DIRECTING ATTENTION IN WORKING MEMORY

Amy Louise Atkinson

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

University of Leeds
School of Psychology

December 2019

INTELLECTUAL PROPERTY AND COPYRIGHT

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

The data presented in **Chapter 2** has been accepted for publication as:

> Atkinson, A. L., Berry, E. D., Waterman, A. H., Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2018). Are there multiple ways to direct attention in working memory? *Annals of the New York Academy of Sciences*, *1424*(1), 115-126.

The data presented in **Chapter 4** has been accepted for publication as:

> Atkinson, A. L., Waterman, A. H., & Allen, R. J. (2019). Can children prioritize more valuable information in working memory? An exploration into the effects of motivation and memory load. *Developmental Psychology*, *55*(5), 967-980.

The data presented in **Chapter 3** is being revised following submission as:

> Atkinson, A. L., Allen, R. J., Baddeley, A. D., Hitch, G. J., & Waterman, A. H. (In Preparation). Can valuable information be prioritised in verbal

working memory? Submitted to *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Only material produced by Amy Louise Atkinson has been used in this thesis.

ACKNOWLEDGEMENTS

Firstly, I'd like to thank my supervisors, Richard Allen and Amanda Waterman, for their guidance, support, and expertise throughout my PhD. Thanks to Richard for buying into my ideas for "just one more" side project, and thanks to Amanda for controlling this to some degree, otherwise this thesis would never have been submitted. I really couldn't have wished for better supervisors. I'd also like to thank Alan Baddeley and Graham Hitch for providing helpful and insightful comments on several chapters; it has been an honour to work with you both. I'm also grateful to Liam Hill for providing useful guidance at the end of my first and second years, and Ed Berry for convincing me to give R a go.

The studies reported in this thesis would not have been possible without the assistance of several undergraduate and Master's projects students. Thanks to everyone who was involved in data collection, particularly Erin Funnell who collected data for numerous projects. I'd also like to thank Alessandra Souza and Klaus Oberauer for welcoming me into their lab at the University of Zurich. It was great to experience a different research environment and to collaborate with you both. This visit, and indeed the PhD in general, would not have been possible without generous funding from the Economic and Social Research Council (ESRC).

My PhD experience would not have been anywhere near as enjoyable without the friends I've made along the way. Greg deserves a special mention for many trips "to the pub!", which included a lot of laughter. Lucy has always been around for cake, and Joyti (or my "PhD sister" as she'd like to be known) has provide many motivational quotes over the years. Alice and Raabia have also been good friends. Thanks are also owed to Zeynep, Mshari, Jack, Emily, and Paul for lots of good times over the last year, as well my third-year office mates, Lydia and Siobhan.

On a personal note, I'd like to thank my family, particularly my mum, dad, and grandma Ivy, for always supporting me in my endeavours, even if they didn't really understand much of my research. My non-PhD friends also deserve thanks for always being there, particularly Jess, Leanne, Laura, Maria, and my godson Toby. Finally I'd like to thank Matthew and our little bump for bringing so much happiness into my life and giving me a final push to get this thesis submitted.

ABSTRACT

Working memory (WM) allows a limited amount of information to be temporarily stored in a state of heightened accessibility for use in ongoing processing. This thesis examined how increasing the relative value of one item affects performance on WM tasks. Chapter 2 explored whether this type of prioritisation differs from probe frequency, whereby one item is more likely to be tested than the rest. This pair of studies demonstrated that probe value and probe frequency are likely to involve distinct cognitive mechanisms. Chapter 3 then examined whether value effects can be observed in an auditory-verbal WM task, as research to date has investigated effects in the visual domain. Significant value effects were observed, which were not reduced in size when rehearsal and attentional resource mechanisms were reduced. Participants did, however, appear to abandon the less valuable items when both rehearsal and executive control resources were disrupted. Next, Chapter 4 investigated whether children can direct their attention to more valuable information in WM when sufficiently motivated to do so. Probe value effects were observed, although these were smaller than those previously reported in adults. There was also evidence that children selectively prioritised the more valuable item when it was likely to considerably enhance performance. Finally, Chapter 5 examined whether prioritising a more valuable item at WM can enhance performance on a surprise LTM task. A value effect was observed at LTM, but only when the more valuable item had been tested at WM. Taken together, these findings demonstrate that value effects in WM are distinct from other attentional manipulations and robust across modalities and groups. Prioritising a more valuable item for a WM test can also enhance longer-term retention if the item is assessed at WM.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ADHD	Attention deficit hyperactivity disorder
AFC	Alternative forced choice
ANOVA	Analysis of variance
BDR	Backward digit recall
BF	Bayes Factor
BOSS	Bank of Standardized Stimuli
cm	Centimetres
D	Cohen’s D
ERPs	Event-related potentials
FDR	Forward digit recall
FoA	Focus of Attention
fMRI	Functional magnetic resonance imaging
GG	Greenhouse-Geisser
ISI	Interstimulus interval
LTM	Long-term memory
M	Mean
ms	Milliseconds
MSE	Mean square error
RT	Response time
SE	Standard error
SES	Socioeconomic status
SL	Spatial location
SP	Serial position
STM	Short-term memory

WM	Working memory

CHAPTER 1

GENERAL INTRODUCTION

"Selection is the very keel on which our mental ship is built…"
James (1890)

Working memory (WM) refers to a system that allows a limited amount of information to be temporarily stored in a state of heightened accessibility for use in ongoing processing (Cowan, 1988). Research on the topic has increased exponentially in recent years, with WM now considered to be one of the most important constructs in cognitive psychology (Cowan, 2014; 2017; Ku, 2018; Melby-Lervåg & Hulme, 2012). This is perhaps unsurprising given that WM is essential for many cognitive processes (Cowan, 2010), and is also highly predictive of a range of important outcomes, including fluid intelligence (Engle, Tuholski, Laughlin, & Conway, 1999) and academic achievement (Alloway & Alloway, 2010; Gathercole, Pickering, Knight & Stegmann, 2004). However, despite its importance, the capacity of WM is severely constrained (Hartshorne, 2008), with individuals typically only able to store 3-4 items at a given time (Cowan, 2010). Relative to young adults, this capacity is reduced in a number of groups, including children (Gathercole, Pickering, Ambridge, & Wearing, 2004), older adults (Dobbs & Rule, 1989), and individuals diagnosed with a range of developmental and neurological conditions (e.g. Martinussen, Hayden, Hogg-Johnson & Tannock, 2005, Owen et al., 1997). Given that WM is limited in capacity, but central to higher-order cognition and broader achievement, research has therefore begun to explore how WM can be enhanced. One approach, that will be examined in this thesis, involves encouraging individuals to prioritise particularly valuable or goal-relevant information. If individuals are able to direct their attention towards particularly important information, interventions could be developed which encourages individuals to focus their attention towards particular items. Furthermore, in real life settings, information encountered often naturally differs in its importance or value (Oberauer & Hein, 2012).

The literature review that follows will discuss several theoretical models of WM, including domain-specific (Baddeley, 2000; Baddeley & Hitch, 1974; Logie, 1995; 2001) and domain-general approaches (Cowan, 1999; Oberauer, 2002). The importance and the limited-capacity nature of the system will then be considered, as well as approaches that have been taken in an attempt to improve WM. The relationship between WM and attention will then be discussed, including a review of existing research which has examined whether individuals can direct their attention to particularly important or goal-relevant information in WM. Finally, the aims and the structure of the thesis will be presented.

1.1. Theories of working memory

The distinction between holding an idea in mind and the ability to retrieve something no longer in mind was first discussed by Locke (1690; Cowan, 2014; Logie, 1996). Similar ideas were proposed by other researchers, including Galton (1883), James (1905), Wundt (1948) and Miller (1956), who also distinguished between a limited-capacity temporary store and a more flexible long-term memory (LTM) system (Baddeley, 1997; Logie & Cowan, 2015). In 1958, Broadbent suggested that the short-term limited capacity system might serves as an interface between sensory input and LTM (Broadbent, 1958). These ideas were developed further by Atkinson and Shiffrin (1968), who proposed the multi-store model of memory. This comprised a sensory store, a short-term store, and a long-term store. It was postulated that information first enters the sensory store, before being transferred into the short-term store. This acts as an interface with the long-term component and is also responsible for strategic maintenance and retrieval processes. Although this model was highly influential, subsequent research highlighted some important shortcomings (Baddeley, 2012; Craik & Lockhart, 1972) including an inability to explain several neuropsychological findings.

Motivated by these inadequacies, Baddeley and Hitch (1974) conducted a series of experiments to explore whether distinct cognitive tasks rely on a common 'WM system', and how this might relate to current beliefs about short-term memory (STM). Participants completed verbal reasoning, comprehension, and LTM free-recall tasks whilst attempting to retain series of digits. It was assumed that the retention of digits would place demands on a STM system, which should impair performance on the cognitive tasks if they rely on the same resources (Baddeley & Hitch, 1974). In the verbal reasoning task, response time (RTs) were significantly increased when participants were required to concurrently retain sequences of six digits. This manipulation also impaired performance on the comprehension and LTM free-recall tasks. From this, it was concluded that a common system is involved in information processing and short-term storage, with a trade-off existing between them.

Based on these findings, Baddeley and Hitch (1974) tentatively proposed the first version of the multi-component model of WM, which comprised several separable components. This contrasted with prior conceptualisations of STM, which tended to assume that the system was unitary in nature. Within this framework (Baddeley & Hitch, 1974), the core component was a limited-capacity workspace that could flexibly store or process information depending on task demands. This sub-system was also thought to play a key role in the recoding of information, the application of maintenance strategies such as chunking, and reconstruction at retrieval. This processor was accompanied by components solely capable of storage. Based on evidence that performance on several cognitive tasks was disrupted by the retention of digits, it was argued that the first of these storage components held speech-like information. An equivalent system was also proposed, capable of retaining visual information.

Over the years that followed, the theoretical framework was developed further, and a more formal model was proposed (see Figure 1.1A; Baddeley, 1983; 1986). The core of this model was a central executive; an attentional system responsible for control processes and the deployment of strategies. Notably, this slightly deviates from the conceptualisation of this component in the 1974 chapter, in which the central executive was considered to be capable of processing *or* storage. This controller is supported by two storage-based subsystems, the articulatory loop (now referred to as the phonological loop; Baddeley, 2000) and the visuo-spatial scratchpad (now typically referred to as the visuo-spatial sketchpad; Baddeley, 2000). These components are responsible for retaining verbal and visuo-spatial information respectively. In the 1986 version, the phonological loop was further separated into the phonological store, capable of holding verbal information, and an articulatory control process, which sub-vocally rehearses information in order to promote retention. The visuo-spatial sketchpad has also since undergone fractionation, with Logie (1995) suggesting that the sub-system comprises a visual cache, responsible for maintaining visual information, and an inner scribe, that rehearses spatial representations (Logie, 1995; 2011).

The model was further adapted in 2000, where Baddeley added a fourth component, the episodic buffer (Baddeley, 2000; see Figure 1.1B). This domain-general storage system is thought to integrate information from a variety of sources. It is believed to bring together information from different modalities (and thus different subsystems), storing information in a multi-dimensional code. It is also considered to play a key role in integrating individual features such as colour and shape into identifiable objects. Finally, the component is assumed to provide a temporary interface between the WM components and the LTM store and is thus thought to be important for learning.

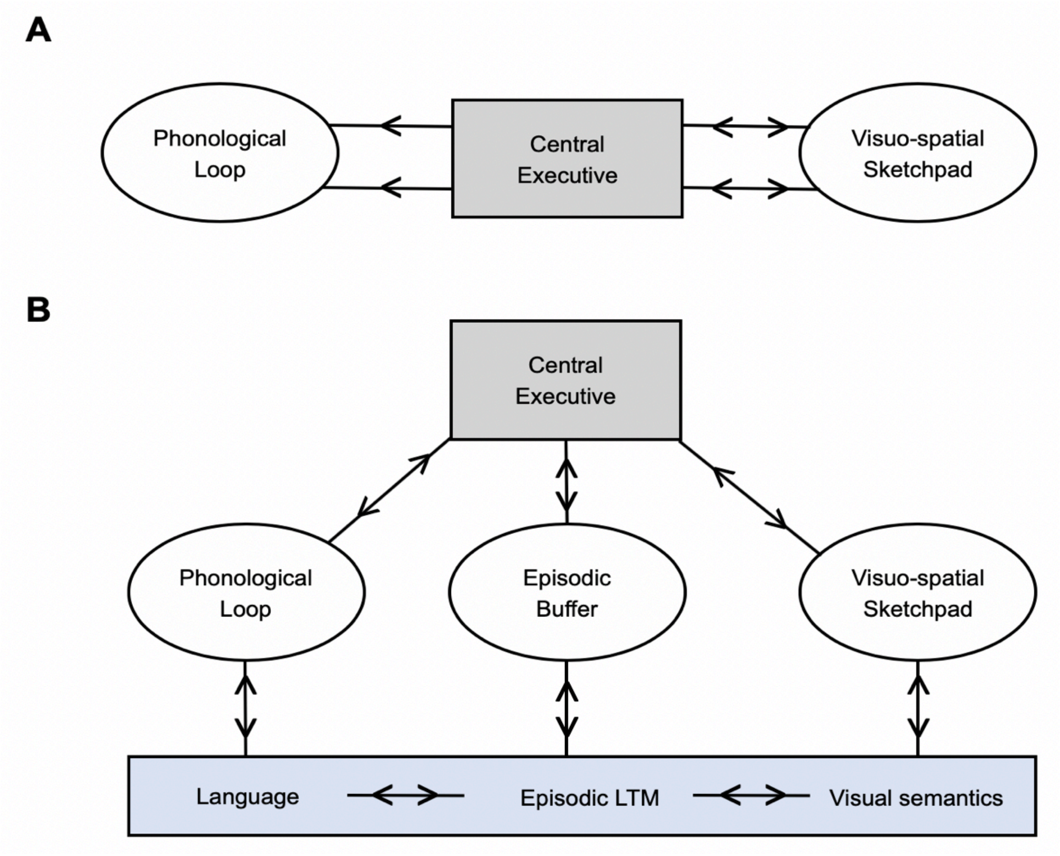


Figure 1.1. A) The three-component version of the multi-component model (Baddeley, 1983; 1986), comprising the central executive, the phonological loop (Baddeley 1986; or articulatory loop; Baddeley, 1983) and the visuo-spatial sketchpad (Baddeley, 1986; or visuo-spatial scratchpad; Baddeley, 1983). B) The more recent, four-component model (Baddeley, 2000), including the central executive, the phonological loop, the visuo-spatial sketchpad, and the episodic buffer. Figure adapted from “Working memory: Theories, models, and controversies” by A. D. Baddeley, 2012, *Annual Review of Psychology*, *63*, 1-29. Copyright 2012 by Annual Reviews.

This model is considered to be the most influential and widely accepted account of WM, and is able to explain a wealth of findings from cognitive and developmental psychology, neuropsychology, and neuroimaging (Baddeley, 2000). However, researchers investigating WM have not accepted all aspects of the model. For instance, although Logie broadly agrees that WM comprises multiple components, he has argued that the central executive is redundant (Logie, 2016), and that executive control is achieved through the interaction of multiple different cognitive functions, including inhibition and task switching.

Other researchers have refuted the fundamental underpinnings of the multi-component approach, suggesting that WM is better defined as a domain-general resource rather than a set of separable components (e.g. Cowan, 1999; Oberauer, 2002). The most influential of these domain-general frameworks has been the embedded processes model (Cowan, 1999). This suggests that WM comprises an activated form of LTM, as well as a focus of attention (FoA) that can store items in a more integrated form (Cowan, 1999; 2016). Information that is relevant to the current task is held in the activated form of LTM. A subset of this information can then be stored in the FoA, whereby information is re-activated and integrated into multi-feature objects or scenes. This FoA is limited in capacity, with the sub-system able to hold approximately three to five items at a given time. This limit can, however, refer to single features (e.g. colour, shape) or whole items that comprise multiple features. The system also comprises a central executive, which controls voluntary processes and directs attention to particular items in order to bring them into the FoA (Cowan, 1999; 2016). Information can also enter the FoA through a non-deliberate or automatic route, due to changes in the environment (e.g. a loud noise) or the presence of stimuli with personal relevance (Cowan, 1999; 2016).

A somewhat similar model is Oberauer's three-embedded-components model (Oberauer, 2002; 2013). In this model, three states of activation are proposed: activated LTM, a region of direct access, and a FoA. Activated LTM holds information that may be relevant for the current task, such as numbers whilst performing a digit span task. The region of direct access is more limited in capacity, holding representations that are bound to their context (e.g. remembering the digits in the current trial and their correct serial order). This region is similar to Cowan’s conceptualisation of the FoA. However, unlike Cowan, Oberauer argues that this region has a fixed capacity. Instead, he posits that limitations arise from interference between temporary bindings (Oberauer, 2013), though he does note that the number of items typically held is approximately four (Oberauer and Hein, 2012). A third component is also proposed, the FoA, which selects an item from the region of direct access. This increases the item’s accessibility relative to the other representations and allows it to be refreshed or manipulated. As this sub-system involves selecting an item from the set, it typically holds one item. However, Oberauer believes that this is not a fixed capacity limit, but that bringing multiple items into the FoA would undermine the function of this sub-system (Oberauer, 2013). This differs slightly from other, similar conceptualisations of the FoA, where it has been suggested that the system has a fixed one-item capacity limit (McElree, 2001; 2006; McElree & Dosher, 1989).

Other WM models also exist, including the time-based resource sharing model (Barrouillet, Bernardin, & Camos, 2004). This proposes that storage and processing of information in WM both compete for a shared attentional resource. As this attentional resource is severely limited, individuals can either engage in processing or maintenance processes at a given time, with attention being rapidly switched between them (Barrouillet et al., 2004; Barrouillet & Camos, 2012). When attention is directed towards processing, information being stored in WM suffers from time-based decay and may be lost completely if it is not refreshed (Portrat, Camos & Barrouillet, 2009). Conversely, Macken, Taylor and Jones (2015) suggest that WM abilities are limited by interactions between the individual and their set of cognitive and perceptual-motor skills, the task, and the material.

This section has outlined the most common theoretical approaches to WM, including domain-specific and domain-general approaches. The next section will explore the ways in which information can be encoded, maintained, and lost from WM.

1.2. How is information encoded, maintained and lost from working memory?

Encoding can be defined as the processing of a stimulus in order to determine its identity and characteristics (Ricker & Cowan, 2014). This is considered to be fast, although information might be liable to loss if no further processing takes place (Bayliss, Bogdanovs, & Jarrold, 2015). In order to stabilise the representations, it is thought that information must be transformed through a process termed ‘WM consolidation’ (Bayliss et al., 2015; Ricker, Nieuwenstein, Bayliss, & Barrouillet, 2018). This attentionally-demanding process converts transient sensory information into durable representations that are more resistant to forgetting and interference (Jolicoeur & Dell’Acqua, 1998; Nieuwenstein & Wyble, 2014; Ricker & Cowan, 2014). Although encoding and consolidation are related, they are considered to be separable processes, with consolidation thought to occur after encoding is complete (Ricker et al., 2018).

Information must then be retained until the retrieval phase, with two major maintenance mechanisms thought to operate within WM (Camos, Mora, & Oberauer, 2011). The first is articulatory rehearsal, in which verbal information is repeated sub-vocally. Verbal information can easily enter this system, whilst visual or spatial information can be recoded into a phonological form to permit articulatory rehearsal (e.g. repeating “red circle” in order to retain a memory for a red circle observed visually; Baddeley, 2003). The second mechanism is attentional refreshing. Although there is some disagreement on how best to define this term, it broadly refers to an attentionally-demanding mechanism that acts to keep representations active and in an accessible state (Camos et al., 2018). This process is thought to strengthen the representation within WM, thus preventing it from fading (Camos et al., 2018). Two types of attentional refreshing have been proposed: fast and slow refreshing. Fast refreshing is thought to occur outside of conscious awareness at a rate of up to 50 milliseconds (ms) per item. In contrast, slow refreshing is thought to reflect a more deliberate act, involving the individual explicitly thinking of the item. The speed of this type of refreshing has not been conceptualised as well, although it is considered to be less than 150-250ms per item (Camos et al., 2018).

Articulatory rehearsal and attentional refreshing are thought to reflect independent mechanisms that are selected adaptively depending on task demands (Camos et al., 2011). The mechanisms are not considered to be mutually exclusive, however, with individuals able to employ both concurrently (Camos et al., 2011; Camos, Lagner, & Barrouillet, 2009).

There is considerable debate within the literature as to what causes loss of information from WM (Lemaire & Portrat, 2018; Souza and Oberauer, 2015). Some argue that information is lost as a result of temporal decay, whereby representations fade over time (Barrouillet & Camos, 2012; Cowan, 2001). In contrast, other researchers believe that information is lost due to interference. This can occur when items that are similar to one another or occur close together in time become confused (Souza and Oberauer, 2015). It can also result from cognitive activity or operations (Lewandowsky, Oberauer, & Brown, 2008). Until recently, researchers have typically accepted a decay or interference view, with evidence of one being taken as evidence against the other (Lemaire & Portrat, 2018; Souza and Oberauer, 2015). However, in recent years, several researchers have noted the substantial evidence supporting both accounts, arguing that both are likely to exist (Altmann & Schunn, 2012; Lemaire & Portrat, 2018). It is also important to note that both temporal decay and interference are considered distinct from removal, an adaptive process whereby individuals remove no longer relevant information from WM (Lewis-Peacock, Kessler, & Oberauer, 2018).

To summarise, information must be encoded and consolidated in WM in order to produce stable representations. Information is then thought to be maintained through verbal rehearsal and/or attentional refreshing. There is ongoing debate regarding how information is lost, with more recent accounts proposing that both temporal decay and interference play a role. The next section will discuss the capacity limits of WM and its important for many cognitive tasks.

1.3. The limitations and importance of working memory

The capacity of WM is considered to be severely limited (e.g. Cowan 2001; 2010; Luck & Vogel, 1997; but see Macken et al, 2015). Some researchers suggest that WM has a fixed capacity of approximately three to five items in young adults, when strategies such as rehearsal are controlled (Cowan, 2001; 2010). This limit was first demonstrated by Luck and Vogel (1997) using the change detection paradigm.

Arrays of items were presented at encoding and test, and participants had to identify whether they were identical or not. Set size was varied, with 1-12 items being presented. When the to-be-remembered items were coloured squares, participants showed near-perfect performance for set sizes 1-3, with performance decreasing thereafter. Estimates of WM capacity were also calculated, revealing that participants were able to hold approximately four items at a given time. This was replicated in numerous follow-up experiments, in which rehearsal was disrupted, encoding time was manipulated, and the to-be-remembered feature was orientation as opposed to colour. The authors were also interested in exploring whether this observed limit relates to individual features or integrated objects. To investigate this, participants completed an experiment, in which coloured bars were presented. In the first condition, they were told that the colour may vary. In the second, they were told that orientation might vary, and in the third they were informed that either feature might change. In this latter condition, participants had to remember twice the number of features than in the former two. Despite this, performance was very similar across conditions. Taken together, this suggests that participants can retain approximately four integrated items at a given time. Although the findings of this study have not always replicated (e.g. Delvenne & Bruyer, 2004; Wheeler & Treisman, 2002), a 3-5 item limit has been supported by research using other methods, including functional magnetic resonance imaging (fMRI; Todd & Marois, 2004), event-related potentials (ERPs; e.g. Ikkai, McCollough, & Vogel, 2010; Luck & Vogel, 2013; Vogel & Machizawa, 2004), and mathematical modelling (Cowan, 2001; 2010).

However, not all researchers agree that WM capacity has a fixed capacity (e.g. Bays, Catalao, & Husain, 2009; Bays & Husain, 2008; Macken et al., 2015).

This is particularly apparent in the visual domain, where it has been suggested that WM may be better explained as a shared resource that can be flexibly allocated across items (Bays et al., 2009; Bays & Husain, 2008). Proponents of this model suggest that there is no upper limit to the number of items that can be stored in WM, but that performance is instead determined by the *quality* of representations (Ma, Husain, & Bays, 2014). Nevertheless, despite disagreement regarding the exact nature of the system, advocates of this model agree that WM is highly constrained.

There are also substantial age-related changes in WM across the lifespan, with children and older adults typically showing poorer abilities than younger adults (Brockmole & Logie, 2013; Gathercole, Pickering, Ambridge et al., 2004). Gathercole, Pickering, Ambridge et al. (2004) investigated how WM develops across childhood and adolescence by assessing four to 15-year-olds on multiple WM measures. Linear increases in performance were observed between the ages of 4 and 14 years, with performance levelling off between 14 and 15 years of age. There is, however, some evidence that developmental changes may continue beyond this age. In a study of over 50,000 participants between 8-75 years of age, Brockmole and Logie (2013) found that visual WM abilities peaked at approximately 20 years of age. Performance than began to drop soon after, with participants in their late 50s exhibiting worse performance than 8-9-year old children. A variety of disorders have also been associated with poor WM, including Parkinson's disease (Lee et al., 2010; Siegert, Weatherall, Taylor, & Abernethy, 2008), Schizophrenia (Forbes, Carrick, McIntosh, & Lawrie, 2009; Lee & Park, 2005) and Attention deficit hyperactivity disorder (ADHD; Martinussen et al., 2005).

