceedings of the National Academy of Sciences*, *104*(47), 18766–18771. https://doi.org/10.1073/pnas.0705618104

Xu, Z., Adam, K. C. S., Fang, X., & Vogel, E. K. (2018). The reliability and stability of visual working memory capacity. *Behavior Research Methods*, *50*(2), 576–588. https://doi.org/10.3758/s13428-017-0886-6

Yang, Y., Coutinho, M. V. C., Greene, A. J., & Hannula, D. E. (2021). Contextual cueing is not flexible. *Consciousness and Cognition*, *93*, 103164. https://doi.org/10.1016/j.concog.2021.103164

Yörük, H., Santacroce, L. A., & Tamber-Rosenau, B. J. (2020). Reevaluating the sensory recruitment model by manipulating crowding in visual working memory representations. *Psychonomic Bulletin & Review*, *27*(6), 1383–1396. https://doi.org/10.3758/s13423-020-01757-0

Zellin, M., Conci, M., Mühlenen, A. von, & Müller, H. J. (2013). Here Today, Gone Tomorrow – Adaptation to Change in Memory-Guided Visual Search. *PLOS ONE*, *8*(3), e59466. https://doi.org/10.1371/journal.pone.0059466

Zellin, M., Conci, M., von Mühlenen, A., & Müller, H. J. (2011). Two (or three) is one too many: Testing the flexibility of contextual cueing with multiple target locations. *Attention, Perception, & Psychophysics*, *73*(7), 2065–2076. https://doi.org/10.3758/s13414-011-0175-x

Zellin, M., von Mühlenen, A., Müller, H. J., & Conci, M. (2014). Long-term adaptation to change in implicit contextual learning. *Psychonomic Bulletin & Review*, *21*(4), 1073–1079. https://doi.org/10.3758/s13423-013-0568-z

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

Zhong, C., Qu, Z., Yang, N., Sun, M., Wang, Y., & Ding, Y. (2024). Susceptibility to Attentional Capture by Target-Matching Distractors Predicts High Visual Working Memory Capacity. *Psychological Science*, *35*(11), 1203–1216. https://doi.org/10.1177/09567976241279520

Zhu, Z., Fan, Z., & Fang, F. (2016). Two-stage perceptual learning to break visual crowding. *Journal of Vision*, *16*(6), 16. https://doi.org/10.1167/16.6.16

Zinchenko, A., Conci, M., Müller, H. J., & Geyer, T. (2018). Predictive visual search: Role of environmental regularities in the learning of context cues. *Attention, Perception, & Psychophysics*, *80*(5), 1096–1109. https://doi.org/10.3758/s13414-018-1500-4

Zinchenko, A., Conci, M., Töllner, T., Müller, H. J., & Geyer, T. (2020). Automatic Guidance (and Misguidance) of Visuospatial Attention by Acquired Scene Memory: Evidence From an N1pc Polarity Reversal. *Psychological Science*, *31*(12), 1531–1543. https://doi.org/10.1177/0956797620954815