Despite these limitations, WM is essential for a wide range of important cognitive activities, such as learning (Alloway, 2006), planning (Cowan, 2010), reasoning (Süß, Oberauer, Wittmann, Wilhelm, & Schulze, 2002) and language comprehension (Daneman & Merikle, 1996). As a result of this, WM is considered central to many everyday activities, including planning, problem solving, and following instructions (Cowan, 2010; Fürst & Hitch, 2000; Gathercole, Durling, Evans, Jeffcock, & Stone, 2008; Jaroslawska, Gathercole, Logie, & Holmes, 2015; Souza & Oberauer, 2015). For instance, when planning the order of shops to visit in a town centre, WM would be required to calculate the optimal route depending on your current location (Cowan, 2010). Similarly, when completing a mental arithmetic problem, one would need to store intermediate values whilst calculating the final values (Fürst & Hitch, 2000). WM abilities are also highly predictive of a range of important outcomes, including fluid intelligence (Engle et al., 1999), general cognitive ability (Gold et al. 2010), and academic achievement in childhood and adolescence (Alloway & Alloway, 2010; Gathercole, Pickering, Knight et al., 2004).

To summarise, WM is considered essential for a myriad of mental tasks and everyday activities. However, the capacity is extremely limited, particularly in children, older adults, and individuals diagnosed with a range of disorders. As such, research has begun to examine how performance on WM tasks can be enhanced. One approach which has been extensively investigated is training. The evidence for this will be discussed in the next section.

1.4. Working memory training

WM training has been the focus of much research in recent years. The majority of these studies have examined the effects of ‘core WM training’ (Morrison & Chein, 2011). This typically involves participants repeatedly engaging in demanding WM tasks aimed at improving domain-general mechanisms (Morrison & Chein, 2011). Although some studies and narrative reviews have reported positive findings (e.g. Klingberg, 2010; Morrison & Chein, 2011), the general consensus is that there is currently no, or very limited, evidence to suggest that training results in any far-transfer effects to outcomes such as fluid intelligence, attention, or academic achievement (Anderson et al., 2018; Melby-Lervåg & Hulme, 2012; Redick et al., 2013; Shipstead, Redick, & Engle, 2010; 2012).

In contrast to the large number of studies which have explored core WM training, a very limited amount of research has examined strategy training (Morrison & Chein, 2011). This involves individuals being taught effective methods to encode, store, and retrieve information (Morrison & Chein, 2011), such as articulatory rehearsal (Comblain, 1994; St Clair-Thompson, Stevens, Hunt, & Bolder, 2010; Turley-Ames & Whitfield, 2003), and elaborative encoding strategies such as mental imagery and story-formation (Carretti, Borella, & De Beni, 2007; McNamara & Scott, 2001; St Clair-Thompson et al., 2010). Within these studies, inconsistent findings have been reported. For instance, St Clair-Thompson et al. (2010) investigated how training 5-8-year old children to use a range of strategies affected performance on WM tasks, mental arithmetic, the ability to follow instructions, and performance on standardised educational assessments. Classes within schools were randomly assigned to an intervention group (receiving a ‘memory booster’ intervention) or a control group (receiving no intervention). As part of the intervention, children were taught to use rehearsal, visual imagery, story formation, and (in some cases dependent upon performance) grouping. Children completed two 30-minute sessions a week over a six to eight-week period. Relative to the children in the control group, those in the intervention group exhibited greater improvements in performance on WM tasks, mental arithmetic, and following instructions. There were, however, no significant effects on standardised tests of mathematics or reading five months after the intervention.

As such, there is currently little evidence that core WM training or strategy training results in any meaningful benefits to performance outside of laboratory settings. Research has therefore begun to investigate ways in which individuals may be able to optimise their performance on WM tasks, rather than attempting to enhance their capacity. In many WM tasks, items are not of equal value or goal-relevance (Oberauer & Hein, 2012; Souza & Oberauer, 2016). One approach might therefore involve focusing attention on the most goal-relevant or valuable information presented. This approach will be explored in this thesis. The next section will introduce the concept of attention, and evaluate its relationship with WM.

1.5. Working memory and attention

The term ‘attention’ can be defined and conceptualised in a number of ways. Within this thesis, attention will be used to refer to a set of resources that select information for processing through bottom-up capture or top-down control (Fougnie, 2008; Yantis, 2000). This definition assumed that attention is not unitary and involves multiple mechanisms. Bottom-up capture occurs when stimuli in the external environment attract attention, whilst top-down control involves the allocation of attention to information that is important or goal-relevant (Connor, Egeth, & Yantis, 2004; Katsuki & Constantinidis, 2014; Pashler, Johnston, & Ruthruff, 2001). Further distinguishing the two, bottom-up capture is considered to be automatic in nature, whilst top-down control is often thought to be voluntary and reliant on executive resources (Katsuki & Constantinidis, 2014; Pinto, van der Leij, Sligte, Lamme, & Scholte, 2013; Theeuwes, 2018).

There are, however, many other definitions of attention within the literature (Chun, Golomb, & Turk-Browne, 2011; Fougnie, 2008; Johnston, McCann, Remington, 1995). Chun et al. (2011) introduced a taxonomy which focuses on delineating the target of attention. This separates the system into external and internal attention. Whilst external attention involves the selection of perceptual information in the environment, internal attention reflects the selection or maintenance of information that is internally generated (Chun et al., 2011). Examples of this may include goals, rules, responses, or memories (Chun et al., 2011). Attention can also be considered in terms of central and perceptual processes (Johnston et al., 1995). Central attention is closely related to executive control and is involved with tasks that are cognitively demanding. There are a limited number of resources available, such that performance is reduced if two tasks are conducted simultaneously that rely on central attention (Fougnie, 2008). In contrast, perceptual attention reflects the selection of a subset of information from the external world. Although these various taxonomies are conceptually similar (Chun et al., 2011), they may be thought of as reflecting different aspects of attention. For instance, whilst bottom-up and top-down mechanisms reflect the ways in which attention is *allocated*, central and peripheral attention reflect *types* of attention, and external and internal attention refer to the *target information* involved.

A close bi-directional relationship has been found between WM and attention, whereby the contents of WM determine what is attended to (Awh and Jonides, 2001), and information that is attended to in the external world is more likely to enter WM (Vogel, Woodman, & Luck, 2005). There is also growing consensus that a sub-region operates within WM, termed the FoA. This sub-region is thought to hold a limited amount of information in a state of heightened accessibility, thus enhancing its retrieval speed and accuracy. As discussed in Section 1.1, the FoA is central to several prominent theories of WM, including Cowan's embedded processes model (Cowan, 1999; 2016) and Oberauer's embedded component theory (Oberauer, 2002; 2013). It is, however, important to note that the conceptualisation of this sub-system differs. Within Cowan's model, the FoA is considered able to store up to four integrated items and 'zoom in and out' on particular representations (Cowan, 2005). In contrast, Oberauer suggests that the FoA typically holds a single item, although this is not a fixed capacity limit (Oberauer, 2013). More recency, the FoA has also been considered in relation to Baddeley's multi-component model (Baddeley, 2000), where it has been suggested that the sub-region may reside within the episodic buffer (Hu, Hitch, Baddeley, & Allen, 2014).

Within this thesis, the FoA will be used refer to a sub-system within WM that is able to hold a subset of information in a highly accessible state (Hu et al., 2014). This definition adopts the assumptions of the FoA proposed by Hu et al. (2014), which suggests that the FoA is equivalent to the modality-general episodic buffer of Baddeley's multi-component model. This conceptualisation argues that the FoA is not restricted to a single item, with the sub-system able to hold around two to three items concurrently (Allen & Ueno, 2018; Hu et al., 2014). This theoretical assumption was selected as it has been used by previous work examining the effects of value in WM (e.g. Allen & Ueno, 2018; Berry et al., 2018; Hitch et al., 2018; Hu et al., 2014; 2016). This definition is, not, however, necessarily inconsistent with other conceptualisations of the FoA, such as those proposed by Cowan (2005) and Oberauer (2002; 2013). For instance, Cowan argues that the FoA can zoom in and out on particular representations depending upon task demands (Cowan, 2005). Moreover, although Oberauer argues that the FoA typically holds a single item, he suggests that this is not a fixed capacity (Oberauer, 2002; 2013).

It is assumed that information can enter the FoA automatically through bottom-up capture or strategically through top-down control. Hu et al. (2014) suggested that each item encoded enters the FoA automatically, but is then displaced by subsequent items. This therefore results in a recency boost, whereby the final item is subject to an automatic boost relative to the earlier items (Berry et al., 2018; Hu et al., 2014; 2016). It is also believed that information can be held in the FoA strategically if an item is particularly valuable or goal-relevant (Allen & Ueno, 2018; Hu et al., 2016). Strategically holding an item in the FoA is thought to enhance memory for this information (Allen & Ueno, 2018; Hitch et al., 2018; Hu et al., 2014; 2016). This process does, however, appear to be reliant on executive resources (Hu et al., 2016), leading researchers to speculate on the mechanisms that may result in strategic memory boosts. Suggestions have included attentional refreshing, whereby the particularly important item is brought into the FoA and refreshed more or for longer periods of time relative to the other items during maintenance (Hitch et al., 2018; Hu et al., 2016; Sandry et al., 2014). Another possibility is that the particularly important item is consolidated into WM better than the other items (Allen & Ueno, 2018; Hitch et al., 2018), a process that is also considered to involve the FoA (De Schrijver & Barrouillet, 2017). As with the broad conceptualisation of the FoA outlined above, these theoretical assumptions are based mainly on work by Baddeley and colleagues (e.g. Allen & Ueno, 2018; Berry et al., 2018; Hitch et al., 2018; Hitch et al., 2014; 2016). Such findings are not, however, inconsistent with other theoretical positions. For instance, Cowan suggested that the FoA can be accessed by both automatic and strategic routes (Cowan, 1999; 2016).

The following sections will review findings from several paradigms which have been used to investigate the FoA and the extent to which individuals are able to direct their attention in WM.

1.6. Visual cueing

The paradigm used most extensively to investigate the relationship between WM and attention has been cueing. Within this paradigm, a series of to-be-remembered items are presented and then tested following a brief delay. Before (pre-cue) or after (retro-cue) the set of items are displayed, a visual cue is presented that informs participants which item will, or is most likely to, be tested at retrieval (see Figure 1.2). Both pre-cues and retro-cues significantly enhance performance and speed retrieval rates relative to conditions in which the cue is absent or uninformative. These effects have been demonstrated using various tasks, including change detection, item recognition, cued recall, and continuous delay estimation (Souza & Oberauer, 2016). Within the visual domain, retro-cue effects are so robust that they have been considered a benchmark that theoretical models of WM should account for (Oberauer et al., 2018). Moreover, although studies using other types of material are less frequent, visual retro-cues have also been reported to enhance memory for auditory stimuli (Backer & Alain, 2012; Kumar et al., 2013) and visually-presented verbal material (such as letters; Krefeld-Schwalb, 2018). Effects have also been reported in a range of populations, including children (Shimi, Nobre, Astle, & Scerif, 2014; Shimi &

Scerif, 2017) and older adults (Loaiza & Souza, 2018; Mok, Myers, Wallis & Nobre, 2016, but see Duarte et al., 2013 & Newsome et al., 2015). Boosts are, however, sometimes associated with slower RTs and decreased accuracy for uncued items presented in the same trial (Astle, Summerfield, Griffin, & Nobre, 2012; Chun et al., 2011; Gressman & Janczyk, 2016; Rerko, Souza, & Oberauer, 2014, but see Landman, Spekreijse & Lamme, 2003; Li & Saiki, 2014; Myers, Chekroud, Stokes, & Nobre, 2018).

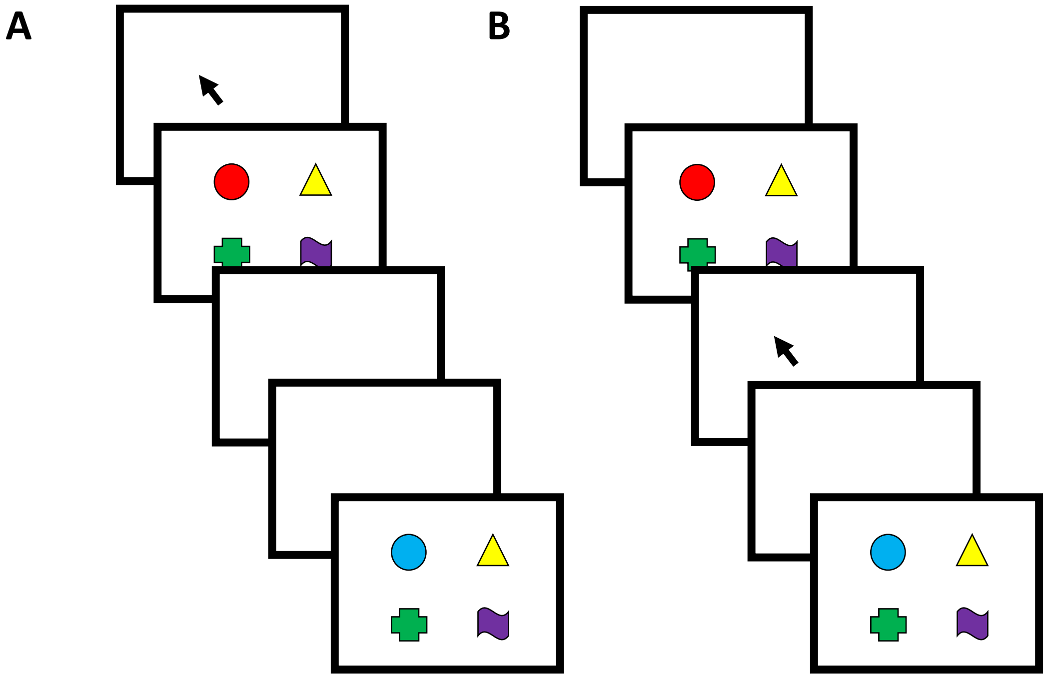


Figure 1.2. The pre-cue (A) and retro-cue (B) manipulations typically used, illustrated using a change detection paradigm. Performance in these conditions is typically compared to a condition in which no cue or an informative cue is presented (e.g. a * in the centre of the screen which does not identify which item is more likely to be tested).

As pre-cues are presented before the array of items, boosts may at least partially reflect a biasing at encoding towards the cued item (Robison & Unsworth, 2017; Wang, Yan, Wang, Olivers, & Theeuwes, 2017). In contrast, retro-cues are only able to influence the maintenance and retrieval phases. There is, however, ongoing debate regarding the precise mechanisms driving retro-cue effects (Souza & Oberauer, 2016). One possibility is that the targeted item may occupy an ‘output-driving state’ in which it is highly accessible and more readily available to guide behaviour (Myers et al., 2018). This may enable it to be prioritised for comparison with the probe relative to the other items (Matsukura, Luck, & Vecera, 2007). Alternatively, boosts may result from attentional refreshing, strengthening of the item-context bindings, or protection from interference (Souza, Rerko, & Oberauer, 2016) or decay (Pertzov, Bays, Joseph, & Husain, 2013). Other hypotheses have also been proposed, which have suggested that the retro-cue may provide a “head start” in the retrieval process (Souza et al., 2016) or allow uncued items to be removed from WM (Souza, Rerko, & Oberauer, 2014). It is currently unclear whether one of these hypotheses or a combination can account for the effects of retro-cues, as there is evidence supporting and refuting each account (Souza & Oberauer, 2016). The mechanisms involved may also differ depending on task factors (Astle et al., 2012; Souza & Oberauer, 2016), including the validity of the cue. For instance, when a cue always identifies the item that will be probed, a successful approach might be to remove the uncued items from WM. However, this strategy may not be useful when the cues are not 100% valid, as this would result in very poor memory for the other items.

There is also ongoing debate regarding the type of attention involved in pre-cues and retro-cues. The boost from pre-cues might result from participants strategically reducing the number of to-be-remembered items (Berryhill, Richmond, Shay, and Olson, 2012). However, if the cue appears abruptly, attention might be automatically directed towards the location of the cued item (Berryhill et al., 2012; Schmidt, Vogel, Woodman, & Luck, 2002). Similarly, retro-cue boosts might result from bottom-up automatic capture or top-down voluntary control. Some studies have found that the effects of retro-cues depend on the validity of the cue (Gunseli, van Moorselaar, Meeter, and Olivers, 2015; Shimi et al., 2014). For instance, Gunseli et al. (2015) found larger boosts when the retro-cue was valid on 80% of the trials, relative to a condition in which the cue was only valid 50% of the time. From this, it has been concluded that individuals use cues strategically depending on task demands (Souza & Oberauer, 2016). As such, the direction of attention towards retro-cues might be considered to be controlled by top-down processes. However, Berryhill et al. (2012) found that performance for cued items was enhanced even when the cues were uninformative. This suggests that cueing effects are not entirely strategic in nature. Further complicating the debate, it has been suggested that type of attention recruited might depend on the type of retro-cue. If the cue is presented as a word (e.g. identifying a colour) or a number (e.g. identifying a serial position), the cue might recruit top-down attentional control (Berryhill et al., 2012). However, there is evidence that other types of cues, such as an arrow, recruit both bottom-up and top-down attention (Berryhill et al., 2012; Ristic & Kingstone, 2006). As arrow cues are very commonly used, the majority of retro-cue effects reported in the literature are likely to reflect a combination of top-down executive control and bottom-up automatic capture.

To summarise, cueing has frequently been used to investigate the relationship between WM and attention. Both pre-cues and retro-cues enhance performance and reduce RTs relative to conditions in which no cues are presented.

1.7. Probe frequency

Another manipulation which has been employed to investigate the relationship between WM and attention is probe frequency (Cowan, Morey, AuBuchon, Zwilling, & Gilchrist, 2010; Gorgoraptis, Catalao, Bays & Husain, 2011; Klyszejko, Rahmati, & Curtis, 2014). This paradigm is similar to cueing, in that one item is more likely to be tested than the other items presented in the trial. However, the way in which the manipulation is employed differs. Whilst cueing paradigms present a cue to indicate which item is likely to be tested on a trial-by-trial basis, participants are informed that one particular item is always more likely to be tested in the probe frequency manipulation. For instance, participants might be told that a particular shape, colour or location will be tested more frequently. The manipulation might also be considered similar to the Hebb repetition effect, in which a presentation array or list is repeatedly presented across a series of trials (e.g. Hebb, 1961; Oberauer, Jones, & Lewandowsky, 2015; Page, Cumming, Norris, Hitch, & McNeil, 2006). However, a key difference is that with the probe frequency manipulation, participants are typically informed of the item that will be tested at the start of the session or block, and the main outcome of interest is whether memory for this item is enhanced relative to those less likely to be tested. In contrast, in the Hebb repetition manipulation, participants are not informed that arrays or lists will be repeated, with the main outcome of interest whether memory for the repeating information is enhanced relative to novel arrays/lists.

This manipulation was employed by Gorgoraptis et al. (2011), who presented participants with four coloured orientation bars sequentially. After a brief period of time, one of the bars was probed, and participants had to reproduce it. Before the start of the task, participants were told that one of the colours was more likely to be tested. In trials where this item was presented, it was tested 62.5% of the time, relative to 12.5% for the other three items. This colour differed across participants but was consistent throughout the experiment. Performance was compared to a control condition, in which the cued colour was not presented and each item was equally likely to be tested (25%). Performance at the targeted item was more precise than the baseline condition, whilst performance at the non-targeted items was significantly worse. This demonstrates that participants are able to direct their attention to more goal-relevant information, but that this negatively affects memory for other items presented within the same sequence.

Similar findings have also been reported in other studies (Cowan et al., 2010; Klyszejko et al., 2014) For instance, Klyszejko et al. (2014) presented participants with series of four coloured dots, and asked them to remember the location of each. Following a brief delay, one of the dots was presented and participants had to indicate which direction it had moved in. Before the start of the experiment, participants were given extensive training about the colour of the dots, where they learned how likely each was to be tested (i.e. white = 100%, red = 75%, blue = 25%, black = 0%). As predicted, accuracy increased as the item was more likely to be assessed. This demonstrates that individuals are able to exert careful control over WM, distributing their attention in a graded manner. There is also evidence that children as young as seven years old can apply probe frequency information to enhance performance, although this ability does not appear to be as developed in this group as 12-13-year olds adolescents and adults (Cowan et al., 2010).

There has been little research investigating whether probe frequency effects in WM are driven by bottom-up automatic capture or top-down strategic processes. It would therefore be useful for further research to investigate this.

1.8. Item value

The final manipulation to be reviewed, which will be the primary focus of this thesis, is the effect of value on memory. Several paradigms have been used to study the effects of value on memory. This includes monetary incentives, value-directed remembering, and notional points. These will be summarised below.

1.8.1. Monetary incentives

Monetary incentives have commonly been used to study the effects of value on memory (Adcock, Thangavel, Whitfield-Gabrieli, Knutson & Gabrieli, 2006; Gruber & Otten, 2010; Murty and Adcock, 2014; Shigemune et al., 2010; Wittmann et al., 2005). For instance, Adcock, et al. (2006; see Figure 1.3A) presented participants with scenes to remember for a recognition test 24 hours later. Before the scene was presented, participants were told that the scene was either high value (worth $5.00) or low value (worth $0.10). At test, participants were presented with the old scenes as well as some new ones, and they had to recognise which items had been encoded previously. If participants recognised the previously encountered scenes as ‘old’, they were given the sum of money that had been associated with each item at encoding. Participants were also asked to give a confidence rating for their judgement, stating whether they ‘remember’, ‘know’, ‘are pretty sure’, or ‘guessing’. Participants recognised more of the high value items relative to the low value items. They were also significantly more likely to report ‘remembering’ or ‘knowing’ the high value items. Similar findings were reported by Gruber and Otten (2010) who presented participants with words. After 15 minutes, participants were asked to categorise the word based on their confidence into ‘remember’, ‘confident old’, ‘not confident old’, ‘not confident new’, ‘confident new’. Recognition accuracy was significantly higher for the valuable items, with the proportion of hits to misses nearly twice as large in the high value condition (0.56 relative to 0.29 of the low value condition). Similarly to Adcock et al. (2006), participants also reported ‘remembering’ more of the high value words. Taken together, this demonstrates that individuals can prioritise high value items in the verbal and visual domains, which results in more accurate recognition and higher confidence ratings.

Although the studies reviewed to date have examined LTM, the effects of monetary incentives have also been studied in WM (Klyszejko et al., 2014). Klyszejko et al. (2014) presented participants with coloured dots to remember for a brief period of time (see Figure 1.3B). These dots were presented sequentially, with each colour being associated with a particular reward value (e.g. white = $1.00, red = $0.75, blue = $0.25, black = $0.00). After a short retention interval, one of the dots was presented to the left or the right of its original position. Participants were asked to indicate the direction in which the dot had moved, with a correct response gaining them the amount of money associated with the colour. Each dot was equally likely to be tested. As predicted, accuracy increased as the colour was associated with a higher monetary value.

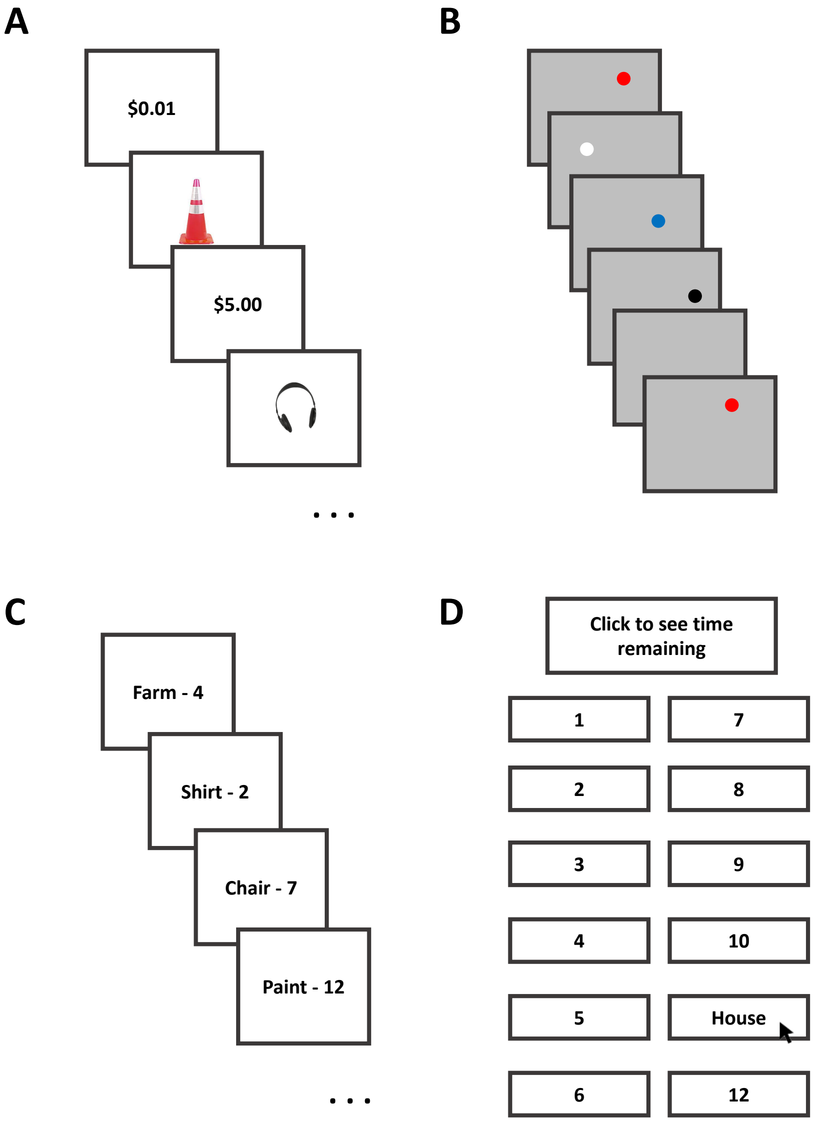


Figure 1.3. The monetary incentives (A and B) and value-directed paradigms (C and D) used to study the effects of value on memory. A) Images (e.g. Adcock et al., 2006) or words (e.g. Gruber & Otten, 2010) are preceded by a monetary value, which participants gain if they recognise the item at retrieval. Other studies have used other slight variations of this, such as presenting the item value before a block of trials (e.g. Shigemune et al., 2010). B) Experimental procedure used by Klyszejko et al. (2014). Participants were shown dots varying in their point values. After a short delay, they were presented with one of the dots and asked whether it had moved slightly to the left or the right. C) A version of the value-directed remembering paradigm where the word and the point values are presented simultaneously (e.g. Castel, Humphreys, et al. 2011). D) Another version of the value-directed remembering paradigm, which is typically used when participants are able to

allocate their own study time. Participants must click on the point value to reveal the word (e.g. Robison & Unsworth, 2017). B is adapted from "Attentional Priority Determines Working Memory Precision" by Z. Klyszejko, M. Rahmati, and C. E. Curtis, 2014, *Vision Research, 105,* p. 71. Copyright 2014 by Elsevier Ltd. D is adapted from "Working memory capacity, strategic allocation of study time, and value-directed remembering" by M. K. Robison & N. Unsworth, 2017, *Journal of Memory and Language, 93*, p. 235. Copyright 2015 by Elsevier Inc.

1.8.2. Value-directed remembering

Another commonly used paradigm is value-directed remembering (Watkins and Blooms, 1999). In this paradigm, participants are shown lists of words to remember which are associated with a particular point value (e.g. Car – 6). In some studies, participants have been shown one word and point value at a time (see Figure 1.3C; e.g. Castel, Humphreys, et al. 2011). In other studies, they have been able to distribute their own time across items during the encoding phase (e.g. Figure 1.3D; e.g. Robison & Unsworth, 2017). If participants correctly recall the word in a subsequent memory test, they gain the number of points associated with it. The overall goal of the task is to gain as many points as possible, although long lists are usually presented which would make it difficult to recall all of the words presented. Across several experiments, it has been shown that participants typically recall significantly more high value items than low value items (e.g. Castel, Benjamin, Craik, & Watkins, 2002; Nguyen, Marini, Zacharczuk, Llano, & Mudar, 2019; Robison & Unsworth, 2017; Siegel & Castel, 2018).

This paradigm has, however, typically explored effects in episodic memory, as opposed to WM. Furthermore, a limitation of this paradigm is that it does not require individuals to encode all of the information, with some studies even allowing participants to allocate their own study time between words (e.g. Robison & Unsworth, 2017). This may result in the abandonment of less valuable items. Indeed, there is evidence to suggest that this strategy is optimal (Atkinson, Baddeley, & Allen, 2017; Robison & Unsworth, 2017). For instance, Robison and Unsworth (2017) found that participants who utilised this approach exhibited better performance than participants who attempted to encode most or all of the items. However, in real life settings, information of reduced value often still needs to be remembered to some extent. For example, consider following a series of instructions whilst baking a cake. The critical piece of information might be the temperature of the oven and the cooking time. One must also remember the gist of the other instructions, however, or the cake may not be edible. Similarly, when studying for an exam, some information may be deemed more important than the rest. Nevertheless, one should still have a basic understanding of the less important information in order to perform well on the test.

1.8.3. Probe value

A third paradigm which has been used to investigate the effects of value on WM is probe value (referred to as strategic prioritisation in some previous literature; Allen & Ueno, 2018; Berry, Waterman, Baddeley, Hitch, & Allen, 2018; Hu et al., 2014; Hu, Allen, Baddeley, & Hitch, 2016; Hitch, Hu, Allen, & Baddeley, 2018). In this paradigm, participants are typically presented with a series of coloured shapes to remember for a brief period of time. Memory for one of the items is then tested using cued-recall. Before encoding, participants are told the point values associated with each position. Although one or more of the items are typically more valuable, each position is equally likely to be tested. As such, participants should attempt to retain all of items, but particularly focus on the more valuable one(s).

This manipulation was introduced by Hu et al. (2014) who presented participants with four coloured shapes to remember sequentially (e.g. red triangle, blue circle, etc; see Figure 1.4A). After a short delay, one item was probed, and participants were asked to recall the shape or the colour. In one condition, participants were told that the final item was associated with the most points and the first was associated with the least, with the middle items worth intermediate values (e.g. 1-2-3-4). In a second condition, the point values were reversed, such that the first item was worth the most points and the final item was worth the least (e.g. 4-3-2-1). After encoding, a to-be-ignored shape was presented (termed a suffix) in some trials. This was either a plausible suffix, whereby the shape was from the same experimental set as the to-be-remembered items (e.g. a yellow star), or an implausible suffix, where a pale coloured irregular shape was displayed. Regardless of the point values associated with each item, a recency effect was observed, in which performance of the final item was superior to the other serial positions (SPs). Moreover, when the first item or the final item was associated with more points, a value effect was observed, whereby performance at this SP was superior to the rest. This was, however, accompanied by costs to the less valuable items, which were not remembered as accurately. This suggests that increasing the value of an item does not improve overall performance, but rather encourages participants to distribute more of their attention towards this item. A further finding from this set of experiments was that both the value and recency boosts were disrupted by a suffix, with the effect being larger when the suffix was plausible. From this, it was concluded that the final item and the more valuable item are held in a privileged state within WM, which was assumed to be the FoA. This renders the items more accessible, but also vulnerable to interference, particularly from stimuli that are similar in nature to the to-be-remembered items. It is, however, important to note that these findings differ from outcomes in the cueing literature (e.g. Makovski & Jiang, 2007; van Moorselaar, Gunseli, Theeuwes, & Olivers, 2014; Souza et al., 2016; Shepherdson, Oberauer, & Souza, 2018), in which it has been suggested that items in the FoA are protected from interference. Taken together, this suggests that the FoA might respond differently depending on task factors (Hitch et al., 2018).

These effects were extended in a recent publication by Hitch et al. (2018). The first experiment in this series examined whether participants can prioritise the middle items, and whether these effects are also vulnerable to a post-stimulus suffix. Participants completed two prioritisation conditions, in which were they were told that the second or the third item was worth more points than the rest. Correct recall of the more valuable item was associated with four points, whereas correct recall of any other probed SP would gain them 1 point. Replicating Hu et al. (2014), probe value and recency boosts were observed, with both of these effects diminished by a suffix. Participants were also able to prioritise multiple items simultaneously, with the cost of this apparent at the final item. This further supports the conclusions drawn by Hu et al. (2014), suggesting that the more valuable item(s) and the final item are held in a privileged state, which that renders items more accessible, but also more vulnerable to interference. Moreover, evidence that recency effects are reduced in size when another item is more valuable suggests that these items compete for access to the FoA.

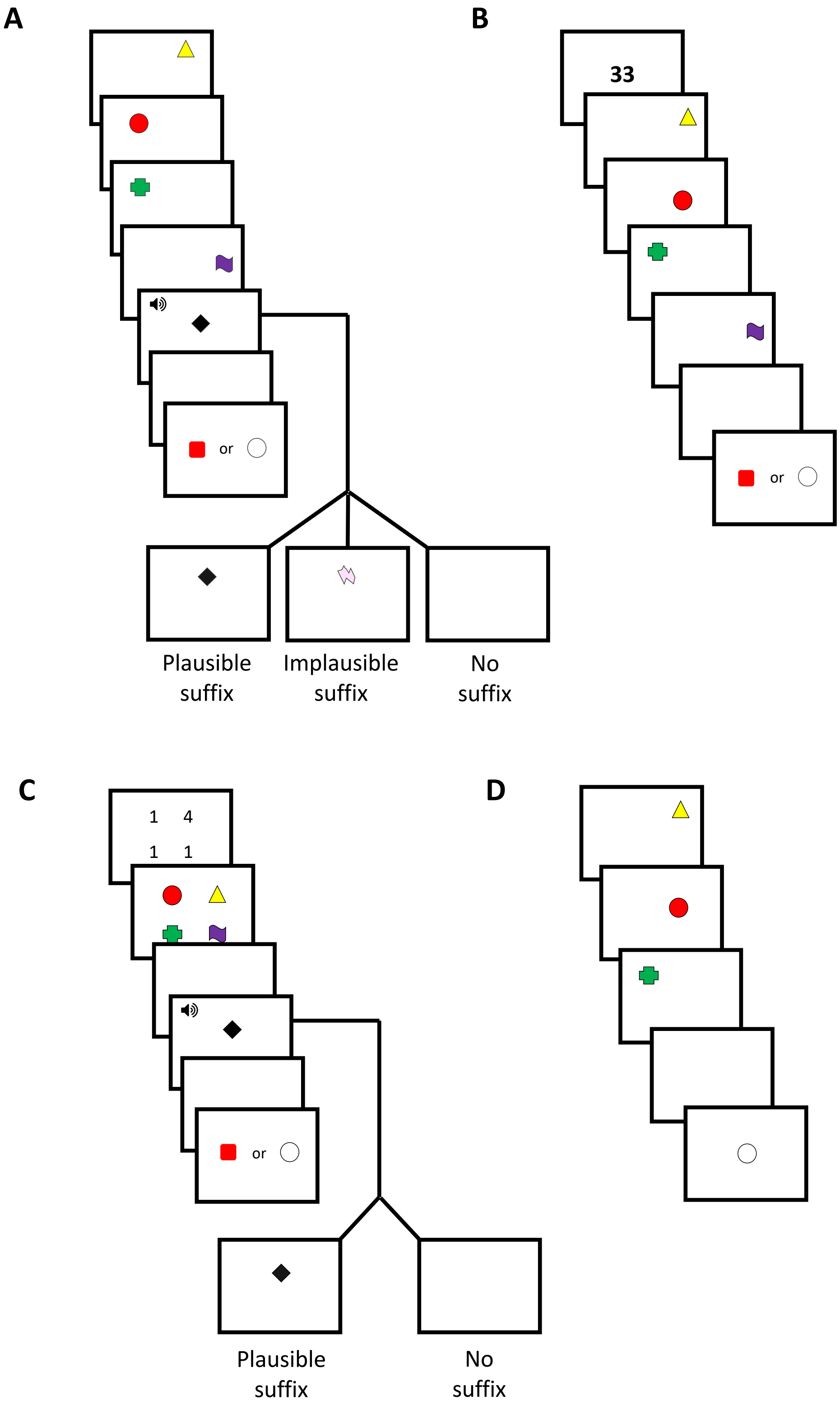


Figure 1.4. The probe value paradigm. A) The procedure used in Hu et al. (2014) and Hitch et al. (2018) to investigate the effects of probe value when items are presented sequentially, as well as vulnerability to a suffix. Participants also repeated 1-2-3-4 during encoding and maintenance in order to reduce verbal recoding. A

sound placed when the suffix was displayed such that participants knew to ignore that item. B) The experimental procedure used in Hu et al. (2016), investigating whether probe value and recency boosts are reliant on executive resources. Participants either had to repeat the two-digit number until the retrieval phase or count upwards in 2s from the number. C) The paradigm used in Allen and Ueno (2018) to investigate whether the probe value effects are presented when items are presented simultaneously. As the value of each position changed on a trial-by-trial basis, four numbers were presented at the same locations as the to-be-remembered items prior to encoding to inform participants how many points each item was worth. As in the previous work, participants repeated 1-2-3-4 during encoding and maintenance. D) The simpler paradigm used in Berry et al. (2018), in which participants were presented with three items to remember and were not told to engage in articulatory suppression. At retrieval, a shape probe was always presented and participants had to retrieve the colour. This figure was adapted from: A) "Executive and perceptual attention play different roles in visual working memory: Evidence from suffix and strategy effects." By Y. Hu, G. J. Hitch, A. D. Baddeley, M. Zhang, and R. J. Allen, 2014, *Journal of Experimental Psychology: Human Perception and Performance, 40*(4), Copyright 2014 by American Psychological Association. B) "Executive control of stimulus-driven and goal-directed attention in visual working memory" by Y. Hu., R. J. Allen. A. D. Baddeley, & G. J. Hitch, 2016, *Attention, Perception, & Psychophysics*, *78*(7), p. 2168. Copyright 2016 by The Psychonomic Society, Inc. C) "Multiple high-reward items can be prioritized in working memory but with greater vulnerability to interference" by R. J. Allen and T. Ueno, 2018, *Attention, Perception, & Psychophysics, 80*, p. 1734. Copyright 2018 by The Psychonomic Society, Inc. D) "The limits of visual working memory in children: exploring prioritization and recency effects with sequential presentation" by E. D. J. Berry, A. H. Waterman, A. D. Baddeley, G. J. Hitch, and R. J. Allen, 2018, *Developmental Psychology*, p. 243. Copyright 2018 by American Psychological Association.

There is also evidence that the cognitive mechanisms underlying probe value and recency boosts differs. This was demonstrated by Hu et al. (2016), who investigated whether the effects are reliant on executive resources. The paradigm used was similar to that employed by Hu et al. (2014; Figure 1.4B) with either the first item (4-3-2-1) or the final item worth the most points (1-2-3-4). No suffix was presented, with participants instead required to engage in a concurrent task during encoding and maintenance. This either involved repeating a two-digit number (low load) or counting upwards in 2s from a two-digit number (high load). If probe value and recency boosts are reduced under high cognitive load, it would suggest that the effects rely on executive control. Conversely, evidence that these boosts are not diminished would suggest that they are relatively cost-free and automatic in nature. As predicted, a probe value effect emerged, though this was diminished or abolished in the high load condition. A recency boost was also observed, which was present in both concurrent task conditions. From this, it was concluded that the way in which the effects emerge appear to differ; whilst probe value boosts appear to be reliant on executive resources, recency effects are obtained in a relatively automatic and cost-free manner. Nevertheless, both probe value and recency effects increase the probability that an item will be held in the FoA, which then, in turn, enhances performance for these items.

All of the studies discussed so far have employed a sequential mode of presentation (Hitch et al., 2018; Hu et al., 2014; 2016). Allen and Ueno (2018) recently extended this by investigating whether individuals are also able to direct their attention when item are encountered simultaneously. Participants were presented with arrays of four items and asked to recall a feature of one probed item following a brief delay. However, unlike in the previous studies (Hitch et al., 2018; Hu et al., 2014; 2016), the more valuable item changed on a trial-by-trial basis (Figure 1.4C). Accuracy was higher when an item was associated with more points, indicating that individuals can prioritise information in WM when items are encountered simultaneously. Participants were also able to direct their attention to multiple items and showed a graded effect when item values varied between 1-4 points (e.g. 1-2-3-4). Furthermore, as might be predicted, the probe value boosts were vulnerable to a suffix when multiple items were prioritised. Taken together, this extends findings from studies using a sequential mode of presentation, and further highlights the vulnerability of probe value effects to interference.

To summarise the studies discussed on probe value so far, individuals are able to prioritise a high value item in visual WM, though this effect appears to rely on executive resources. A recency effect is also observed, which may be obtained relatively automatically. These boosts are thought to occur as the valuable item(s) and the final item are held in a privileged state, which is assumed to be the FoA. This results in these items being more accessible, but also increases their vulnerability to interference.

All of the studies discussed above have investigated probe value effects in visual WM (Allen & Ueno, 2018; Berry et al., 2018; Hitch et al., 2018; Hu et al., 2014; 2016). To date, only one set of experiments has investigated effects in verbal WM (Sandry, Schwark, & MacDonald, 2014). In this set of experiments, participants were visually presented with three letters to remember. In some of the trials, one letter was presented in red, which indicated that this item was worth more points relative to the other items. After a brief period of time, participants completed a 2-AFC choice test in which they had to identify which letter had been presented during the encoding phase. In order to encourage participants to verbally recode the items, the letters were presented in upper case at test. In the first experiment, overall performance was near ceiling (Mean proportion correct = 0.93), thus preventing a sensitive analysis of effects. In order to reduce overall performance levels, Experiment 2 employed articulatory suppression. In this experiment, significant probe value effects were observed, suggesting that individuals can prioritise more valuable information in verbal WM. There was also a recency effect, whereby performance at that last time was higher than the other SPs. This also suggests that neither the probe value nor the recency effect are reliant on verbal rehearsal. However, care should be taken when drawing conclusions from this study regarding the extent to which individuals can prioritise in verbal WM. Letters were presented in lower case during encoding and upper-case during test to encourage participants to verbally recode the information. However, participants may have been able to retain the information in a visual format until test, or at least recruit assistance from the visual domain. This may be particularly likely within this experiment, as research has demonstrated that the use of articulatory suppression can prevent verbal recoding (Baddeley, Lewis, & Vallar, 1984; Baddeley & Hitch, 1994; Schendel & Palmer, 2007). Further research is therefore needed to investigate whether individuals can prioritise information in a purely verbal WM task.

Whilst research has indicated that children aged 7-10 years old can benefit from visual cueing (Shimi et al., 2014; Shimi & Scerif, 2017), only one set of studies has investigated probe value effects in this population (Berry et al., 2018). This set of experiments employed a similar manipulation to that used in studies investigating adults' ability to prioritise in the visual domain (Hu et al., 2014; 2016; Hitch et al., 2018), though using sequences of three rather than four items (Figure 1.4D). As in Hu et al. (2014, 2016), they were either informed that the first (Experiment 1 and 2) or the final item (Experiment 3) was more valuable. Participants also completed several additional measures of WM capacity, including forward digit recall (FDR), backward digit recall (BDR) and the Corsi blocks tapping task. Across all three experiments, no significant probe value effects were observed, with Bayes Factor (BF) analysis indicating evidence of no effect. This was taken as evidence that children do not or cannot prioritise items in WM based on item value, possibly as a result of undeveloped executive resources. In contrast, a significant recency effect emerged in all experiments. Performance at the final position was also unrelated to individual differences in WM capacity, unlike performance at the first or middle position. Taken together, this supports claims that the recency boost is obtained relatively automatically (e.g. Hu et al., 2016). It would, however, be useful for further research to investigate whether children are able to prioritise more valuable information in WM in a variety of task contexts.

More broadly, it is important to consider whether value effects (implemented through the use of notional points or monetary incentives) differ from other manipulations used to investigate the relationship between WM and attention. Often, terms such as ‘cueing’ and ‘prioritisation’ have been used interchangeably within the literature to refer to the three broad manipulations outlined above (cueing, probe frequency, and probe value; e.g. Gorgoraptis et al., 2011; Klyszejko et al., 2014). Although they are conceptually similar (Allen & Ueno, 2018), and all encourage participants to direct their attention to representations within WM, there may be subtle differences between them. Indeed, there is some evidence that probe value and cueing effects might differ, with probe value boosts vulnerable to interference (Hu et al., 2016), whilst cueing appears to protect items from interference (Souza et al., 2016). It is, however, unclear whether probe value and probe frequency effects encourage involve the same or distinct underlying cognitive mechanisms. It would therefore be beneficial for further research to investigate this. Moreover, although effects of value have been explored in WM (e.g. Allen & Ueno, 2018; Hitch et al., 2018; Hu et al., 2014; 2016) and LTM (e.g. Castel et al., 2002; Castel, Humphreys, et al. 2011; Robison & Unsworth, 2017), it is unclear whether prioritising an item for a WM test affects longer-term retention.

1.9. Thesis outline and aims

This thesis will examine whether adults and children can prioritise more valuable information in WM (termed probe value effects hereafter), and the consequences of this over the short- and long-term. This manipulation was selected as it is relatively less explored compared to other manipulations, such as visual cueing. Furthermore, this manipulation could be considered more analogous to real world settings than cueing, as information encountered in everyday tasks might naturally differ in value. For instance, if a teacher is giving a series of classroom instructions, a child might know that the completion of one action is particularly important and may earn them some kind of reward, such as a sticker or some ‘table points’. Interventions could also be developed based on value manipulations; whereby more important information is associated with higher point values. Although such interventions are unlikely to enhance WM capacity, they might optimise memory efficiency by encouraging individuals to direct their attention towards particularly important information. Such an approach is not particularly feasible with visual cueing, as it would require the re-design of environments in order to implement the cues.

The thesis contains four experimental chapters, covering eight behavioural experiments. These will be outlined below.

1.9.1. Chapter 2: Do probe value and probe frequency manipulations encourage individuals to direct attention in similar ways?

As discussed above, terms such as ‘cueing’ and ‘prioritisation’ have been used to interchangeably to refer to boosts obtained from cues, probe frequency, and value-based manipulations. There is some evidence that probe value and cueing effects differ from each other (Hu et al., 2016; Souza et al., 2016). However, it is currently unclear whether probe value and probe frequency manipulations involve the same or distinct cognitive mechanisms. This chapter therefore reports two experiments which examined this.

In both experiments, participants were presented with coloured shapes and asked to recall the colour of one probed item after a brief delay. Experiment 1 manipulated probe value and probe frequency orthogonally, with both manipulations targeted at the same item. As part of the probe value manipulation, participants were either told that all items were equally valuable or that one item was worth more points. As part of the probe frequency manipulation, participants were informed that one item was more likely to be tested than the rest or that all items were equally likely to be assessed. If both effects reflect the same cognitive mechanisms, one might expect probe value effects to be smaller when the more valuable item is also more likely to be tested. However, if they reflect different cognitive mechanisms, the probe value and probe frequency effects should be independent with no interaction between them. Experiment 2 then aimed to strengthen these findings by investigating whether probe frequency effects are reliant on executive resources, as probe value effects in visual WM appear to be (Hu et al., 2016). Evidence that probe frequency effects are not reliant on executive resources would suggest that these boosts are more automatic in nature relative to probe value effects.

1.9.2. Chapter 3: Prioritising valuable information in verbal working memory

The rest of the thesis will focus on the probe value manipulation. Although several studies have demonstrated that individuals are able to direct their attention to more valuable information in visual WM (Hitch et al., 2018; Hu et al., 2014; 2016), research has not investigated whether individuals can also prioritise more valuable information in a purely auditory-verbal WM task. This chapter will report three experiments which examined this ability in young adults. An immediate serial recall task was employed, in which participants were auditorily presented with a series of digits and asked to recall them in the correct serial order following a brief delay. As with the probe value manipulation in Chapter 2, participants were either told that one item was more valuable than the rest or that all items were equally valuable.

Experiment 3 aimed to establish whether participants can direct their attention to more valuable information in a purely verbal WM task. Two follow-up experiments were then conducted to further elucidate the conditions in which individuals can prioritise in an immediate serial recall verbal task. Experiment 4 investigated whether individuals could prioritise information when rehearsal was disrupted, whilst Experiment 5 examined the ability when the ability to rehearse and draw on executive resources were both reduced.

1.9.3. Chapter 4: Can children prioritise more valuable information in working memory?

As discussed above, Berry et al. (2018) recently reported that 7-10-year-old children are not able to prioritise more valuable information in WM. This contrasts with findings from the cueing literature, in which children as young as 7-years-old have been found to benefit (Shimi et al., 2014; Shimi & Scerif, 2017). However, one possibility is that children did not show probe value effects in Berry et al. (2018) as they were not sufficiently motivated to do so. Moreover, probe value effects in children may vary as a function of memory load, with significant effects only observed when more items are presented. This would be in line with the cueing literature, in which effects appear to vary as a function of memory load in both adults (Astle et al., 2012; Kuo, Stokes, & Nobre, 2012; Nobre, Griffin, & Rao, 2008; Souza et al., 2014; van Moorselaar, Olivers, Theeuwes, Lamme, & Sligte, 2015) and children (Shimi & Scerif, 2017).

This chapter reports two experiments which investigated children's ability to prioritise more valuable information in WM. Seven to 10-year-old completed a cued-visual WM task, similar to that employed in Chapter 2. In Experiment 6, items were presented sequentially to closely mirror the design used by Berry et al. (2018). For the first time, Experiment 7 then investigated whether children are able to prioritise more valuable information during simultaneous presentation. As with the experiments in the previous chapters, participants were either told that one item was worth more points, or that all items were equally valuable. However, unlike in previous work, the memory task was embedded within a child-friendly context, with participants told that they could use the points collected to play a specially designed game at the end of the experiment. They were also told they would receive a prize if they collected a sufficient number of points. These features were added to particularly motivate the children to take account of the point values associated with each item and to adhere to the probe value instructions. For the first time, memory load was also manipulated to investigate whether this affects children's ability to prioritise important information in WM.

1.9.4. Chapter 5: Investigating the durability of probe value boosts

Previous research has investigated whether value affects WM (e.g. Allen & Ueno, 2018; Hitch et al., 2018; Hu et al., 2014; 2016; Sandry et al., 2014) and LTM (e.g. Adcock et al., 2006; Gruber & Otten, 2010; Shigemune et al., 2010). However, it is currently unclear whether prioritising an item for a WM test produces long-term benefits. It is also unclear whether testing a more valuable item at WM would yield larger prioritisation benefits at LTM. These questions were examined in this chapter. Participants encoded images of everyday objects for a WM test after a short delay. As in the previous experiments, they were told that one item was more valuable than the rest or that all items were of equal value. Following a series of filler tasks, participants completed a surprise LTM test in which memory for the everyday objects were assessed. Some of these items had been tested at WM, whilst others had not.

CHAPTER 2

DO PROBE VALUE AND PROBE FREQUENCY MANIPULATIONS INVOLVE THE SAME COGNITIVE MECHANISMS?

The experiments presented within this chapter are published as: Atkinson, A. L., Berry, E. D., Waterman, A. H., Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2018). Are there multiple ways to direct attention in working memory? *Annals of the New York Academy of Sciences*, *1424*(1), 115-126.

2.1. Introduction

As discussed in Chapter 1, terminology such as “cueing” and “prioritisation” have been used interchangeably to refer to the visual cueing, probe frequency, and probe value manipulations (e.g. Gorgoraptis et al., 2011; Klyszejko et al., 2014; Myers et al., 2018). Although these manipulations all encourage individuals to direct attention in WM, and can be considered conceptually similar (Allen & Ueno, 2018), the way in which the effects emerge may differ. Cueing typically involves participants directing their attention based on a visual stimulus (e.g. an arrow) that points towards or selects one of the to-be-remembered objects just before or after item presentation (Souza & Oberauer, 2016). This cue often informs participants which item will, or is most likely to, be tested at retrieval. In contrast, the probe frequency manipulation involves participants being told that an item associated with a specific feature (e.g. a shape, colour, location, or SP) is more likely to be assessed than the rest of the items (Cowan et al., 2010; Gorgoraptis et al., 2011; Klyszejko et al., 2014). Finally, in the probe value manipulation, participants are told that one item in the trial is worth a higher reward (e.g. more ‘points’) than the others (Berry et al., 2018; Hitch et al., 2018; Hu et al., 2014; 2016).

There is some existing evidence that probe value and visual cueing manipulations might encourage participants to direct their attention in different ways (Hitch et al., 2018). Hu et al. (2016) investigated whether probe value effects are reliant on executive resources. Participants were presented with series of coloured shapes sequentially, and asked to retrieve a feature of one of the shapes (e.g. colour) following a brief delay. Before the start of each block, participants were told that the value of each SP differed, with the first or the final item being worth the most notional points. During encoding and maintenance, participants either completed a simple concurrent task (repeating a two-digit number) or a more complex concurrent task (counting upwards in 2s). The probe value boost was substantially reduced or abolished in the complex concurrent task condition, suggesting that the effect is reliant on executive resources.

In contrast, evidence that cueing boosts are reliant on either effortful or automatic processing is less conclusive. Some studies have found that boosts are smaller when the cue does not predict the item that will be tested as accurately (i.e. cue validity is reduced). This has been taken as evidence that participants use cues strategically in order to enhance performance (Gunseli et al., 2015; Shimi et al., 2014; Souza & Oberauer, 2016). However, it is possible that visual cues might be attended to automatically when they are valid in all, or most cases, but less so when they predict the item that will be tested on a smaller percentage of the trials. There is also some debate whether retro-cue benefits are reliant on executive resources, with some studies finding that effects are reduced under concurrent task conditions (Janczyk & Berryhill, 2014) whilst others have found they are not (Hollingworth & Maxcey-Richard, 2013; Makovski, 2012). Furthermore, it has been suggested that the type of retro-cue might determine whether attention is directed strategically or automatically. Berryhill et al. (2012) suggested that central, semantic based retro-cues which identify a particular item (e.g. “red”) might be classed as strategic in nature, as individuals must encode the cue and then direct their attention based on it (but see Hommel, Pratt, Colzato, & Godijn, 2001). Conversely, visual cues that appear abruptly, just before or just after the set of to-be-remembered items might draw attention towards a particular location automatically (Berryhill et al., 2012; Schmidt et al., 2002; Woodman, Vecera, & Luck, 2003). If this cue appears at the spatial location of a to-be-remembered object, memory for this item might be automatically enhanced (Berryhill et al., 2012; Schmidt et al., 2002; Woodman et al., 2003). Finally, spatial symbols, such as arrow cues, might result in boosts due to both automatic and strategic processes (Berryhill et al., 2012). Individuals might strategically and voluntarily attend to the arrow cue if they know that the item presented at that spatial location will, or is likely to, be tested. However, there is also evidence that overlearned spatial symbols (such as arrows) might automatically attract attention towards particular locations, even when they are not predictive (Berryhill et al., 2012; Hommel et al., 2001). Given that arrow cues are very commonly used (Berryhill et al., 2012; Janczyk & Berryhill, 2014), the majority of retro-cue effects reported in the literature are likely to reflect a combination of automatic and strategic processes. This contrasts with probe value effects, which are considered to be reliant on executive resources and entirely strategic in nature (Hu et al., 2016).

There is also evidence that probe value and cueing effects differ in their vulnerability to interference. Several studies have demonstrated that probe value effects are reduced by the presence of a to-be-ignored visual suffix presented after encoding (Allen & Ueno, 2018; Hitch et al., 2018; Hu et al., 2014). This is particularly true if the suffix is plausible, and similar in nature to the set of to-be-remembered items (Hu et al., 2014). In direct contrast to this, there is some evidence that retro-cues *protect* the item from interference (Makovski & Jiang, 2007; Makovski, Sussman, & Jiang, 2008; Matsukura et al., 2007, van Moorselaar et al., 2014; Souza & Oberauer, 2016). For instance, some studies have found that the effects of interference (such as an irrelevant visual mask) are reduced after a retro-cue is presented, relative to trials where no such cue is presented (Makovski & Jiang, 2007; van Moorselaar et al., 2014).

Taken together, these studies provide converging evidence that probe value and visual cueing effects encourage participants to direct their attention in different ways. But what about probe value and probe frequency; do they involve the same or distinct cognitive mechanisms? As discussed in Chapter 1, little research has investigated the basis of probe frequency effects and whether the manipulation encourages individuals to direct their attention strategically or involve more automatic processes. However, existing literature provides some suggestions. The probe frequency manipulation can be considered similar to visual cueing, as both typically identify the item that is most likely to be tested at retrieval. However, the key difference is that cueing manipulations present a visual stimulus that appears on a trial-by-trial basis. In contrast, the probe frequency manipulation instead informs participants which feature (e.g. shape, colour, SP, etc) is more likely to be tested at the start of the block (Cowan et al., 2010; Gorgoraptis et al., 2011; Klyszejko et al., 2014). As discussed above, effects from arrow cues are considered to be at least partially automatic because the presentation of an overlearned visual symbol encourages participants to look towards a particular spatial location (Berryhill et al., 2012; Hommel et al., 2001). As the probe frequency manipulation does not present a visual stimulus, it would be reasonable to suggest that probe frequency effects are likely to be entirely strategic in nature (as existing research suggests probe value effects are; Hu et al., 2016).

However, it may be that repeatedly testing a particular feature primes attention towards it on future trials. For instance, if a particular colour (e.g. blue) is tested on 75% of trials, attention might be somewhat automatically directed towards that particular item because it is goal-relevant. The same process might also occur for other features, such as shape or SP. Further research is therefore needed in order to investigate this. Such findings would provide some insights into whether probe value and probe frequency manipulations involve the same or different cognitive mechanisms. Evidence that they involve different cognitive mechanisms would have important implications for the relationship between WM and attention by suggesting that studies employing these distinct manipulations should not be directly compared.

The current chapter reports two experiments which aimed to examine whether probe value and probe frequency encourage involve the same cognitive mechanisms, and whether probe frequency effects are relatively automatic or strategic in nature. The first experiment examined whether probe value and probe value effects independently enhance performance on a cued-recall visual WM task.

The second then examined whether probe frequency effects are reliant on executive resources, as probe value effects appear to be (Hu et al., 2016).

In line with most previous research in the area (e.g. Berry et al., 2018; Hitch et al., 2018; Hu et al., 2014; 2016; Sandry et al., 2014), sequential presentation was used in both experiments. This yields SP curves, which allows more fine-grained analysis of effects relative to a simultaneous mode of presentation (Hu et al., 2016). Using this paradigm, a recency effect is generally observed (e.g. Hitch et al., 2018; Hu et al., 2014; 2016). This is sometimes, but not always, accompanied by a small primacy effect (Hitch et al., 2018; Hu et al., 2014; 2016). For instance, Hu et al. (2014) found a recency effect regardless of whether participants were instructed to prioritise the first or final item. A primacy effect was, however, only observed when participants were encouraged to direct their attention towards the first item. Studies investigating prioritisation of information in visual WM have commonly targeted the first or the final item (e.g. Berry et al., 2018; Hu et al., 2014; 2016). In order to avoid the marked recency effect observed across studies, the manipulations used within both experiments were targeted at the first item. As the comparison in both experiments was between a differential probe value/probe frequency condition and the same SP in a control condition, any small primacy effect is likely to be present across conditions and is thus unlikely to confound the results.

2.2. Experiment 1

Participants completed a cued-recall visual WM task in which four coloured shapes were presented sequentially. After a brief delay, one shape was probed and participants had to verbally recall its colour. Probe value and probe frequency were manipulated orthogonally. As part of the probe value manipulation, participants were either told that the first item was worth more points than the rest (differential probe value) or that all items were worth the same number of points (equal probe value). As in previous research (Allen & Ueno, 2018; Berry et al., 2018; Hitch et al., 2018; Hu et al., 2014; 2016), the point system was notional. As part of the probe frequency manipulation, the first item was more likely to be tested than the rest (i.e. 70% of the time; differential probe frequency) or as likely to be tested as the other items (i.e. 25% of the time; equal probe frequency). The validity used in the differential probe frequency condition was similar to that used in previous research (e.g. Cowan et al., 2010: 80%; Gorgoraptis et al., 2011: 62.5%, Klyszejko et al., 2014: 75%), and also ensured that the other three non-targeted SPs were each tested the same number of times. Participants were told the frequency information before the start of the trial block. Both manipulations were targeted at the first SP as studies have consistently found significant probe value effects when this item is targeted (Hu et al., 2014; 2016).

Based on previous findings, it was predicted that no overall effect of probe value would be observed, but that there would be a significant interaction between probe value and SP. It was anticipated that this would, at least in part, be driven by significant probe value boosts at the first SP, whereby memory for this item would be higher in the differential probe value condition than the equal probe value condition (Allen & Ueno, 2018; Hu et al., 2014; 2016; Hitch et al., 2018). It was also predicted that probe frequency effects would emerge at SP1, with participants exhibiting higher accuracy for this item in the differential probe frequency condition than the equal probe frequency condition. Such findings would be in line with previous research which has indicated that individuals can direct attention in WM based on how frequently an item is tested (Cowan et al., 2010; Gorgoraptis et al., 2011; Klyszejko et al., 2014). However, previous studies have tested a particular colour (Klyszejko et al., 2014; Gorgoraptis et al., 2011) or shape (Cowan et al., 2010) more frequently than the other items, rather than a SP. The current study would therefore extend these findings by demonstrating that probe frequency effects can emerge based on an item's SP, as well as an object's physical feature (such as colour; Gorgoraptis et al., 2011; Klyszejko et al., 2014, or shape; Cowan et al., 2010).

Of particular interest was whether an interaction would emerge at SP1. Evidence of an interaction between probe value and probe frequency would indicate that these manipulations are not independent, and that the size of the probe value boost might differ depending on probe frequency. More specifically, evidence that probe value effects are smaller in the differential probe frequency condition would suggest that participants experience less benefit from increased probe value when they are already motivated to direct attention to this item. Such findings might be taken as evidence that probe value and probe frequency encourage individuals to direct attention in similar ways. Alternatively, probe value and probe frequency effects might be additive instead of interactive, suggesting that the manipulations might involve distinct cognitive mechanisms.

2.2.1. Method

2.2.1.1. Participants

Forty-four young adults took part (aged 18-30 years; Mean (*M)* = 20.42; *standard deviation* (*SD)* = 1.15; 9 males). Participants were native English speakers, had no known learning difficulties, and had normal or corrected-to-normal vision with no colour blindness. Participants were students at the University of Leeds, who were reimbursed for their time with course credits. None of the participants had completed Experiment 1. The experiment was approved by the School of Psychology Ethics Committee at the University of Leeds (Ethics reference number: 16-0303).

2.2.1.2. Materials

Four items were presented sequentially in each trial. Stimuli were created by randomly pairing a shape from a pool of six options (circle, triangle, arch, arrow, flag, cross) with a colour from a pool of six options (red, yellow, green, blue, purple, black). No shape or colour was repeated within the same trial. All stimuli subtended a visual angle of approximately 1.5°, based on a viewing distance of 50 centimetres (cm). Shapes were presented on a white background at one of eight points around a 2° imaginary circle positioned at the centre of the screen. Locations were selected randomly, with the constraint that no location could be used more than once per trial. The test cue was an outline of one of the stimuli presented during the encoding phase. This was displayed in the centre of the screen.

2.2.1.3. Design and procedure

The study employed a 2 x 2 x 4 within subject design, manipulating probe value (differential, equal), probe frequency (differential, equal) and SP (1-4). Participants completed four blocks of 40 trials; one for each combination of probe value and probe frequency. The order of probe frequency blocks and the order of probe value blocks within the probe frequency conditions was counterbalanced. In the equal probe frequency condition, the first SP was as likely to be tested as the other items (25% of the time; 10 trials). In the differential probe frequency condition, the first SP was assessed 70% of the time (28 trials), whilst each of the other three items were each probed 10% of the time (4 trials). The SPs tested were randomly distributed within the blocks. At the start of each block, participants completed four practise trials to familiarise themselves with the condition.

Each condition commenced with the provision of written instructions. In the differential probe value conditions, participants were told that correct recall of the first item would earn them four points, whilst correct recall of the other items would earn them one point. In the equal probe value conditions, they were told each item was worth one point. The points were part of a notional reward system, with the number of points accrued not tallied or related to an actual reward. During the instructions, participants were also informed about the probe frequency manipulation. This was implemented to ensure that all participants were equally aware that one of the items would be tested more frequently than the rest, and that there would not be individual differences in the number of trials taken to detect a pattern. In the equal probe frequency conditions, participants were told that all items would be tested the same number of times. In the differential probe frequency condition, they were told that the first item would be tested more often than the other items, although they were not told the precise percentage of times it would be assessed.

The experimental paradigm used is displayed in Figure 2.1. Each trial began with presentation of the word ‘la’, which participants were asked to whisper until the retrieval phase to disrupt verbal recoding (Baddeley, 1986). Following a key press, a fixation cross appeared for 500ms, which informed participants that the shapes were about to appear. Next, four coloured shapes were displayed sequentially. Each shape displayed for 500ms, with an interstimulus interval (ISI) of 250ms. After a delay of 1000ms, the outline of one shape was presented and participants verbally recalled the original colour of the shape. Their response was recorded by the experimenter, who then pressed the space bar to progress onto the next trial. Participants were reminded of the probe value and probe frequency instructions after every ten trials. Participants were not given feedback regarding performance on the task.

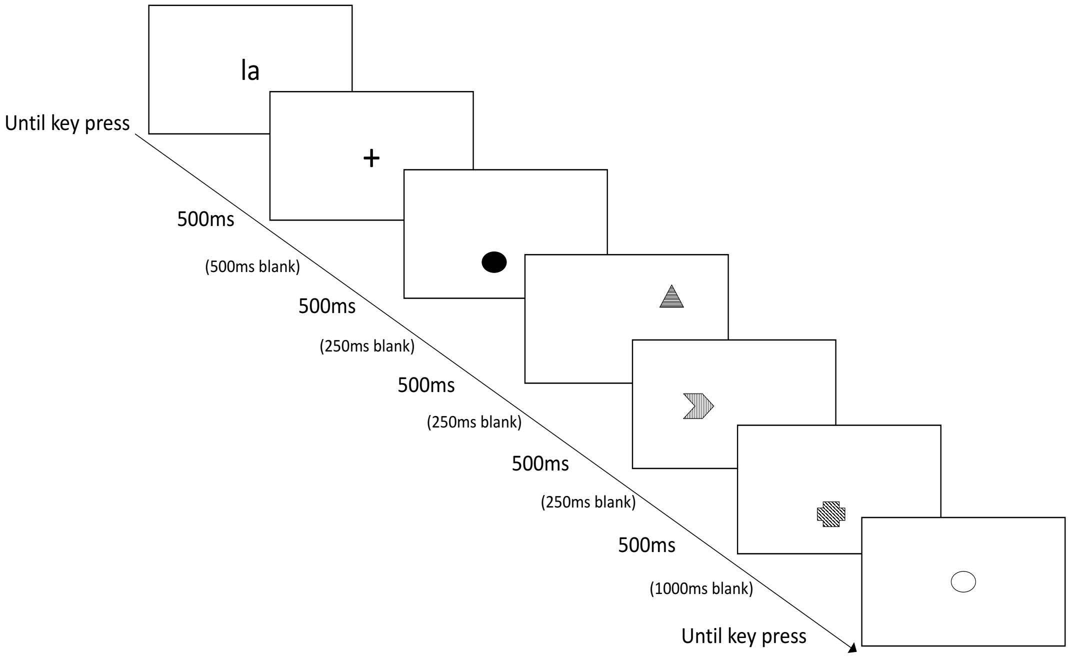


Figure 2.1. The experimental paradigm used in Experiment 1. The different types of shading reflect different colours (e.g. red, green, blue). Experiment 2 was identical, except that participants were presented with a two-digit number prior to encoding rather than the word “la”. Figure not to scale.

2.2.1.4. Data analysis

The data within this thesis was analysed using frequentist methods, such as analysis of variance (ANOVA), t-tests, and correlations. However, several issues have recently been raised with the exclusive use of p-values (Lakens, McLatchie, Isager, Scheel, & Dienes, 2018; Wasserstein & Lazar, 2016). For instance, these tests often lead to a dichotomisation, whereby values lower than 0.05 are taken as evidence for an effect, whereas values larger than 0.05 are not (Dhaliwal & Campbell, 2010; Wasserstein & Lazar, 2016). This dichotomisation is problematic, as it ignores any uncertainty that is inherent in statistical analysis (Dhaliwal & Campbell, 2010; Wasserstein & Lazar, 2016). Furthermore, obtaining a p-value larger than 0.05 does not provide evidence of no effect (Barchard, 2015; Rouder, Morey, Speckman, & Province, 2012), although this is often illogically concluded (Lakens et al., 2018). This therefore prevents one from concluding that there is no difference between conditions or group (Barchard, 2015); a question that may be of interest in some cases.

Several suggestions have been posited to remedy the shortcomings of p-values (Benjamin et al., 2018; Cumming, 2014; García & Puga, 2018; Goodman, 2001; Halsey, 2019; Lakens et al., 2018; Quatto, Ripamonti, & Marasini, 2019; Quintana & Williams, 2018). One suggestion has been to report BFs alongside p-values in order to provide the reader with additional information (García & Puga, 2018; Goodman, 2001; Halsey, 2019; Lakens et al., 2018; Quatto et al., 2019; Quintana & Williams, 2018). BF analysis assesses the strength of evidence for the alternative hypothesis relative to the null hypothesis (Leppink, O'Sullivan, &

Winston, 2017; Wagenmakers et al., 2018), and can therefore be interpreted as the weight of evidence for or against an effect (Faulkenberry, 2018). BF analysis also provides a test of equivalence between conditions or groups, thus allowing one to conclude that a manipulation has no effect on an outcome measure (Barchard, 2015; Lakens et al., 2018; Wagenmakers et al., 2018).

BF analysis was therefore performed alongside the frequentist analysis. This was conducted in R (R Core Team, 2016) using the BayesFactor package (Morey et al., 2018). Bayesian ANOVAs were run using the default priors (Rouder et al., 2012), with the number of iterations set at 500,000. Setting the number of iterations to such a large number increases the precision of the estimates (Rouder, Morey, Verhagen, Swagman, & Wagenmakers, 2017). All possible models were assessed, which means that a model could contain an interaction even if the main effects were absent. As appropriate, follow-up BF t-tests were then conducted. When reporting the BF analysis, the most likely model given the data is described relative to the null model including only random effects of participant. BFs for all main effects and interactions are also reported. If the effect or interaction was included in the most likely model, the BF was calculated by comparing the most likely model to a model excluding that effect. If the effect or interaction was not included in the most likely model, the BF was calculated by comparing the most likely model to a model including all of the effects featured in the most likely model plus the effect of interest. BF_{10} values describe how many times more likely the alternative hypothesis is to the null hypothesis. For instance, a BF_{10} value of 6.21 states that the alternative hypothesis (of an effect) is 6.21x more likely than the null hypothesis (of no effect). BF_{10} values larger than 1 provide evidence of an effect, whilst values lower than 1 provide evidence of no effect. Naturally, BFs occur on a continuous scale (Lee &

Wagenmakers, 2014), which avoids issues such as dichotomisation that frequently occurs with p-values (Wasserstein & Lazar, 2016). However, some guidelines have been created in order to allow researchers to evaluate the level of evidence for or against an effect (Lee & Wagenmakers, 2014; Jeffreys, 1961). Lee and Wagenmakers' (2014) guidelines (adapted from Jeffreys (1961) guidelines) are displayed in Table 2.1.

As suggested by Morey and Rouder (2011), BF_{01} values are also reported when the BF analysis indicates favour of no effect (i.e. the BF_{10} value is less than 1). BF_{01} values reflect how many times more likely the null hypothesis (of no effect) is relative to the alternative hypothesis (of an effect). This is calculated by taking the inverse of BF_{10}. For instance, if the BF analysis indicates that BF_{10} is equal to 0.21, the reciprocal can be taken to produce the BF_{01} value ($1/0.21 = 4.76$). This indicates that the null hypothesis (of no effect) is 4.76x more likely than the alternative hypothesis (of an effect).

A further suggestion is that p-values and BFs should be supplemented with effect size. This provides a measure of the magnitude of an effect and allow one to assess the practical significance of findings (Wasserstein & Lazar, 2016; Wetzels et al., 2011). Effect sizes are therefore reported alongside all effects.

In the current chapter, the dependent variable was accuracy, determined by the proportion of trials where participants responded correctly (the number of trials responded correctly divided by the number of trials responded correctly plus the number of trials responded incorrectly). Findings are first reported across SP. Further planned analysis was then conducted at SP1, as this is the position at which the manipulations were targeted.

Table 2.1.

Guidelines for interpreting Bayes Factors, set out by Lee & Wagenmakers (2014)

BF_{10} value	Interpretation
100+	Extreme evidence for an effect
30-100	Very strong evidence for an effect
10-100	Strong evidence for an effect
3-10	Moderate evidence for an effect
1-3	Anecdotal evidence for an effect
1	No evidence
0.33-1	Anecdotal evidence against an effect
0.1-0.33	Moderate evidence against an effect
0.03-0.1	Strong evidence against an effect
0.01-0.1	Very strong evidence against an effect
<0.01	Extreme evidence against an effect

Note: these guidelines were based on Jeffrey (1961) guidelines, with a few alterations. In Jeffreys (1961), the “anecdotal” label was termed “not worth more than a bare mention”, the “moderate” label was termed substantial, and the “extreme” label was termed “decisive”.

2.2.2. Results

2.2.2.1. Across serial positions

Mean accuracy as a function of probe value, probe frequency, and SP is displayed in Figure 2.2. A 2 (Probe value: differential vs equal) x 2 (Probe frequency: differential vs equal) x 4 (SP: 1-4) within-subjects ANOVA revealed no main effect of probe value (*Differential M* = 0.52, standard error (*SE)* = 0.02; *Equal M* = 0.53, *SE* = 0.02;

($F(1, 43) = 0.38, p = .54$, mean square error (*MSE)* $= 0.038, \eta_p^2 < .01$; $BF_{10} = 0.10$, $BF_{01} = 10.00$) or probe frequency (*Differential* $M = 0.51, SE = 0.02$; *Equal* $M = 0.54$, $SE = 0.02$; $F(1, 43) = 2.57, p = .12, MSE = 0.063, \eta_p^2 = .056$; $BF_{10} = 0.51, BF_{01} = 1.96$), demonstrating that neither manipulation affected overall performance on the task. A main effect of SP emerged ($F(3, 129) = 40.07, p < .001, MSE = 0.064, \eta_p^2 = .48$; $BF_{10} > 10{,}000$). Pairwise comparisons (corrected using Bonferroni-Holm) revealed significant differences between SP1 ($M = 0.66, SE = 0.02$) and SP2 ($M = 0.42, SE = 0.02$; $p < .001$), SP1 and SP3 ($M = 0.43, SE = 0.02$; $p < .001$), SP1 and SP4 ($M = 0.59, SE = 0.03$; $p = .034$), SP2 and SP4 ($p < .001$) and SP3 and SP4 ($p < .001$). Significant interactions emerged between probe value and SP ($F(3, 129) = 25.01, p < .001, MSE = 0.036, \eta_p^2 = .37$; $BF_{10} > 10{,}000$), and probe frequency and SP (*Greenhouse-Geisser (GG) corrected* $F(2.54, 109.31) = 19.15, p < .001, MSE = 0.042, \eta_p^2 = .31$; $BF_{10} > 10{,}000$), indicating that the effects of probe value and probe frequency differed depending upon the SP tested. No interactions emerged between probe value and probe frequency *(*$F(1, 43) = 0.11, p = .74, MSE = 0.032, \eta_p^2 < .01$; $BF_{10} = 0.13, BF_{01} = 7.69$) or probe value, probe frequency and SP (*GG corrected* $F(2.51, 107.82) = 0.25, p = .83, MSE = 0.037, \eta_p^2 < .01$; $BF_{10} = 0.04, BF_{01} = 25.00$). These outcomes were corroborated by BF analysis, which revealed that the most likely model included a main effect of SP, as well as interactions between probe value and SP, and probe frequency and SP ($BF_{10} > 10{,}000$ relative to the null model with random effects of participant only),

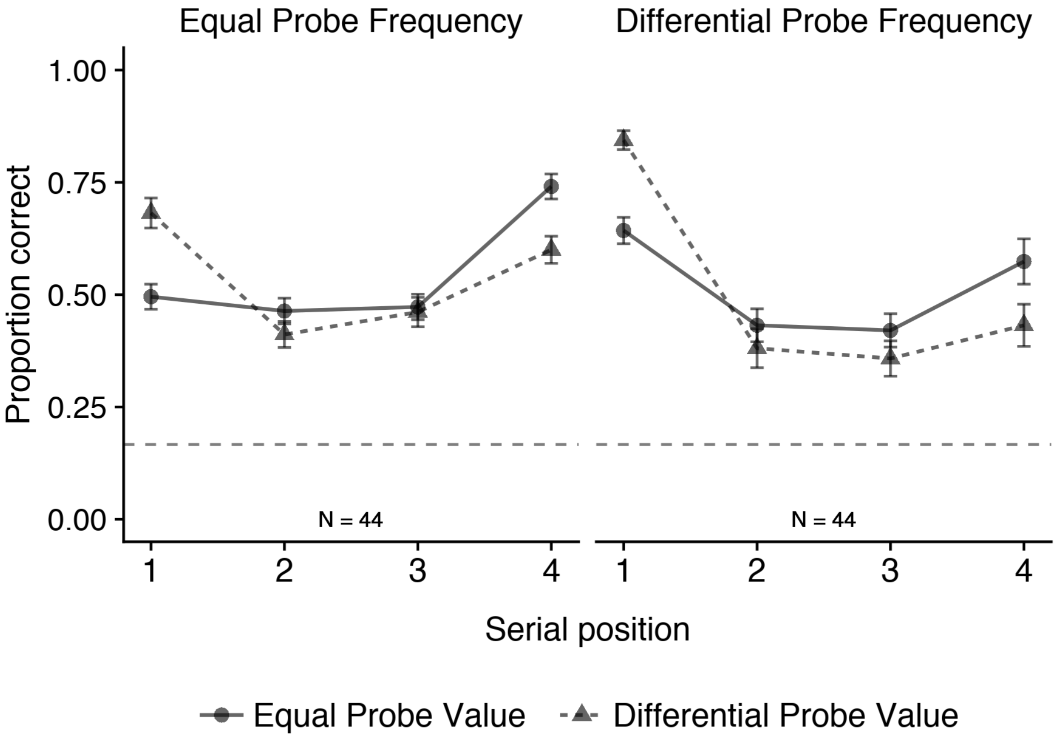


Figure 2.2. Mean proportion correct in Experiment 1, as a function of probe value, probe frequency, and SP. Error bars denote SE.

To investigate the interaction between probe value and SP, a series of paired sample t-tests (corrected using Bonferroni-Holm) were conducted. Mean proportion correct as a function of probe value and SP is displayed in Table 2.2. At SP1, participants performed significantly better in the differential probe value condition ($t(43) = 8.85$, $p < .001$, *Cohen's d* $(d) = 1.33$; $BF_{10} > 10{,}000$). The pattern was reversed at SP4, with participants exhibiting higher accuracy in the equal probe value condition ($t(43) = -4.08$, $p < .001$, $d = -0.61$; $BF_{10} = 130.60$). No significant differences between probe value conditions were found at SP2 ($t(43) = -1.71$, $p = 0.19$, $d = -0.26$; $BF_{10} = 0.62$; $BF_{01} = 1.64$) or SP3 ($t(43) = -1.38$, $p = 0.19$, $d = -0.21$; $BF_{10} = 0.39$; $BF_{01} = 2.56$). In summary, this indicates that increasing the value of the first item boosted performance at SP1, had no significant effect on performance at SP2 and SP3, and negatively affected performance at SP4.

Table 2.2.

Mean accuracy (and SE) in Experiment 1 as a function of probe value and SP, collapsed across probe frequency conditions.

	SP1	SP2	SP3	SP4	Across SPs
Differential probe value	.76 (.02)	.40 (.03)	.41 (.03)	.52 (.03)	.52 (.02)
Equal probe value	.57 (.02)	.45 (.02)	.45 (.03)	.66 (.03)	.53 (.02)

A series of paired sample t-tests (corrected using Bonferroni-Holm) were also conducted to investigate the interaction between probe frequency and SP. Mean proportion correct as a function of probe frequency and SP is displayed in Table 2.3. At SP1, higher accuracy was observed in the differential probe frequency condition ($t(43) = 6.05, p < .001, d = 0.91; BF_{10} > 10{,}000$). The opposite pattern of results was observed at SP3 and SP4, with participants exhibiting significantly higher accuracy in the equal probe frequency condition (SP3*:* $t(43) = -2.47, p = .035, d = -0.37; BF_{10} = 2.45$; SP4*:* $t(43) = -4.51, p < .001, d = -0.68; BF_{10} = 450.47$). No effect of probe frequency emerged at SP2 ($t(43) = -0.87, p = .39, d = -0.13; BF_{10} = 0.23; BF_{01} = 4.35$). To summarise, this demonstrates that increasing the likelihood of the first item being assessed enhanced accuracy at SP1, had no significant effect at SP2, and reduced accuracy at SP3 and SP4.

Table 2.3.

Mean accuracy (and SE) in Experiment 1 as a function of probe frequency and SP, collapsed across probe value conditions.

	SP1	SP2	SP3	SP4	Across SPs
Differential probe frequency	.74 (.02)	.41 (.03)	.39 (.03)	.50 (.04)	.51 (.02)
Equal probe frequency	.59 (.03)	.44 (.02)	.47 (.03)	.67 (.03)	.54 (.02)

2.2.2.2. Serial position 1

As both manipulations were targeted at SP1, further analysis was conducted at this SP to explore whether an interaction emerged between probe value and probe frequency. Mean performance for individual participants, as well as aggregated performance across participants in shown in Figure 2.3, as a function of probe value and probe frequency. A 2 (Probe value: differential vs equal) x 2 (Probe frequency: differential vs equal) within-subjects ANOVA revealed a significant main effect of probe value ($F(1, 43) = 78.28, p < .001, MSE = 0.021, \eta_p^2 = .65$; $BF_{10} > 10{,}000$), with participants exhibiting higher accuracy in the differential probe value condition. There was also a significant main effect of probe frequency ($F(1, 43) = 36.57, p < .001, MSE = 0.029, \eta_p^2 = .46$; $BF_{10} > 10{,}000$), with participants exhibiting higher accuracy in the differential probe frequency condition. No significant interaction emerged between probe value and probe frequency ($F(1, 43) = 0.17, p = .69, MSE =$

0.015, $\eta_p^2 < .01$; $BF_{10} = 0.24$, $BF_{01} = 4.17$), suggesting that probe value boosts do not differ depending on probe frequency and that the manipulations have independent effects. BF analysis supported these conclusions, with the most likely model including main effects of probe value and probe frequency ($BF_{10} > 10{,}000$ relative to the null model with random effects of participant only). [1]

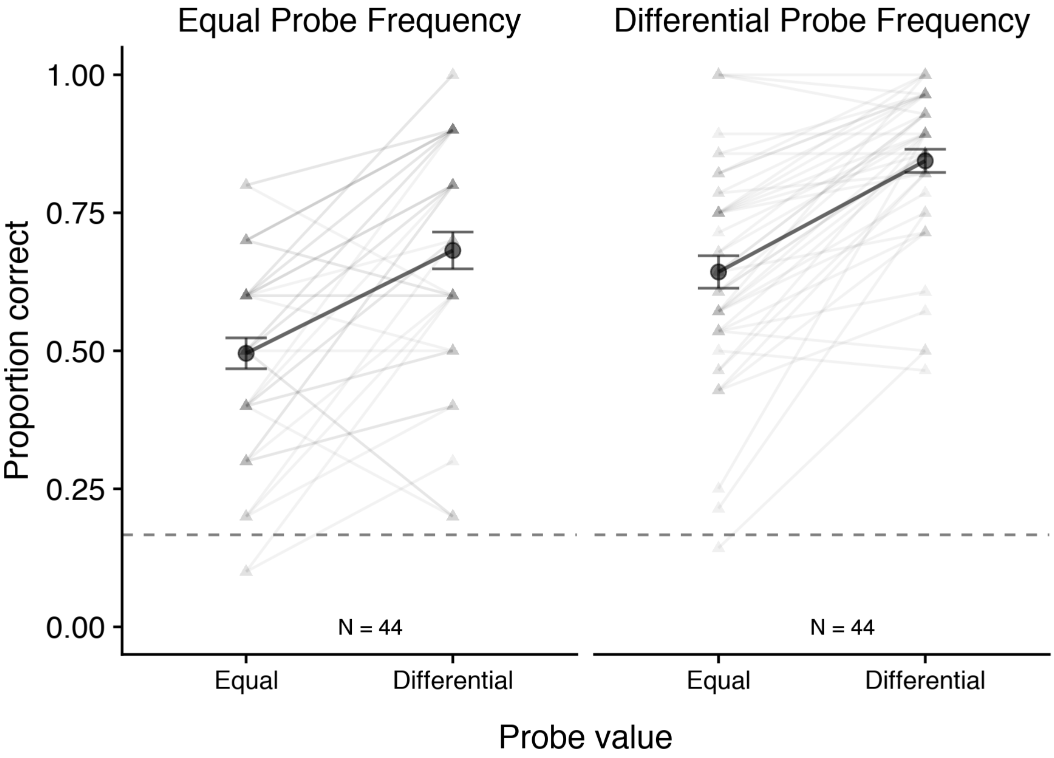


Figure 2.3. Mean performance at SP1 for individual participants as a function of probe value and probe frequency. The lighter grey lines with triangular points display mean performance for individual participants. Meanwhile, the darker grey line reflect the mean across participants, whilst the error bars denote SE.

---

[1] As shown in Figure 2.3, three participants exhibited at ceiling performance (i.e. 100%) at SP1 in the condition where this item was more likely to be tested, but not more valuable. To ensure that these participants did not prevent the emergence of an interaction between probe value and probe frequency at SP1, the analysis was re-ran excluding these participants. There were no differences to the main conclusions.

2.2.3. Discussion

In line with previous research (Allen & Ueno, 2018; Hitch et al., 2018; Hu et al., 2014; 2016), a significant probe value effect emerged, whereby participants exhibited better performance at SP1 in the differential probe value condition than the equal probe value condition. This therefore provides further evidence that individuals can direct their attention to more valuable items in a cued-recall visual WM task. Significant probe frequency effects were also observed, with participants exhibiting higher accuracy at SP1 in the differential probe frequency condition relative to the equal probe frequency condition. This extends previous findings, by demonstrating that probe frequency effects can emerge based on an item's SP, as well as an object's physical feature (such as colour; Gorgoraptis et al., 2011; Klyszejko et al., 2014, or shape; Cowan et al., 2010). There was no overall effect of probe value and probe frequency, with both manipulations coming at a cost to some of the other items. This suggests that neither manipulation increases WM capacity. Evidence that costs emerged in the probe frequency condition should, however, be treated with caution, as there were only a small number of trials testing SPs 2-4 in the differential probe frequency conditions.

Importantly, no significant interaction emerged between probe value and probe frequency across SPs or at the SP in which the manipulations were targeted (SP1). BF analysis also provided evidence of no interactions. This indicates that probe value effects are not affected by the frequency with which the more valuable item is assessed at retrieval. This contrasts with findings from the cueing literature, in which benefits appear to depend on the validity of the cues (Gunseli et al., 2015; Shimi et al., 2014). Perhaps more importantly, this indicates that probe value and probe frequency are independent in their impacts on working memory, and that they might involve distinct cognitive mechanisms. Evidence for this is preliminary, however, warranting additional research to explore this possibility further.

Previous research has suggested that probe value effects are reduced if participants engage in an attentionally-demanding concurrent task during encoding and maintenance (Hu et al., 2016). It would be useful to establish whether such a task also reduces probe frequency boosts. This would provide further insights into probe frequency effects, whilst also further exploring whether probe value and probe frequency are likely to involve distinct cognitive mechanisms. Experiment 2 therefore investigated this.

2.3. Experiment 2

Experiment 2 aimed to examine whether probe frequency effects in WM are reliant on executive resources, as probe value effects appear to be (Hu et al., 2016). As no visual stimulus is presented in the probe frequency manipulation (unlike the cueing paradigm), it would be reasonable to predict that the effect might be obtained strategically. However, it might be that repeatedly testing a particular feature automatically directs participants attention towards it.

This research question has not been explored to date, although a series of studies have suggested that individuals might automatically encode frequency information during memory tasks (Hasher & Zacks, 1979). This is based on evidence that the accuracy of frequency judgements are unaffected by age (Attig and Hasher, 1980), intentionality (Hasher & Chromiak, 1977; Hasher, Zacks, Rose, & Sanft, 1987,; Zacks, Hasher, & Sanft, 1982), or the completion of a concurrent task (Zacks et al., 1982; but see Jonides, Naveh-Benjamin, & Palmer, 1985; Naveh-Benjamin & Jonides, 1985). Furthermore, evidence from amnesic patients suggests that this group can use recurring patterns to enhance performance on motor tasks, despite not being explicitly aware that a pattern is being repeated (Nissen, Willingham & Hartman, 1989). This suggests that individuals might also apply frequency information automatically. As such, probe frequency boosts in WM might occur in a relatively cost-free manner, placing minimal reliance on executive resources. Such findings would provide evidence of a dissociation between probe value and probe frequency, providing further evidence that the manipulations might involve distinct cognitive mechanisms. This was therefore explored in Experiment 2.

A 2 (Probe frequency: equal, differential) x 2 (Concurrent task: simple, complex) x 4 (SP: 1-4) within-subjects design was employed. To allow comparisons with the probe value literature, the concurrent tasks used were identical to Hu et al. (2016). In the simple concurrent task condition, participants repeated a two-digit number during encoding and maintenance. In the complex concurrent task conditions, participants counted upwards in steps of two during these phases. As in Experiment 1 and several previous studies (Hu et al., 2014; 2016), the probe frequency manipulation was targeted at the first SP. Performance at this position was therefore of particular interest. Evidence of an interaction between probe frequency and concurrent task at SP1, with performance in the differential probe frequency condition particularly affected by an increase in concurrent task complexity, would suggest that boosts are reliant on executive control. These outcomes would be in line with findings from the probe value literature (Hu et al., 2016). Conversely, evidence of no interaction between probe frequency and concurrent task would suggest that effects are not reliant on executive resources, contrasting with findings on probe value (Hu et al., 2016). Such outcomes would suggest that the manipulations might involve distinct cognitive mechanisms, further supporting the conclusions drawn from Experiment 1.

2.3.1. Method

2.3.1.1. Participants

Twenty-four young adults participated (aged 18-35 years; *M* = 22.11, *SD* = 3.58; 10 males). Participants were either paid or given course credit. The experiment was approved by the School of Psychology Ethics Committee at the University of Leeds Ethics reference number: 16-0303).

2.3.1.2. Materials

The materials used were identical to Experiment 1, except that participants were presented with a randomly selected number between 20-99 at the start of each trial, as opposed to the word ‘la’.

2.3.1.3. Design and procedure

The study employed a 2 x 2 x 4 within-subjects design, manipulating probe frequency (equal, differential), concurrent task (simple, complex) and SP (1-4). Participants completed four blocks of 40 trials; one for each combination of probe frequency and concurrent task. The order of probe frequency blocks and the order of concurrent task blocks within the probe frequency conditions were counterbalanced. The SPs tested were randomly distributed within the blocks.

Participants were told that correctly recalling the colour of the shape tested would gain them one point. The probe frequency instructions were the same as in Experiment 1. The experimental paradigm was also identical to Experiment 1, except that participants were presented with a number between 20-99 at the start of the trial as opposed to the word ‘la’. In the simple concurrent task conditions, participants were asked to repeat the number until retrieval. In the complex concurrent task conditions, participants were asked to count upwards in steps of two from the number until the retrieval phase (e.g. 45, 47, 49). This concurrent task manipulation is identical to that employed by Hu et al. (2016).

2.3.1.4. Data analysis

The dependent variable was accuracy, determined by proportion correct (the number of trials responded correctly divided by the number of trials responded correctly plus the number of trials responded incorrectly). Findings are first reported across SPs, followed by further planned analysis at SP1.

2.3.2. Results

2.3.2.1. Across serial positions

Mean proportion correct as a function of probe frequency, SP, and concurrent task is displayed in Figure 2.4. A 2 (Probe frequency: differential vs equal) x 2 (Concurrent task: simple vs complex) x 4 (SP: 1-4) within-subjects ANOVA revealed no significant main effect of probe frequency (*Differential* $M = 0.48, SE = 0.03$; *Equal* $M = 0.50, SE = 0.02$; $F(1, 23) = 0.39, p = .54, MSE = 0.056, \eta_p^2 = .016$; $BF_{10} = 0.14$, $BF_{01} = 7.14$), indicating that increasing the likelihood of the first item being assessed did not affect overall performance on the task. There was, however, a main effect of concurrent task, ($F(1, 23) = 77.86, p < .001, MSE = 0.023, \eta_p^2 = .77$; $BF_{10} >$ 10,000), with higher accuracy in the simple concurrent task condition ($M = 0.56, SE = 0.02$) relative to the complex concurrent task condition ($M = 0.42, SE = 0.21$). There was also a significant main effect of SP (*GG corrected* $F(2.19, 50.39) = 10.48, p < .001$, $MSE = 0.081, \eta_p^2 = .31$; $BF_{10} >$ 10,000), with pairwise comparisons (corrected using Bonferroni-Holm) revealing significant differences at SP1 ($M = 0.51, SE = 0.03$) and SP2 ($M = 0.39, SE = 0.03; p = .012$), SP2 and SP4 ($M = 0.61, SE = 0.04; p < .001$) and SP3 ($M = 0.45, SE = 0.03$) and SP4 ($p < .001$). A significant interaction emerged between probe frequency and SP ($F(3, 69) = 17.79, p < .001, MSE = 0.052, \eta_p^2 = .44$; $BF_{10} >$ 10,000), indicating that the effects of probe frequency differed depending on the SP tested. In contrast, there was no interaction between concurrent task and SP ($F(3, 69) = 1.31, p = .28, MSE = 0.037, \eta_p^2 = .054$; $BF_{10} = 0.09, BF_{01} = 11.11$). There were also no interactions between probe frequency and concurrent task ($F(1, 23) <$ $0.01, p = .96$, $MSE = 0.034$, $\eta_p^2 < .01$; $BF_{10} = 0.16, BF_{01} = 6.25$), and probe frequency, concurrent task, and SP ($F(3, 69) = 0.27, p = .85, MSE = 0.033, \eta_p^2 =$ $.012$; $BF_{10} = 0.07, BF_{01} = 14.29$). These findings were corroborated by BF analysis, which revealed that the most likely model included main effects of concurrent task and SP, as well as an interaction between probe frequency and SP ($BF_{10} > 10{,}000$ relative to the null model with random effects of participant only).

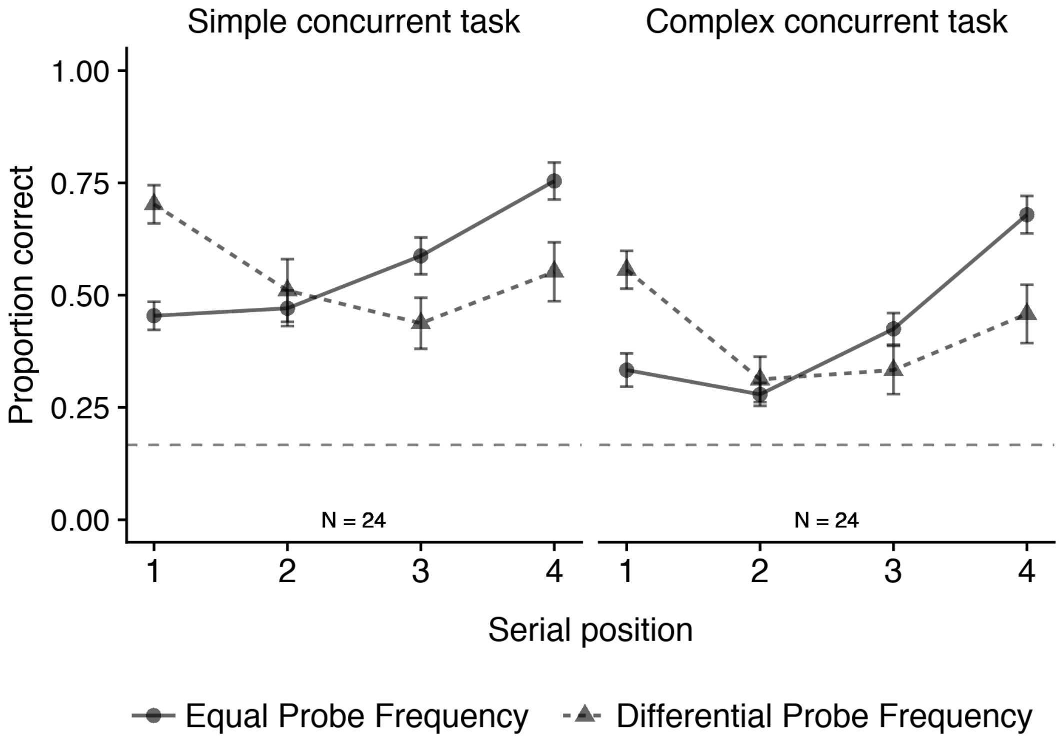


Figure 2.4. Mean accuracy in Experiment 2, as a function of probe frequency, concurrent task, and serial position. Error bars denote SE.

To investigate the interaction between probe frequency and SP, a series of t-tests (corrected using Bonferroni-Holm) were conducted. Accuracy as a function of probe frequency and SP is displayed in Table 2.4. Accuracy was higher in the differential probe frequency condition at SP1 ($t(23) = 6.63, p < .001, d = 1.35$; $BF_{10} > 10{,}000$). This pattern was reversed at SP3 ($t(23) = -2.78, p = .021, d = -0.57$; $BF_{10} = 4.63$) and SP4 ($t(23) = -3.65, p = .004, d = -0.74$; $BF_{10} = 27.09$), with participants exhibiting higher accuracy in the equal probe frequency condition. No significant effect of probe frequency emerged at SP2 ($t(23) = 0.76, p = .46, d = 0.15$; $BF_{10} =$

0.28, BF_{01} = 3.57). To summarise, these outcomes indicate that increasing the likelihood of the first item being tested enhanced memory at SP1, had no significant effect at SP2, and impaired memory at SP3 and SP4.

Table 2.4.

Mean accuracy (and SE) in Experiment 2 as a function of probe frequency and SP, collapsed across concurrent task conditions.

	SP1	SP2	SP3	SP4	Across SPs
Differential probe frequency	.63 (.04)	.41 (.05)	.39 (.05)	.51 (.06)	.48 (.03)
Equal probe frequency	.39 (.02)	.38 (.03)	.51 (.03)	.72 (.04)	.50 (.02)

2.3.2.2. Serial position 1

As in Experiment 1, further planned analysis was conducted at SP1 to explore whether an interaction emerged between probe frequency and concurrent task. Mean performance for individual participants (as well as mean performance across participants) is displayed in Figure 2.5. A 2 (Probe frequency: differential vs equal) x 2 (Concurrent task: simple vs complex) within-subjects ANOVA revealed a main effect of probe frequency ($F(1, 23) = 43.89, p < .001, MSE = 0.03, \eta_p^2 = .66$; $BF_{10} >$ 10,000), with participants exhibiting higher accuracy in the differential probe frequency condition relative to the equal probe frequency condition. There was also a main effect of concurrent task ($F(1, 23) = 18.66, p < .001, MSE = 0.023, \eta_p^2 = .45$;

BF_{10} = 438.42), with participant exhibiting higher accuracy in the simple concurrent task condition. There was, however, no significant interaction between probe frequency and concurrent task ($F(1, 23) = 0.24, p = .63, MSE = 0.016, \eta_p^2 = .01$; BF_{10} = 0.31, BF_{01} = 3.23), suggesting that the probe frequency boosts observed were not affected by concurrent task. BF analysis revealed that the most likely model included main effects of probe frequency and concurrent task (BF_{10}> 1000 relative to the null model with random effects of participant only).

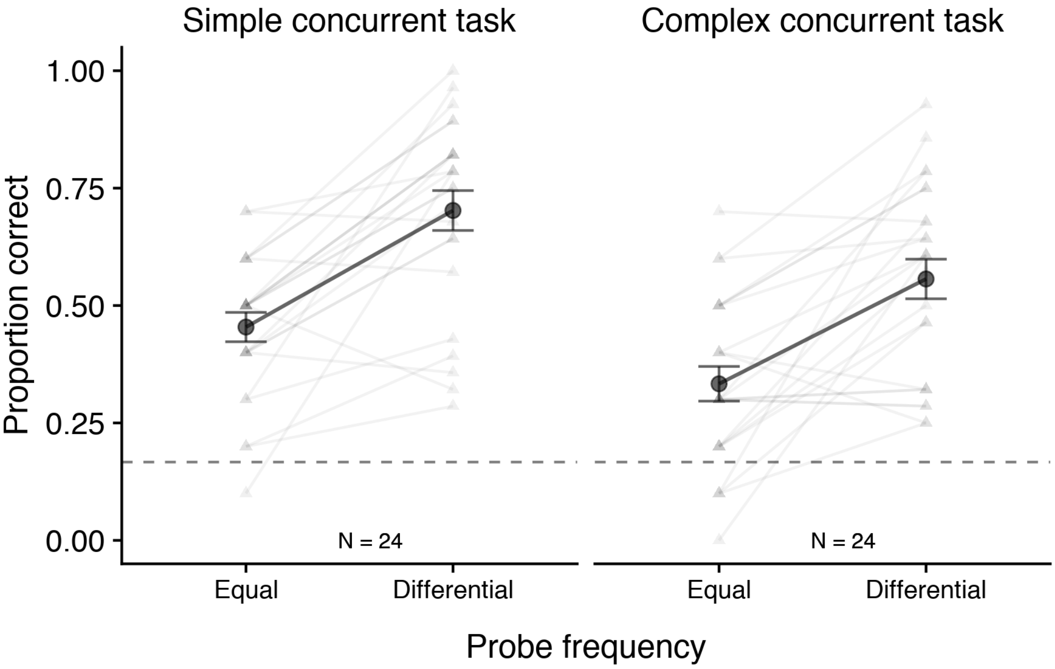


Figure 2.5. Mean performance at SP1 in Experiment 2 for individual participants, as a function of probe frequency and concurrent task. The lighter grey lines with triangular points reflect mean performance for individual participants, whilst the darker grey line with circular points displays mean accuracy across participants. Error bars denote SE.

2.3.3. Discussion

Replicating Experiment 1, significant probe frequency effects were observed at SP1. Importantly, however, there was no interaction between probe frequency and concurrent task at SP1, suggesting that probe frequency effects are not reduced by an attentionally-demanding concurrent task. These outcomes suggest that probe frequency boosts are not reliant on executive resources during encoding and maintenance, and might be obtained in a relatively cost-free and automatic manner (Hasher & Zacks, 1979; 1984; Zacks & Hasher, 2002). This directly contrasts with research exploring probe value (Hu et al., 2016), which has revealed that these effects are significantly reduced or abolished when participants complete an attentionally-demanding concurrent task. This provides further evidence that the probe value and probe frequency manipulations might involve distinct cognitive mechanisms.

Evidence that probe frequency boosts are not reliant on executive resources is in line with some findings from the retro-cue literature (Hollingworth & Maxcey-Richard, 2013; Makovski, 2012), which has reported that effects are not reduced when participants completed a concurrent task. It does, however, contrast with Janczyk and Berryhill (2014) who found that retro-cue effects are significantly reduced if participants engage in an attentionally-demanding concurrent task during cue onset and encoding.

Increased probe frequency came at a cost to some SPs that were less likely to be tested (SP3 and SP4). Such findings are in line with previous studies, which have found that enhancements to a targeted item can come at a cost to other items presented within the same trial (Astle et al., 2012; Chun et al., 2011; Gressmann &

Janczyk, 2016; Gunseli et al., 2015; Hitch et al., 2018; Hu et al., 2014; 2016; Pertzov et al., 2013; Rerko et al., 2014).

Across conditions, accuracy at SP4 was significantly higher than the other SPs that were less likely to be assessed (SP2 and SP3), supporting previous findings that this item holds a privileged status within WM (Berry et al., 2018; Hu et al., 2014). As in Experiment 1, these outcomes should, however, be interpreted with caution, as participants completed only a small number of trials at SPs 2-4 in the differential probe frequency conditions.

2.4. General discussion

This chapter reports a pair of experiments which explored how attention can be directed in a cued-recall visual WM task. Of particular interest was whether probe value boosts are affected by the probability with which the more valuable item is tested at retrieval (i.e. probe frequency) or whether these manipulations yield independent effects. In Experiment 1, probe value and probe frequency boosts were observed, although the two effects were additive. This demonstrates that probe value boosts are not affected by probe frequency, which contrasts with findings from the retro-cue literature (Gunseli et al., 2015; Shimi et al., 2014). It also provides some preliminary evidence that probe value and probe frequency manipulations have independent effects on WM performance. This latter finding was further supported by Experiment 2, which indicated that probe frequency boosts in cued-recall visual WM tasks are not reliant on executive resources during encoding and maintenance, unlike probe value effects (Hu et al., 2016).

Taken together, these findings might suggest that the manipulations involve distinct underlying cognitive mechanisms. But how might the boosts emerge? Probe value boosts are thought to reflect the more valuable item being retained in the FoA for longer periods of time or more frequently relative to less valuable items (Hitch et al., 2018; Hu et al., 2014; 2016) As probe value boosts appear to rely on executive resources during encoding and maintenance (Hu et al., 2016), effects are likely to result from a process that relies on executive control occurring during one or both of these stages. One possibility is that probe value boosts occur because the more valuable item is consolidated into WM better than the other items. However, this is unlikely to be the sole mechanism, as consolidation is thought to protect items from interference (De Schrijver & Barrouillet, 2017), whereas probe value effects are highly vulnerable to interference (Allen & Ueno, 2018; Hitch et al., 2018; Hu et al., 2014). Alternatively, probe value might bias attentional refreshing (Hitch et al., 2018; Hu et al., 2016), a process that retains information by reactivating decaying memory traces (Barrouillet, Bernardin, Portrat, Vergauwe, & Camos, 2007; Camos et al., 2018; Vergauwe & Cowan, 2015; Vergauwe & Langerock, 2017). The more valuable item might be refreshed more frequently or for longer periods of time (Hitch et al., 2018), thus keeping the representation active. If the more valuable item is active in memory when perceptual interference is experienced (e.g. a post-stimulus suffix), this could distort the representation in memory, leading to erroneous recall that reduces or abolishes the probe value effect. Some evidence for this was provided by Hu et al. (2014), who found that participants often recalled a feature of the to-be-ignored suffix when asked to recall a feature of the more valuable item.

In contrast, probe frequency effects do not appear to rely on executive resources, suggesting that boosts might occur in a cost-free and relatively automatic manner (Hasher & Zacks, 1979; 1984; Zacks & Hasher, 2002). Effects are therefore unlikely to result from processes that rely on executive control, such as enhanced consolidation (Ricker et al., 2018) or attentional refreshing (Camos et al., 2018; Camos et al., 2011). Although speculative, one possibility is the item that is most likely to be probed is automatically tagged as being more important. This might occur because participants are told this item will be tested more frequently, or because they become explicitly or implicitly aware of this throughout the trials. The WM system might then respond to this information by holding it in the FoA automatically. Alternatively, probe frequency effects could result from the more goal-relevant item being encoded with greater strength (McElree, 2001) or prioritised for comparison with the probe at retrieval (Astle et al., 2012). These latter possibilities have been suggested as explanations of visual cueing effects in WM (Souza & Oberauer, 2016), which may also be somewhat automatic in nature (Berryhill et al., 2012; Hommel et al., 2001).

Whilst these outcomes fit with the starting predictions, it could nevertheless be argued that the probe value and probe frequency may involve the same underlying mechanism. Experiment 1 revealed additive effects when probe value and probe frequency were employed together, which are interpreted as indicating the operation of independent underlying mechanisms. However, these outcomes might be expected if the manipulations involve the same mechanism, but neither fully saturate it. Instead of involving distinct mechanisms, the probe value and probe frequency manipulations might therefore activate the same mechanism but in different ways. Experiment 2 would then indicate that activation of this mechanism is somewhat automatic when probe frequency is increased, but under more strategic control when probe value is manipulated (Hu et al., 2016). To delineate between these possibilities, it would be useful for additional research to further explore how the probe value and probe frequency manipulations differ.

As discussed above, the current experiments suggest that probe frequency effects in WM are obtained relatively automatically. However, it seems unlikely that participants would completely ignore the frequency information, rather than using it strategically in order to enhance performance. One possibility is that participants engage in effortful and strategic processing under normal circumstances, but that this does not increase the effects above and beyond the automatic effects observed (Hasher & Zacks, 1979). This therefore does not result in a reduced probe frequency effect under complex concurrent task conditions. A second possibility is that the complex concurrent task used in the current study reduced, rather than completely abolished, executive resources. This may therefore have allowed participants to use any residual executive control to strategically and effortfully direct their attention towards the item most likely to be tested. This could be investigated in further research by examining whether probe frequency effects emerge when participants complete a more cognitively demanding concurrent task, such as counting backwards in 3s (Allen, Baddeley, & Hitch, 2006 (Exp. 4); Jaroslawska, Gathercole, & Holmes, 2018). Regardless of these findings, however, the current experiments provide clear evidence that probe frequency effects are relatively more automatic in nature than probe value effects.

It would also be beneficial for research to investigate whether probe frequency effects exhibit the characteristics of automatic processes outlined by Hasher & Zacks (1979). This includes few developmental trends, whereby effects do not change substantially across the lifespan (Hasher & Zacks, 1979). Emotional states that have been shown to affect attentional capacity, such as depression, stress, or high arousal should also not affect the presence or size of the effect (Hasher & Zacks, 1979). Furthermore, awareness or intentionality to encode/apply the information should have no significant effect on the size of the effect (Hasher & Zacks, 1979). Evidence that probe frequency boosts possess these characteristics would provide further support for the claim that these effects are obtained automatically in WM (Hasher & Zacks, 1979). Moreover, evidence that probe frequency effects emerge even when participants are not aware that a certain SP is more likely to be tested would draw stronger parallels with previous work, such as the Hebb repetition effect (e.g. Hebb, 1961; Oberauer et al., 2015; Page et al., 2006), which is observed even in the absence of explicit awareness (McKelvie, 1987).

However, somewhat refuting this, Cowan et al. (2010) found that probe frequency effects increase during childhood. In this study, 7-8-year-old children, 12-13-year olds children, and adults were asked to remember the location of arrays of four or six shapes for a brief period of time. The shapes presented were either circles or triangles. In some conditions, participants were told that either the circles or the triangles were more likely to be assessed. When four items were presented, all groups could use the probe frequency information to enhance performance. However, when six items were presented, the 7-8-year olds were less able to effectively distribute their attention relative to the other groups. This suggests that there may be a developmental shift in the ability to direct attention in WM based on probe frequency. However, within this study, multiple triangles and circles were presented in each array, meaning that participants needed to direct their attention to multiple items simultaneously. Participants might therefore have utilised alternative strategies, such as filtering out the less relevant shape, rather than focusing attention on the shape most likely to be tested. It would therefore be useful for further research to investigate whether the probe frequency effects observed in the current paradigm are invariant to the effects of age.

Further research could also explore how probe frequency effects change as validity increases or decreases. For instance, one might expect the size of the probe frequency boost to become larger if the difference between the items was increased (e.g. one item was tested 85% of the time, and the others were each tested 5% of the time). Similarly, the size of the effect might decrease if there was a smaller difference between the items (e.g. the item most likely to be tested was assessed 40% of the time, whilst the other items were each tested 20% of the time). Evidence of this would be in line with the visual cueing literature, whereby the size of cueing effects increase alongside cue validity (Gunseli et al., 2015; Shimi et al., 2014). Within the cueing literature, this is taken as evidence that cues encourage individuals to direct their attention strategically (Gunseli et al., 2015; Shimi et al., 2014; Souza & Oberauer, 2016). However, it is possible that the automaticity of effects differs depending on the validity of the instruction. For instance, if the item that is most likely to be tested is assessed on only 40% of the trials, probe frequency effects might not be obtained automatically. Under such circumstances, participants may voluntarily direct their attention to the item most likely to be tested as doing so may improve overall performance. Conversely, if the item that is most likely to be tested is assessed on a larger percentage of trials (e.g. 85%), the effects might be obtained automatically. These suggestions are speculative, however, warranting further research to investigate how the size and nature of the probe frequency effects differ depending on validity.

Finally, it would be useful for research to explore whether other characteristics of probe value and probe frequency differ. Whereas visual cueing appears to protect an item from perceptual interference (Makovski & Jiang, 2007; van Moorselaar et al., 2014; Souza & Oberauer, 2016), probe value effects are vulnerable to interference (Allen & Ueno, 2018; Hitch et al., 2018; Hu et al., 2014). Probe frequency effects might be likened to visual cueing benefits to some degree, as both inform participants which item is most likely to be tested. Furthermore, visual cues (at least using symbols such as arrows) and probe frequency are both considered to yield effects that are somewhat automatic. As such, it might be predicted that probe frequency protect an item from perceptual interference, as it has been suggested visual cues do (Makovski & Jiang, 2007; van Moorselaar et al., 2014; Souza & Oberauer, 2016).

Nevertheless, the current findings have important implications for the relationship between WM and attention by demonstrating that all forms of attentional direction are not functionally equivalent. Terminology such as "cueing" and "prioritisation" therefore not be used interchangeably, as this can encourage researchers to compare outcomes across paradigms. These findings may also have important practical implications, indicating that individuals can direct attention to more important information in WM. Probe value and frequency manipulations can be implemented simultaneously, resulting in larger effects than if one manipulation was used in isolation. This might be particularly useful for everyday tasks that rely on WM, such as learning (Alloway, 2006), planning (Cowan, 2010) and language comprehension (Daneman & Merikle, 1996). However, before these manipulations are implemented in an applied context, further research would first be required to investigate how they impact memory for information in real-life tasks. It is also important to note that the orientation of attention does not appear to increase WM capacity (Hitch et al., 2018; Hu et al., 2014; 2016). As such, directing attention to a particular item might negatively affect memory for other items held within the system.

2.5. Conclusions

In summary, these experiments suggest that the manipulation of probe value and probe frequency encourages may involve distinct cognitive mechanisms. Whilst probe value effects appear to depend on executive resources during encoding and maintenance (Hu et al., 2016), probe frequency effects do not. This suggests that probe frequency effects in WM rely on relatively more automatic processes, whilst probe value effects are driven by strategic and effortful processes (Hu et al., 2016). It is, however, possible that participants did not completely ignore the probe frequency information within the current experiments, but that strategically directing attention based on probe frequency does not yield an additional benefit above and beyond automatic processes. Alternatively, the complex concurrent task might not have completely abolished participants' executive control resources, allowing them to direct any residual resources to strategically remembering the item that was most likely to be tested. It would be beneficial for additional research to further explore this. However, regardless of the outcomes of this further research, the current studies provide clear evidence that probe frequency effects in visual WM are relatively more automatic than probe value effects.

Evidence that probe value and probe frequency effects differ highlights the importance of exploring probe value and probe frequency effects further, as both are likely to provide important insights into the relationship between WM and attention. By examining such effects, one can delineate the characteristics that are manipulation-dependent, and characteristics that describe how attention can be allocated in WM more generally. Finally, these findings indicate that researchers should avoid generalising findings across paradigms, as this may result in erroneous conclusions.

CHAPTER 3

PRIORITISING VALUABLE INFORMATION IN VERBAL WORKING MEMORY

The experiments reported within this chapter have been submitted to the *Journal of Experimental Psychology: Learning, Memory, and Cognition* as Atkinson, A. L., Allen, R. J., Baddeley, A. D., Hitch, G. J., & Waterman, A. H. Can valuable information be prioritised in verbal working memory? The manuscript is currently being revised in line with reviewer comments and will be re-submitted to the journal in the near future.

3.1. Introduction

Chapter 2 demonstrated that probe value and probe frequency effects in WM are likely to result from distinct cognitive mechanisms. Whilst directing attention based on probe value is considered to be under full strategic control (Hu et al., 2016), orienting attention based on probe frequency might be relatively more automatic. As there is also evidence that probe value manipulations differ from cueing (e.g. Hitch et al., 2018; Hu et al., 2016; Makovski & Jiang, 2007; Souza & Oberauer, 2016), the subsequent empirical chapters examined probe value effects in WM more closely. The current chapter examined whether individuals can direct their attention to a more valuable item in auditory-verbal WM, and whether the cognitive mechanisms underpinning such effects are similar to those reported in the visual domain.

Studies investigating how attention can be directed in LTM have often used verbal material, although the stimuli have typically been visually presented (e.g. Castel, Humphreys, et al. 2011; Gruber and Otten, 2010; Middlebrooks, Kerr, & Castel, 2017; Robison & Unsworth, 2017). For instance, stimuli such as words have commonly been used in the value-directed remembering literature (Castel, Humphreys, et al. 2011; Middlebrooks et al., 2017; Robison & Unsworth, 2017). Using such material, these studies have demonstrated that individuals can retain valuable information more effectively in LTM, even when attention is divided between the memory task and another task (Middlebrooks et al., 2017). Several studies reporting that monetary reward enhances LTM have also used verbal material, such as words (Adcock et al., 2006; Gruber and Otten, 2010).

In contrast to this, research examining how individuals orientate attention in WM has primarily used purely visual stimuli, such as coloured shapes (Souza & Oberauer, 2016). However, if the FoA is modality-general, as is claimed in several prominent WM models (Cowan, 2005; Hitch et al., 2018; Hu et al., 2014; Oberauer, 2013), individuals should be able to direct their attention to information regardless of the domain. Indeed, there is some evidence that cueing enhances retention of sounds in WM (Backer & Alain, 2012; Backer, Binns & Alain, 2015; Kumar et al., 2013). Backer et al. (2015) asked participants to listen to an auditory scene, which included four sounds: a non-speech human sound (e.g. laugh), an animal noise (e.g. dog bark), music (e.g. piano), and a man-made object (e.g. drill). Each sound was played from a different speaker, positioned at four different spatial locations. After encoding, a visual cue was presented that was either 100% valid or uninformative (i.e. it did not identify which item would be tested). In one block, the retro-cues were spatial in nature, identifying the location of the sound that would be tested (e.g. 1-5, where 1-4 refer to locations and 5 reflects an uninformative cue). In another block, the retro-cue was semantic, identifying the type of item that would be tested (e.g. M for music). After a short delay, a sound was played at a particular location, and participants had to indicate whether this probe was the same as the one presented during encoding. Relative to the uninformative cues, both the spatial and semantic retro-cues resulted in faster RTs, although only the spatial retro-cues enhanced accuracy. Other studies have, however, found that non-spatial cues boost memory for sounds (Kumar et al., 2013).

Cueing boosts in WM have also been found using visually presented letters or words (Krefeld-Schwalb, 2018; Niklaus, Singmann, & Oberauer, 2019; Shepherdson et al., 2018). Krefeld-Schwalb (2018) found that both 100% and 75% valid retro-cues enhanced performance when letters were used as stimuli. In line with some findings from the visual domain (Hollingworth & Maxcey-Richard, 2013; Makovski & Jiang, 2007; Rerko & Oberauer, 2013, but see Janczyk & Berryhill, 2014), these retro-cue effects were preserved under concurrent task conditions. Replicating these effects, Shepherdson et al. (2018) found that participants responded faster and more accurately when a retro-cue was presented, and the material was letters. However, when the stimuli were words, retro-cues enhanced RTs but not accuracy. Taken together, these studies suggests that cueing enhances memory for auditory and visually-presented verbal information, but the nature of the effects might depend on the type of material and cues used (Backer et al., 2015; Shepherdson et al., 2018).

Although several studies have examined whether individuals can prioritise more valuable information in visual WM (Allen & Ueno, 2018; Berry et al., 2018; Hitch et al., 2018; Hu et al., 2014; 2016), little research has investigated the effects using non-visual material. It is possible that task factors, such as the type of information presented, may be critical in determining how individuals prioritise more valuable information, and the way in which the FoA is used more generally (Hitch et al., 2018). For this reason, Hitch et al. (2018) warn against generalising the conclusions from a single paradigm, and instead argue that divergent approaches should be taken in order to yield converging evidence. Such an approach would allow one to identify general characteristics of probe value effects and the FoA, as well as the factors that vary as a function of task features or demands (Hitch et al., 2018). For instance, whilst probe value effects emerge in visual WM when rehearsal is disrupted (Hitch et al., 2018 Hu et al., 2014; 2016), directing attention to verbal or auditory information might be more reliant on rehearsal. Moreover, there is some evidence to suggest that visuo-spatial WM is more reliant on executive control (Gray et al., 2017) and affected by dual-task interference to a greater extent (Morey, 2018; Morey, Morey, van der Reijden, & Holweg, 2013). As such, it is possible that probe value effects in visual WM might be more reliant on executive resources, relative to the prioritisation of more valuable auditory or verbal material.

To date, only one study has explored whether individuals can prioritise verbal information (i.e. letters) in WM (Sandry et al., 2014). In this pair of experiments, participants were visually presented with a series of three lower case letters on screen. After a brief delay, participants had to recognise one item (now presented in upper case) in a two-alternative forced choice test. In some trials, one letter appeared in red, which indicated that correct recognition of this item would gain them more points, but incorrect responses would result in a greater loss of points. In other trials, all of the letters were black, indicating that all of the items were worth the same reward and penalty. The difference between black and red letters was large, with black letters associated with a gain/loss of three points depending on whether participants responded correctly, whilst red letters were associated with a gain/loss of 25 points. In the first experiment, participants performed near ceiling, thus preventing a sensitive analysis of probe value effects. In the second experiment, participants engaged in articulatory suppression, a non-demanding concurrent task, which involves repetition of irrelevant verbal information (Baddeley et al., 1984; Camos et al., 2009; Richardson & Baddeley, 1975). This task disrupts an individual's ability to rehearse (Baddeley, 1986; Camos et al., 2009), and was implemented to decrease overall performance levels. Under suppression, significant probe value effects were observed, with participants remembering the item associated with greater reward and costs better than the item at the same SP in the control condition, where all of the items were equally valuable. This suggests that individuals can direct attention to more valuable information in verbal WM. These outcomes were also taken as evidence that probe value effects are not reliant on verbal rehearsal. However, as items were presented visually, participants may have held the information in visual WM until retrieval (as opposed to verbal WM), or at least recruited assistance from this domain. This may be likely, as articulatory suppression can prevent verbal recoding of visual information (Baddeley et al., 1984; Baddeley & Hitch, 1994; Schendel & Palmer, 2007). Further research is therefore needed to explore whether individuals are able to direct their attention to more valuable information in verbal WM, and the mechanisms that underpin such effects.

The current experiments therefore examined whether individuals can direct attention to more valuable information in an auditory-verbal WM task. Auditory presentation of items avoids, as much as possible, any reliance on visual coding, whilst also omitting the need for verbal recoding (Baddeley et al., 1984; Baddeley & Hitch, 1994; Schendel & Palmer, 2007). Experiment 3 examined whether participants can strategically prioritise items from one of the serial positions within the sequence. Experiments 4 and 5 then applied concurrent task logic to examine how the ability to recall more and less valuable items might change when either verbal processing (Experiment 4) or both verbal processing and executive control resources (Experiment 5) are disrupted during encoding. These questions were explored using an immediate serial recall task, as this method is commonly used to assess verbal WM in the literature. Digits were selected as the stimulus set, as digit span tasks feature in a range of cognitive batteries (Fliessbach, Hoppe, Schlegel, Elger, & Helmstaedter, 2006; Torralva, Roca, Gleichgerrcht, Bekinschtein, & Manes, 2009; Wechsler, 1997), and are also widely used in experimental research across the discipline (Conway et al., 2005). Use of digits might also reduce reliance on strategies such as association and story-formation, as these approaches may be more easily implemented when more meaningful verbal stimuli, such as words, are used.

A further aim of the experiments was to examine participants' awareness of value effects. Perceived effects have not yet been explored in this literature, but are important to consider from a practical perspective as individuals are unlikely to employ a strategy they consider to be non-beneficial or even harmful to overall performance. Berryhill et al. (2012; Experiment 4) explored whether participants were aware of the benefits of retro-cues. Participants completed 100% valid retro-cue or neutral-cue trials and were then asked to give a confidence judgement after each. Significantly higher confidence ratings were given for the retro-cue trials, indicating that participants were aware that cueing enhances performance. Based on this, it might be predicted that individuals would have a good understanding of the benefits of prioritisation. However, it is not clear whether they would be aware of the negative effects to the low value items.

3.2. Experiment 3

Experiment 3 explored whether individuals can direct attention to more valuable information in an auditory-verbal WM task. Participants listened to a series of digits and attempted to recall them in the correct serial order following a brief delay. Nine items were presented in each trial in order to reduce the likelihood of ceiling effects. This also allowed us to explore the effects of value on tasks that are beyond capacity limits, which is when such an approach is likely to be most beneficial.

Participants completed four value conditions. In three of these conditions, one of the serial positions was more valuable than the other items in the sequence (SP3, SP5 or SP7). There was also a Control condition in which all of the items were worth the same number of points. These conditions were completed in different blocks. This range of positions was selected to sample across the sequence while avoiding early or late positions (which can be subject to ceiling effects and engagement of additional mechanisms in verbal memory). These positions were also selected to investigate whether value effects are observed regardless of which SP participants was more valuable, as has been reported in the visual domain (Hitch et al., 2018). At the end of the experiment, participants then filled out a questionnaire assessing perceived effects of value.

As participants could be told that SP3, SP5, or SP7 was more valuable across the blocks, the names of the value levels was altered in this chapter in order to aid understanding. Instead of the levels being referred to as differential probe value (where one item was more valuable) and equal probe value (where items were worth the same number of points), the levels were referred to as Prioritise-SP3, Prioritise-SP5, Prioritise-SP7, and Control. In the Prioritise-SP conditions, the number specifies the SP participants were asked to prioritise. As in Chapter 2 and previous research, correct recall of this item would earn participants 4 points, whilst correct recall of the other eight items gained them 1 point. In the Control condition, all nine of the items were worth 1 point.

Based on previous findings (Chapter 2; Hitch et al., 2018; Hu et al., 2014; 2016; Middlebrooks et al., 2017; Sandry et al., 2014), it was predicted that value boosts would emerge, with participants exhibiting higher accuracy for the more valuable item, relative to performance at the same serial position in the control condition where all items were of equal value. It was also predicted that costs would emerge to less valuable items. This would provide evidence that individuals can orient attention towards more valuable information in a purely verbal WM task. Taken together with previous studies from the visual domain (e.g. Chapter 2; Allen & Ueno, 2018; Hitch et al., 2018; Hu et al., 2014; 2016), such findings would also be consistent with suggestions that the FoA is modality-general and can hold both visual and verbal representations (Cowan, 2005; Hitch et al., 2018; Hu et al., 2014; Oberauer, 2013).

3.2.1. Method

3.2.1.1. Participants

Twenty-four young adults took part (aged 18-25 years; *M* = 21.14, *SD* = 2.04; 3 males). In this experiment, and all subsequent experiments within this chapter, participants were native English speakers with no known learning difficulties, normal or corrected-to-normal vision, and no colour blindness. Three participants were excluded for reaching length nine on the FDR task. The analysis was therefore run on data from 21 participants (*M. age* = 21.10, *SD* = 1.99; 3 males).

3.2.1.2. Design, materials, and procedure

A 4 x 9 within-subjects design was employed, which manipulated value (Prioritise-SP3, Prioritise-SP5, Prioritise-SP7, Control) and SP (1-9).

Participants first completed a FDR task. This involved listening to a series of digits being read at a steady pace by the experimenter and then repeating them aloud. Each sequence contained series of digits from 0 to 9. Participants first completed two practise trials in which three digits were presented. After this, they completed the experimental trials, which progressed in difficulty from three digits to nine digits. Three trials were presented at each length. Participants continued this task until they responded incorrectly on two out of three trials at a given sequence length. This task was used to exclude participants who reached length nine, as it is likely they would have performed at ceiling in the main task.

Next, participants completed the verbal WM task. In each trial, nine digits were presented (ranging from 1 to 9) auditorily, with the constraint that no digit could appear more than once within the same sequence. Digits were read by a female voice. Participants completed four blocks of 12 trials; one for each condition. The order of the condition blocks was fully counterbalanced. At the start of each block, participants completed two practise trials. Participants were allowed to adjust the volume of the audio to ensure the digits were audible.

Each condition commenced with the provision of written instructions. Participants completed three differential value conditions, in which one of the SPs was more valuable. This was either the third (Prioritise-SP3), the fifth (Prioritise-SP5) and the seventh (Prioritise-SP7) digit. They also completed a Control condition, in which all items were equally valuable. In the differential value conditions, participants were told that correct recall of the more valuable item (i.e. SP3, SP5 or SP7) would earn them four points, whilst correct recall of the other items would earn them one point. This was based on previous research in visual WM (e.g. Chapter 2; Hitch et al., 2018). In the Control condition, participants were told that each item was worth one point.

The experimental paradigm is displayed in Figure 3.1. To begin the trial, participants pressed the space bar. A fixation cross then appeared at the centre of the screen for 1000ms, which was followed by a delay of 1000ms. Nine digits were then played through headphones, separated by an ISI of approximately 1000ms. In the Prioritise-SP3, Prioritise-SP5, and Prioritise-SP7 conditions, a yellow star appeared at the centre of the screen (measuring approximately 4.5cm x 7cm) 500ms before the more valuable digit was played. This remained on screen whilst the digit was played and for approximately 500ms afterwards. In the Control condition, no visual cue was presented. After all the digits had been played, there was a post-stimulus retention interval of approximately 1500ms. This was followed by a white speech bubble with a black outline (measuring approximately 4cm x 5.5cm), displayed at the centre of the screen. This prompted participants to recall the digits in the correct serial order. Participants were told they would only receive points for recalling an item if it was in the correct SP. If they couldn’t remember a digit, participants were told to either guess or say ‘blank’.

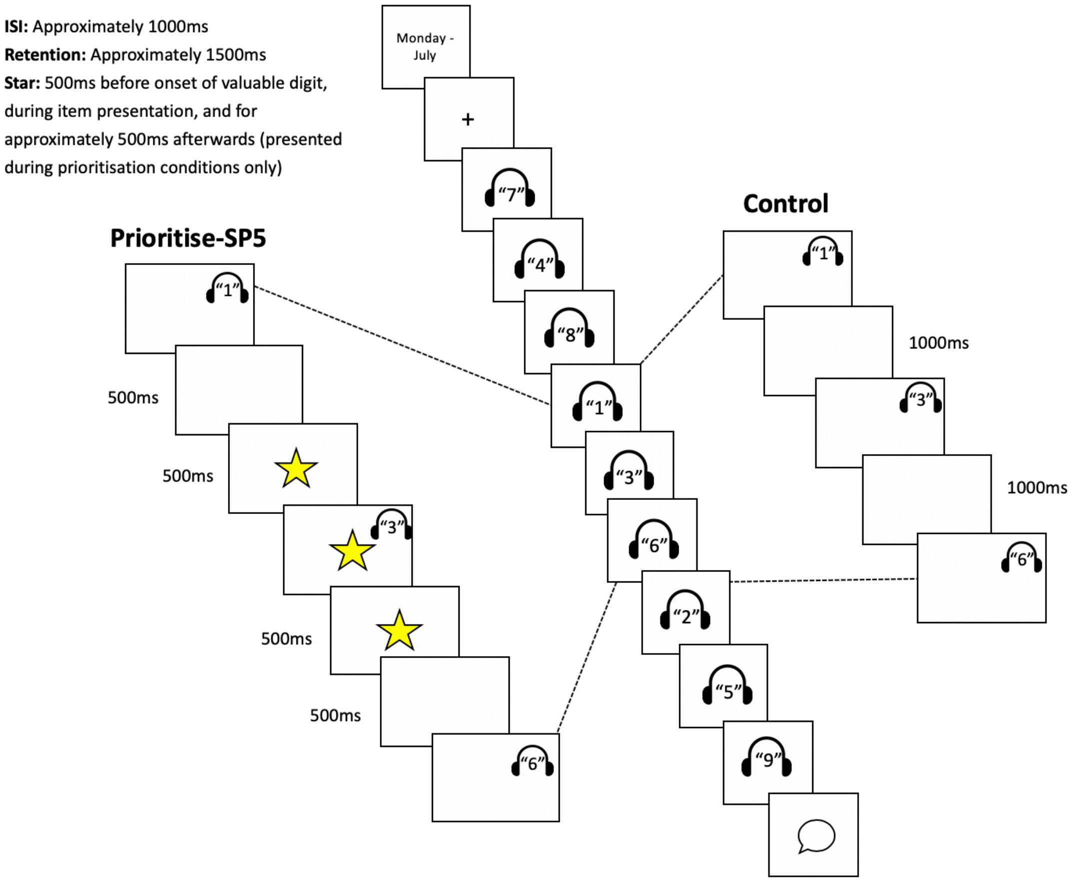


Figure 3.1. Schematic illustration of the experimental paradigm (with Prioritise-SP5 and Control trials as illustrative examples). Participants were first presented with a day of the week and month of the year for 1000ms in Experiment 4 & 5 only. They then heard a series of digits through headphones, with an inter-stimulus interval of approximately 1000ms. In the Prioritise-SP3, Prioritise-SP5 and Prioritise-SP7 conditions, a star was displayed for 500ms before the onset of the digit, during item presentation and for approximately 500ms afterwards. In the Control condition, a blank screen was displayed during this time.

At the end of the experiment, participants completed a questionnaire to assess perceived effects of prioritisation, comprised of 9-point Likert scales (see Appendix

A). This assessed whether participants believed that prioritisation affected memory of the more valuable digit, as well as the other, less valuable items. The questionnaire was separated into four sections, with one for each condition. Low values represented a perceived negative effect of prioritisation, whilst high scores reflected a perceived positive effect.

3.2.1.3. Data analysis

The dependent variable for the main verbal WM task was accuracy at each SP, determined by mean proportion correct (the number of trials responded correctly divided by the number of trials responded correctly plus the number of trials responded incorrectly). Of particular interest was whether prioritisation enhanced memory of the more valuable item relative to the item at the same SP in the control condition. Also of interest was whether recall of other, less valuable items was affected. To investigate this, a composite cost score was calculated for each differential value condition. This was calculated by averaging performance at the eight less valuable SPs (e.g. SPs 1-2 & 4-9 in the Prioritise-SP3 condition). This was then compared to a composite score for the Control condition, calculated by averaging accuracy at the same SPs. This method was employed, as opposed to assessing effects at each SP separately, to avoid the need for a large number of comparisons.

The questionnaire comprised 9-point Likert scales, with a value of 5 indicating “no effect”. To aid in interpretation, 5 was subtracted from each value, such that 0 would represent “no effect”. Negative values therefore reflect perceived costs, whilst positive values signify perceived benefits.

As in Chapter 2, all of the data were analysed using frequentist statistics, as well as BF analysis.

3.2.2. Results

3.2.2.1. Across serial positions

Mean proportion correct (and SE) as a function of value and SP is displayed in Figure 3.2A. Figure 3.2B displays performance in the Prioritise-SP3, Prioritise-SP5, and the Prioritise-SP7 conditions relative to the Control condition.

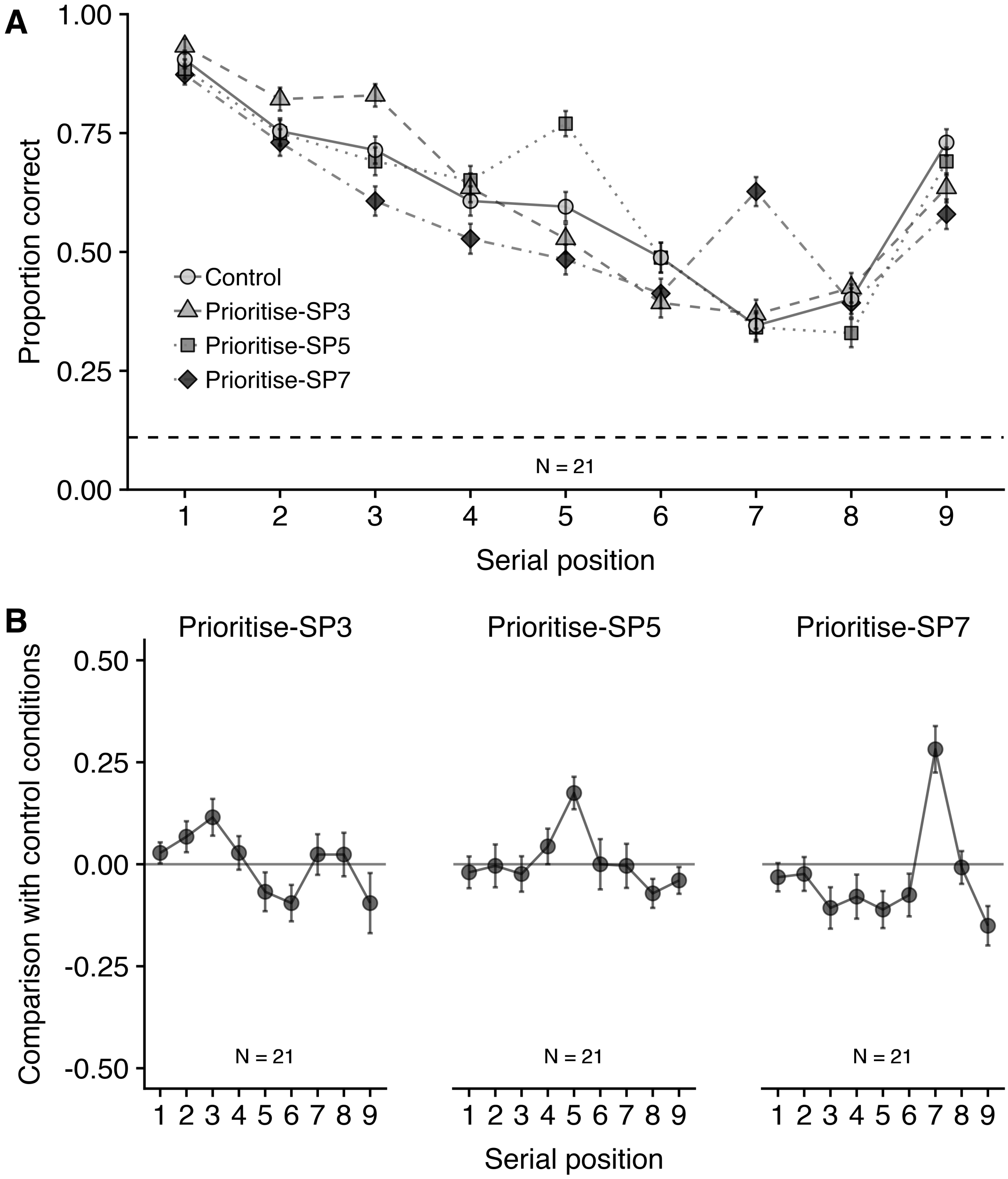


Figure 3.2. Performance in Experiment 3. A) Mean proportion correct (and SE) as a function of value and SP. The dotted grey line (at 0.11) indicates chance guessing rate based on the pool of nine digits. B) Comparison between each value condition and the control condition (centred at 0).

A 4 (Value: Prioritise-SP3, Prioritise-SP5, Prioritise-SP7, Control) x 9 (SP: 1-9) within-subjects ANOVA revealed no significant effect of value (F(3, 60) = 1.11, MSE = .06, p = .354, η_p^2 = .05; BF_{10} = 0.09). There was, however, a significant effect of SP (*GG corrected F*(2.64, 52.87) = 50.20, *MSE* = .15, $p < .001$, $\eta_p^2 = .72$; $BF_{10} >$ 10,000). Of most interest, there was also a significant interaction between value and SP (*GG corrected F*(8.52, 170.45) = 6.98, *MSE* = .06, $p < .001$, $\eta_p^2 = .26$; $BF_{10} >$ 10,000). The BF analysis indicated that the most likely model included effects of SP, as well as an interaction between value and SP ($BF_{10} >$ 10,000 relative to a model containing participant only).

To explore the interaction further, paired sample t-tests, comparing accuracy at the more valuable SP with accuracy at the same SP in the Control condition, were run to establish whether a value-related boost was observed in each case. Figure 3.3 displays the boost individual participants obtained at each of the three targeted SPs. At SP3, recall of the item at SP3 was higher in Prioritise-SP3 (*M* = .83, *SE* = .02) than the Control condition (*M* = .71, *SE* = .04; *t*(20) = 2.55, $p = .019$, *d* = 0.56; BF_{10} = 2.97). Similarly, performance at SP5 was better in the Prioritise-SP5 (*M* = .77, *SE* = .04) relative to the Control condition (*M* = .60, *SE* = .04; *t*(20) = 4.39, $p < .001$, *d* = 0.96; BF_{10} = 109.02). Recall of the item at SP7 was also higher in Prioritise-SP7 (*M* = .63, *SE* = .05) relative to the Control condition (*M* = .35, *SE* = .06; *t*(20) = 4.94, $p < .001$, *d* = 1.08; BF_{10} = 339.59).

3.2.2.2. Size of the value boost

Analysis was also conducted to investigate whether the size of the prioritisation boost at the valuable position differed between conditions. Composite boost scores were calculated by subtracting performance at the targeted position in the Control condition from performance at the targeted position in each differential value condition (e.g. Performance at SP3 in the Prioritise-SP3 condition minus performance at SP3 in the Control condition). A one-way ANOVA revealed that the effects of value differed depending on the SP targeted ($F(2, 40) = 4.31$, $MSE = .04$, $p = .020$, $\eta_p^2 = .18$, $BF_{10} = 3.35$). A series of Bonferroni-Holm corrected paired sample t-tests revealed that the boost was larger when SP7 was targeted ($M = .28$, $SE = .06$) relative to SP3 ($M = .12$, $SE = .05$; $t(20) = 2.87$, $p = .028$, $d = 0.63$; $BF_{10} = 5.28$). There was no significant difference in boosts between SP3 and SP5 ($t(20) = 1.08$, $p = .291$, $d = 0.24$; $BF_{10} = 0.38$, $BF_{01} = 2.63$), or SP5 and SP7 ($t(20) = 1.80$, $p = .174$, $d = 0.39$; $BF_{10} = 0.89$, $BF_{01} = 1.12$).

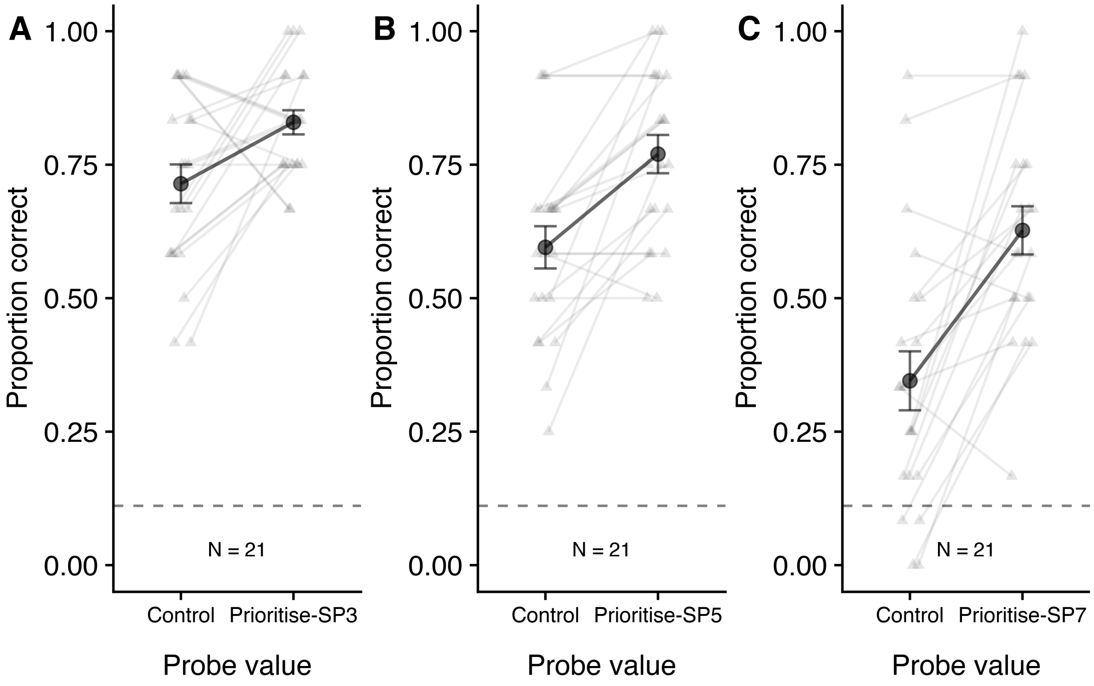


Figure 3.3. The accuracy boost individual participants obtained at each of the targeted positions in Experiment 3. The lighter, grey lines with triangular points display the data for individual participants, whilst the thicker, darker line reflects mean performance. The error bars denote SE, whilst the dashed line at 0.11 reflects chance guessing level (based on the pool of nine digits). A) Performance at SP3 in the Control and the Prioritise-SP3 conditions. B) Performance at SP5 in the Control and the Prioritise-SP5 conditions. C) Performance at SP7 in the Control and the Prioritise-SP7 conditions.

3.2.2.3. Effects to less valuable items

Composite scores were calculated for each differential value condition by averaging performance at the eight less valuable SPs. These scores were compared to composite scores for the Control condition, which were calculated by averaging performance at the same SPs (i.e. omitting the critical SP in each case). The composite scores are shown in Figure 3.4. A series of paired-samples t-tests was conducted to compare composite scores in the differential value and Control conditions, with p-values corrected for multiple comparisons using Bonferroni-Holm. There was no difference between composite scores in the Control condition ($M = 0.60$, $SE = .03$) and Prioritise-SP3 ($M = 0.59$, $SE = .03$; $t(20) = -0.36$, $p = .963$, $d = -0.08$; $BF_{10} = 0.24$, $BF_{01} = 4.17$), or between Control ($M = 0.62$, $SE = .03$) and Prioritise-SP5 ($M = 0.60$, $SE = .03$; $t(20) = -0.72$, $p = 0.963$, $d = -0.16$; $BF_{10} = 0.29$, $BF_{01} = 3.45$). There was a significant difference between composite scores in the Control ($M = 0.65$, $SE = .03$) and Prioritise-SP7 ($M = 0.58$, $SE = .04$) conditions ($t(20) = -2.72$, $p = .040$, $d = -0.59$; $BF_{10} = 4.00$), whereby participants exhibited higher accuracy in the Control condition.

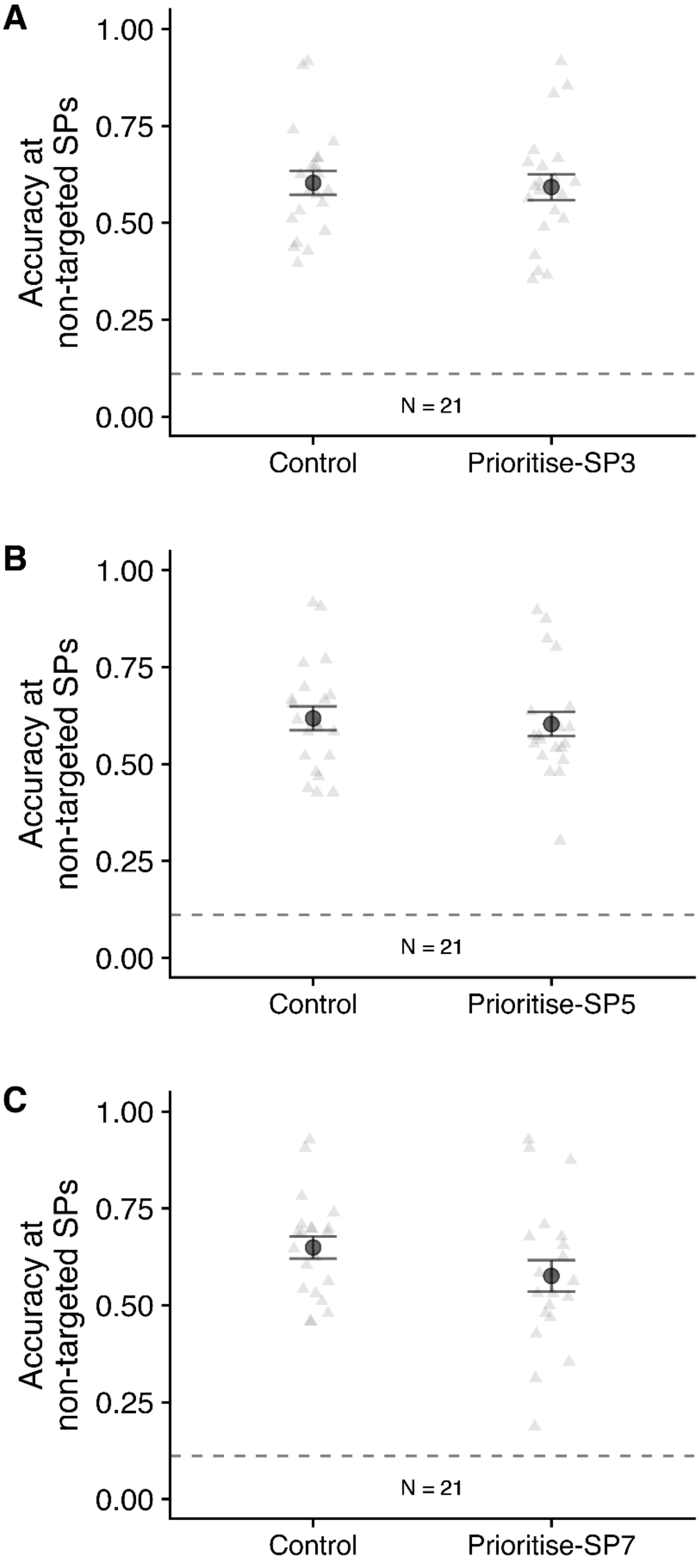


Figure 3.4. The composite cost scores in Experiment 3, calculated by averaging performance at the less valuable SPs in the Prioritise-SP3 (A), Prioritise-SP5 (B), and Prioritise SP7 (C) conditions. The average for the same SPs in the Control

condition is also plotted. The lighter, triangular points reflect the mean cost for individual participants, whilst the darker, circular points reflects the mean across participants. The error bars denote SE.

3.2.2.4. Questionnaire

In the questionnaire, participants were asked whether they believed prioritisation affected memory for the more valuable digit, as well as other less valuable items. These ratings were transformed, such that 0 reflects no effect of value, whilst negative scores indicate perceived costs and positive scores indicate perceived benefits. Mean perceived effects are displayed in Table 3.1. A series of one-sample t-tests were conducted to explore whether the mean ratings differed significantly from 0, with p-values corrected for multiple comparisons using Bonferroni-Holm. Participants perceived benefits to the more valuable item in the Prioritise-SP3 ($t(20) = 3.79$, $p = .006$, $d = 0.83$; $BF_{10} = 32.18$) and Prioritise-SP5 conditions ($t(20) = 3.67$, $p = .006$, $d = 0.80$; $BF_{10} = 25.13$). There was, however, no perceived benefit to the more valuable item in the Prioritise-SP7 condition ($t(20) = 1.56$, $p = .135$, $d = 0.34$; $BF_{10} = 0.64$, $BF_{01} = 1.56$). Participants perceived that prioritising any item (i.e. SP3, SP5 or SP7) resulted in costs to the less valuable items (*Prioritise-SP3:* $t(20) = -3.74$, $p = .006$, $d = -0.82$; $BF_{10} = 28.97$; *Prioritise-SP5:* $t(20) = -3.93$, $p = .005$, $d = -0.86$; $BF_{10} = 42.69$; *Prioritise-SP7:* $t(20) = -3.01$, $p = .014$, $d = -0.66$; $BF_{10} = 6.83$).

Table 3.1.

Mean (and SE) ratings from the questionnaire in Experiment 3, regarding perceived effects of value. Negative values reflect perceived costs, whilst positive values indicate perceived benefits. Values could range from -4 to 4.

	Prioritise-SP3	Prioritise-SP5	Prioritise-SP7
Effect to the more valuable item	1.29 (0.34)	1.24 (0.34)	0.71 (0.46)
Effect to the less valuable items	-1.05 (0.28)	-1.48 (0.38)	-1.48 (0.49)

3.2.3. Discussion

Value boosts were observed, whereby accuracy at the targeted position was higher when that item was associated with more points, relative to a condition in which all items were equally valuable. These boosts were observed in all conditions but were larger towards the end of the sequence (i.e. Prioritise-SP7) than towards the beginning (i.e. Prioritise-SP3). There were significant costs to the less valuable items in the Prioritise-SP7 condition, but not in the Prioritise-SP3 and Prioritise-SP5 conditions. This demonstrates that prioritisation might not always come at a cost to the less valuable items in verbal WM. Analysis of the questionnaire responses revealed perceived boosts to the more valuable digit in the Prioritise-SP3 and Prioritise-SP5 condition, but not in the Prioritise-SP7 condition. In contrast, participants perceived costs to the less valuable items in all three differential value conditions. This indicates that individuals have some awareness of the benefits and associated with prioritisation. However, participants sometimes underestimated the benefits and consistently overestimated the costs, suggesting that they may have a disproportionately negative view of the strategy.

Evidence that value effects emerged demonstrates that individuals can direct attention to more valuable information in an auditory-verbal WM task. This extends previous findings that have used visually presented verbal information in WM (Sandry et al. 2014) and LTM (Middlebrooks et al., 2017) tasks. It also extends findings from the visual domain (Chapter 2; Hu et al., 2014; 2016; Hitch et al., 2018) demonstrating that value effects are not modality-specific and can be observed across WM domains.

What might be the maintenance mechanisms involved in supporting the ability to prioritise certain items other overs? In the visual domain, such effects are typically observed even though participants are asked to concurrently articulate irrelevant verbal information during presentation and maintenance (Chapter 2; Hu et al., 2014; 2016; Hitch et al., 2018), and are therefore unlikely to critically involve verbal rehearsal. Similarly, Sandry et al. (2014) observed that probe value effects for visually presented verbal material remained intact under a similar form of articulatory suppression. However, it is unclear whether the same applies when to-be-remembered information is aurally encountered, with no corresponding visual component to the task. Experiment 4 therefore examined the effects of value to more and less valuable items if participants' ability to verbally rehearse is disrupted by articulatory suppression (Baddeley, 1986; Camos et al., 2009).

3.3. Experiment 4

Experiment 4 investigated whether the ability to prioritise more valuable items within verbal sequences is preserved if participants engaged in articulatory suppression during encoding and maintenance. If boosts are reduced or abolished under such conditions, this would indicate that value effects are critically dependent on verbal rehearsal when information is presented aurally. In contrast, evidence that individuals are able to direct their attention under suppression would be consistent with previous findings using visual modes of presentation (Chapter 2; Hitch et al., 2018; Hu et al., 2014; 2016; Sandry et al., 2014), suggesting that effects are not solely reliant on rehearsal.

Participants completed the same verbal WM task as was employed in Experiment 3. Whilst doing so, participants completed no concurrent task or a simple concurrent task (articulatory suppression). As concurrent task was added as an additional variable, the number of value conditions was reduced to two (Prioritise-SP5 and Control) in order to prevent the experiment from becoming too lengthy, which might introduce fatigue effects. SP5 was selected, as this represents the mid-point in the sequence and generated reliable prioritisation boosts in Experiment 3 with a large effect size (Cohen, 1988). The boost obtained at this position was also intermediate in size relative to when SP3 and SP7 were targeted.

3.3.1. Method

3.3.1.1. Participants

Twenty-four young adults took part (aged 18-24 years; *M* = 21.03, *SD* = 1.48; 5 males). Two participants were excluded for reaching length nine on the FDR task, one participant was removed for not following the instructions correctly and one participant for equipment failure. The analysis was therefore run on data from 20 participants (*M. age* = 21.00, *SD* = 1.49; 5 males).

3.3.1.2. Design, materials, and procedure

A 2 x 2 x 9 within-subjects design was employed, which manipulated value (Prioritise-SP5 vs. Control), concurrent task (no concurrent task vs. simple concurrent task) and SP (1-9). The materials and procedure were identical to Experiment 3, except that in the simple concurrent task conditions, participants were presented with a randomly selected day of the week and a month of the year (e.g. Monday, July) at the centre of the screen for 1000ms before the fixation cross (Calia, Darling, Havelka, & Allen, 2019). Participants were told to repeat this aloud (e.g. *Monday, July, Monday, July*) until the test probe was displayed. This was implemented in order to disrupt verbal rehearsal of the digits.

The questionnaire was also adapted, with a section for each value and concurrent task condition (see Appendix B).

3.3.2. Results

3.3.2.1. Across serial positions

Mean proportion correct (and SE) as a function of value, SP, and concurrent task are displayed in Figure 3.5A. Performance in each Prioritise-SP5 condition relative to the equivalent control condition is displayed in Figure 3.5B as a function of SP and concurrent task.

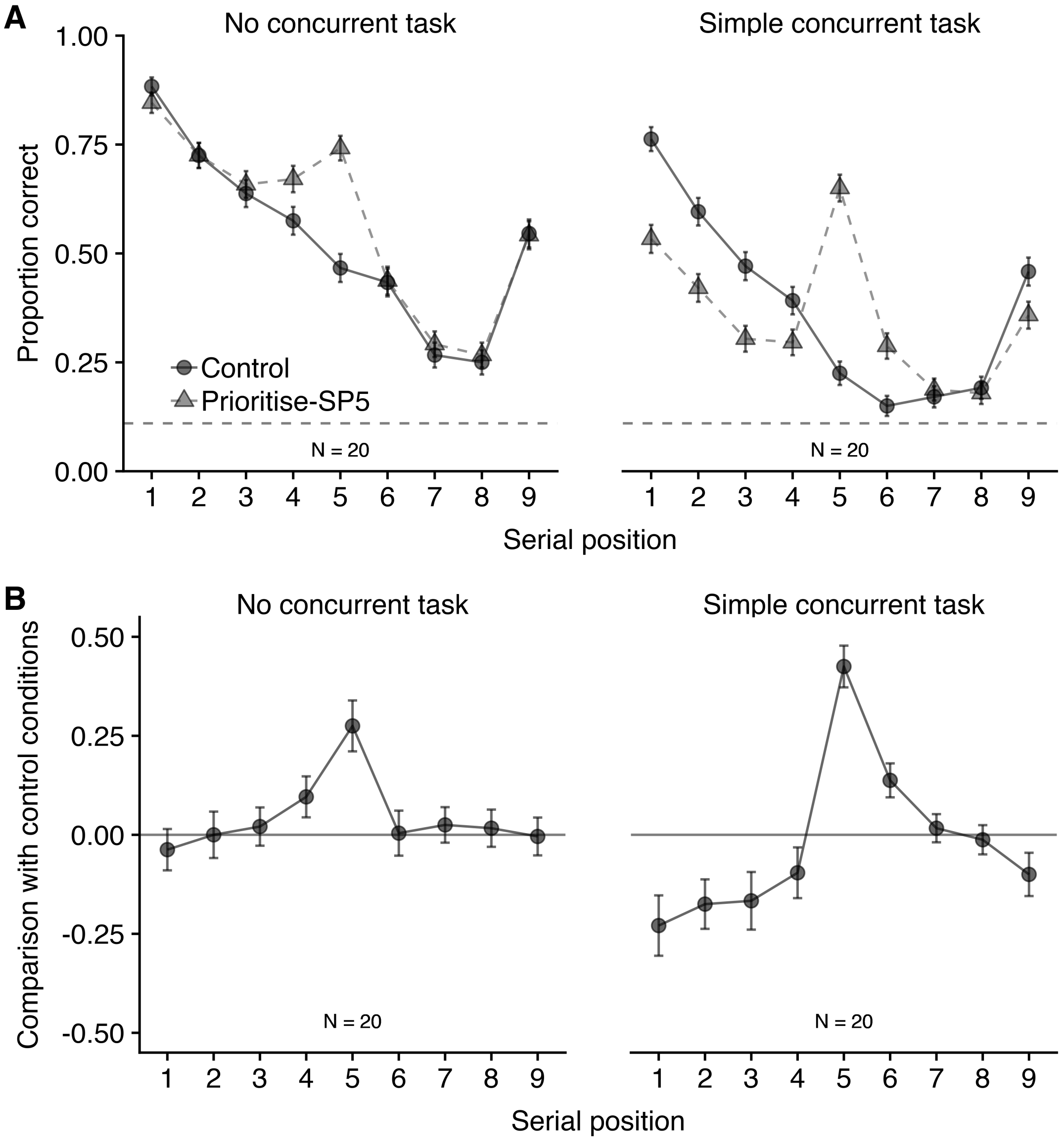


Figure 3.5. Performance in Experiment 4. A) Mean proportion correct (and SE) as a function of value, concurrent task, and SP. The dotted grey line (at 0.11) indicates chance guessing rate, based on the pool of nine digits. B) Comparison between Prioritise-SP5 and Control condition (centred at 0) as a function of concurrent task and SP.

A 2 (Value: Prioritise-SP5 vs Control) x 2 (Concurrent task: no concurrent task vs simple concurrent task) x 9 (SP: 1-9) within-subjects ANOVA revealed no significant effect of value, (*F*(1, 19) = .32, *MSE* = .07, *p* = .577, η_p^2 = .02; BF_{10} = 0.10, BF_{01} = 10.00. There was, however, a significant effect of SP (*GG corrected F*(2.59, 49.27) = 38.02, *MSE* = 0.20, *p* < .001, η_p^2 = .67; BF_{10} > 10,000), characterised by primacy and recency effects, as well as superior performance at SP5. There was also a significant effect of concurrent task (*F*(1, 19) = 50.12, *MSE* = 0.12, *p* < .001, η_p^2 = .73; BF_{10} > 10,000), with participants exhibiting higher accuracy in the no concurrent task condition (*M* = .55, *SE* = .03) relative to the simple concurrent task condition (*M* = .37, *SE* = .03). The interaction between value and concurrent task approached significance (*F*(1, 19) = 3.97, *MSE* = .05, *p* = .061, η_p^2 = .17), with the BF analysis slightly in favour of an effect (BF_{10} = 1.32). There were significant interactions between value and SP (*GG corrected F*(3.10, 58.95) = 12.01, *MSE* = .09, *p* < .001, η_p^2 = .39; BF_{10} > 10,000), and concurrent task and SP (*GG corrected F*(3.85, 73.23) = 3.92, *MSE* = .05, *p* = .007, η_p^2 = .17; BF_{10} = 1.48). A three-way interaction between value, concurrent task, and SP also emerged (*GG corrected F*(3.51, 66.61) = 4.66, *MSE* = .05, *p* = .003, η_p^2 = .20; BF_{10} = 1.88). The BF analysis indicated that the most likely model included main effects of concurrent task and SP, as well as two-way interactions between value and SP, concurrent task and SP, and value and concurrent task, and a three-way interaction between value, concurrent task and SP (BF_{10} > 10,000 relative to a model containing participant only). In order to explore these effects further, additional analyses were conducted at SP5 only (the targeted position) and a composite score of all other SPs (less valuable items).

3.3.2.2. Serial position 5

Participant-level data for SP5 (as well as mean performance levels) is displayed in Figure 3.6, as a function of value and concurrent task. A 2 (Value: Prioritise-SP5 vs Control) x 2 (Concurrent task: no current task vs simple concurrent task) within-subjects ANOVA revealed a significant main effect of value ($F(1, 19) = 46.97$, $MSE = .05$, $p < .001$, $\eta_p^2 = .71$; $BF_{10} > 10{,}000$), whereby participants exhibited greater accuracy at SP5 in the Prioritise-SP5 condition ($M = .35$, $SE = .03$) than the Control condition ($M = .70$, $SE = .05$). There was also a significant main effect of concurrent task ($F(1, 19) = 18.88$, $MSE = .03$, $p < .001$, $\eta_p^2 = .50$; $BF_{10} = 251.37$), with participants exhibiting greater accuracy in the no concurrent task condition ($M = .60$, $SE = .04$) relative to the simple concurrent task condition ($M = .44$, $SE = .03$).A significant interaction between value and concurrent task also emerged ($F(1, 19) = 6.65$, $MSE = 0.02$, $p = .018$, $\eta_p^2 = .26$; $BF_{10} = 1.22$). BF analysis revealed that the most likely model included main effects of value and concurrent task, as well as an interaction between value and concurrent task ($BF_{10} > 10{,}000$ relative to a model containing participant only).

Further analysis (corrected using Bonferroni-Holm) indicated that accuracy was significantly higher in the Prioritise-SP5 relative to the Control conditions when participants completed no concurrent task (*Prioritise-SP5* $M = .74$, $SE = .05$; *Control* $M = .47$, $SE = .05$, $t(19) = 4.28$, $p < .001$, $d = 0.96$; $BF_{10} = 80.30$), and the simple concurrent task (*Prioritise-SP5* $M = .65$, $SE = .05$; *Control* $M = .23$, $SE = .03$, $t(19) = 8.07$, $p < .001$, $d = 1.81$; $BF_{10} > 10{,}000$). The effect size was, however, nearly twice as large in the simple concurrent task condition.

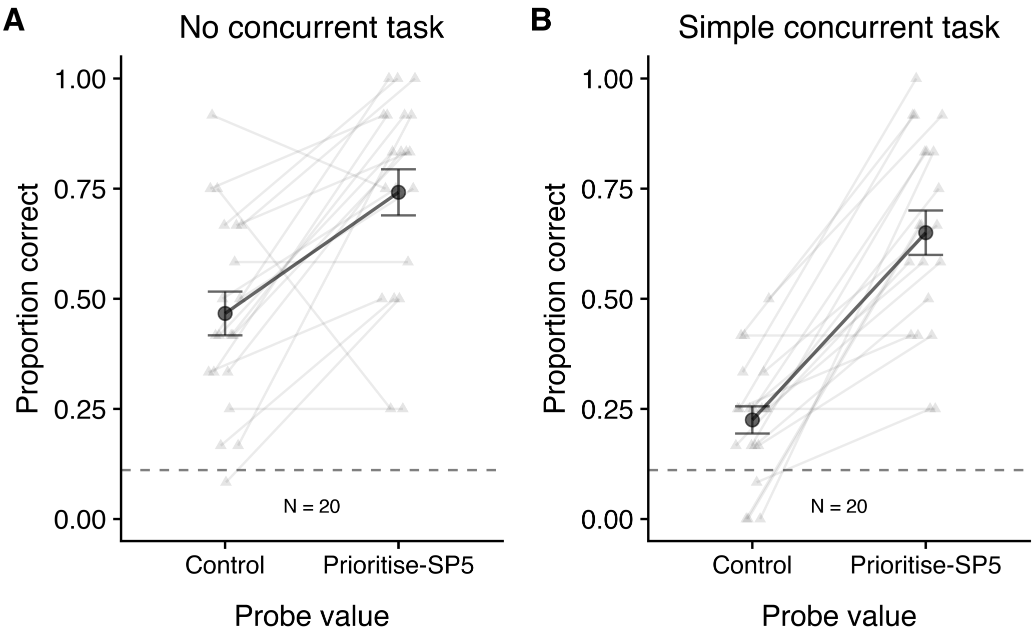


Figure 3.6. Individual participants' data at SP5 in Experiment 4 as a function of value and concurrent task. Part A displays the data for the no concurrent task condition, whilst part B displays the data for the simple concurrent task condition. The lighter grey lines with triangular points reflect individual participants' data, whilst the thicker, darker line with circular points displays the mean across participants. The error bars on this line denote SE. The dashed line at 0.11 reflects chance guessing level (based on the pool of nine digits).

3.3.2.3. Analysis of the less valuable items

As in Experiment 3, composite scores were calculated for each condition by averaging performance across all SPs except SP5. The composite costs are presented in Figure 3.7, as a function of value and concurrent task. A 2 (Value: Prioritise-SP5 vs Control) x 2 (Concurrent task: no concurrent task vs simple concurrent task) within-subjects ANOVA revealed no significant main effect of value ($F(1, 19) = 2.70$, $MSE < .01$, $p = .117$, $\eta_p^2 = .13$, $BF_{10} = 0.57$; $BF_{01} = 1.75$). There was a significant main effect of concurrent task ($F(1, 19) = 49.04$, $MSE = .01$, $p < .001$, η_p^2

= .72; BF_{10} > 10,000) with participants exhibiting reduced accuracy for less valuable items in the simple concurrent task condition (*M* = .36, *SE* = .03) than the no concurrent task condition (*M* = .55, *SE* = .04). There was a significant interaction between value and concurrent task (*F*(1, 19) = 7.45, *MSE* < .01, *p* = .013, η_p^2 = .28; BF_{10} = 1.83). Follow-up paired samples t-tests (corrected using Bonferroni-Holm) revealed no significant costs in the no concurrent task condition (*t*(19) = .60, *p* = .558, *d* = 0.13; BF_{10} = 0.27, BF_{01} = 3.70), but significant costs in the simple concurrent task condition (*t*(19) = -3.01, *p* = .015, *d* = -0.67; BF_{10} = 6.70). The BF analysis indicated that the most likely model includes a main effect of concurrent task and an interaction between value and concurrent task (BF_{10} > *10,000* relative to a model containing participant only).

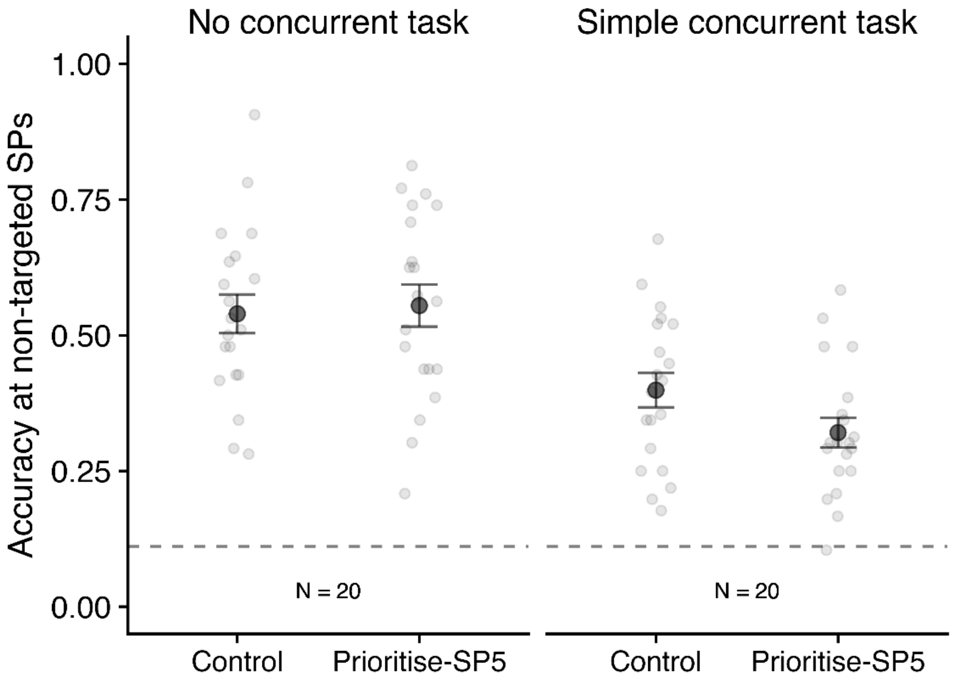


Figure 3.7. The composite cost scores in Experiment 4, as a function of value and concurrent task. These values were calculated by averaging accuracy at the non-targeted SPs (i.e. SPs 1-4 and 6-9). The lighter, triangular points reflect the mean

cost for individual participants, whilst the darker, circular points reflects mean costs across participants. Error bars display SE.

3.3.2.4. Questionnaire

As in Experiment 3, ratings were transformed, such that 0 reflects no perceived effect of value, negative scores reflect perceived costs, and positive scores reflect perceived benefits. Mean perceived effects are displayed in Table 3.2. Four one-sample t-tests were conducted to explore whether ratings differed significantly from zero, with p-values corrected for multiple comparisons using Bonferroni-Holm. Participants believed that memory for the more valuable digit was enhanced by value in the no concurrent task condition ($t(19) = 3.82$, $p = .003$, $d = 0.85$; $BF_{10} = 32.00$), but not in the simple concurrent task condition ($t(19) = 1.31$, $p = .205$, $d = 0.29$; $BF_{10} = 0.49$, $BF_{01} = 2.04$). Participants believed that memory for the less valuable items was negatively affected by prioritisation in the simple concurrent task condition ($t(19) = -4.64$, $p < .001$; $d = -1.04$; $BF_{10} = 166.03$), but not in the no concurrent task ($t(19) = -1.88$, $p = .152$, $d = -0.42$, $BF_{10} = 1.01$).

Table 3.2.

Mean (and SE) ratings from the questionnaire in Experiment 4, regarding perceived effects of value. Negative values reflect perceived costs, whilst positive values indicate perceived benefits. Values could range from -4 to 4.

	No concurrent task	Simple concurrent task
Effect to the more valuable item (SP5)	1.60 (0.42)	0.70 (0.53)
Effect to the less valuable items	-0.60 (0.32)	-1.80 (0.39)

3.3.3. Discussion

Accuracy at SP5 was significantly higher when this position was associated with more points (Prioritise-SP5), relative to a condition in which all items were equally valuable (Control). This replicates Experiment 3 and further demonstrates that individuals can direct their attention to more valuable information in verbal WM. These value boosts were not reduced by concurrent verbal articulation and were in fact larger in size under suppression. This demonstrates that individuals are able to prioritise more valuable verbal information when rehearsal is disrupted.

Replicating outcomes from the same condition in Experiment 3 (Prioritise-SP5), there were no significant value costs in the no concurrent task condition. This therefore provides further evidence that directing attention to a particular item does not always significantly affect others within the sequence. Costs did, however, emerge in the simple concurrent task condition, suggesting that rehearsal may be used to retain less valuable information. Rehearsal processes are thought to be relatively cost-free (Camos & Barrouillet, 2014; Camos et al., 2011), and may thus serve as an ideal way to retain less valuable information.

Analysis of the questionnaire indicated perceived boosts to the more valuable item when no concurrent task was performed, but not in the simple concurrent task condition. Thus, in direct contrast to the observed results, participants had the subjective belief that the concurrent verbal task removed benefits to the more valuable item. This supports the conclusions drawn from Experiment 3, demonstrating that individuals might underestimate the benefits of prioritisation. In line with the behavioural results, participants believed that prioritisation came at a cost to the less valuable items in the simple concurrent task condition, but not in the no concurrent task condition.

Evidence that value boosts emerged in the simple concurrent task condition suggests that effects might involve mechanisms other than rehearsal, such as attentional refreshing, a process whereby decaying memory traces are kept active via executive control (Camos et al., 2018; Sandry et al., 2014). This possibility was investigated in Experiment 5.

3.4. Experiment 5

Experiment 4 demonstrated that boosts to valuable items can be maintained when rehearsal is disrupted. Within the literature a distinction has been drawn between maintenance based on verbal rehearsal, and maintenance based on attentional processes such as refreshing (e.g. Camos et al., 2009). Thus, one possibility is that the more valuable item is retained through attentional refreshing rather than rehearsal (Chapter 2; Hu et al., 2014). If so, sufficient executive attentional resources would be needed in order to prioritise the valuable information. Some support for this hypothesis has been provided in the visual domain, with Hu et al. (2016) demonstrating that probe value boosts were reduced when the complexity of the secondary task was increased (thereby disrupting recruitment of executive control for the main task). This was taken as evidence that effective reward-based prioritisation in visual WM requires executive resources.

Experiment 5 therefore investigated whether individuals can prioritise more valuable information within the current paradigm when executive resources are reduced. Participants were instructed to prioritise a valuable digit whilst completing either a simple or a complex verbal task (adapted from Calia et al., 2019). Based on previous research indicating that executive resources are involved in immediate serial recall, an overall main effect of concurrent task was expected, with poorer performance in the complex condition (Baddeley, Hitch & Allen, 2009; Calia et al., 2019; St. Clair-Thompson & Allen, 2013). Moreover, if this type of prioritisation is critically reliant on executive resources, one might expect an interaction at the more valuable position, whereby the value boost is reduced or abolished under complex concurrent task conditions. It was anticipated that performance would be at floor if nine items were presented in the complex concurrent task condition. As such, the number of items encountered was reduced to seven. Increased value was allocated to SP4, to ensure that the more valuable item continued to be the middle digit in the sequence (in line with Experiment 4).

3.4.1. Method

3.4.1.1. Participants

Twenty-four young adults took part (aged 18-21 years; *M* = 19.34; *SD* = 0.88; 2 males). Participants had not taken part in Experiments 3 or 4.

3.4.1.2. Design, materials, and procedure

A 2 x 2 x 7 within-subjects design was employed, which manipulated value (Prioritise-SP4 vs Control), concurrent task (simple verbal vs. complex verbal) and SP (1-7).

The materials and procedure were similar to Experiment 4, with some exceptions. The no concurrent task condition was dropped, and a complex verbal task condition was introduced. In every trial, participants were presented with a randomly selected day of the week and a month of the year (e.g. Monday, July) at the start of each trial. In the simple verbal task condition, participants had to repeat this pair aloud (e.g. *Monday, July, Monday, July, Monday, July…*) until the test probe was displayed, as in Experiment 4. In the complex task condition, participants had to repeat the day of the week and month of the year presented on screen and then cycle through the days of the week and months of the year in chronological order (e.g. *Monday, July, Tuesday, August, Wednesday, September…*). This complex concurrent task was implemented (as opposed to the concurrent task used in Chapter 2) as it was anticipated that asking participants to count upwards in 2s would cause confusion and interfere with to-be-remembered digits.

As the complex concurrent task was designed to disrupt executive control as well as verbal rehearsal, we predicted that overall accuracy would be lower than with the simple concurrent task. Given that some SPs were approaching chance in the previous experiment, the number of digits presented was reduced from nine to seven. However, despite this reduction in memory load, it was anticipated that participants would not perform near ceiling in this experiment due to the absence of a no concurrent task condition. The initial FDR screening tool was therefore not administered[2]. Finally, the questionnaire was also adapted, with a section for each value and concurrent task condition (see Appendix C).

3.4.2. Results

3.4.2.1. Across serial positions

Mean proportion correct (and SE) as a function of probe value, SP, and concurrent task are displayed in Figure 3.8A. A comparison of each value condition against the equivalent control condition (as baseline) is illustrated in Figure 3.8B.

---

[2] This was confirmed by observing participants' mean performance across SPs, with the highest proportion correct in any condition being 0.86.

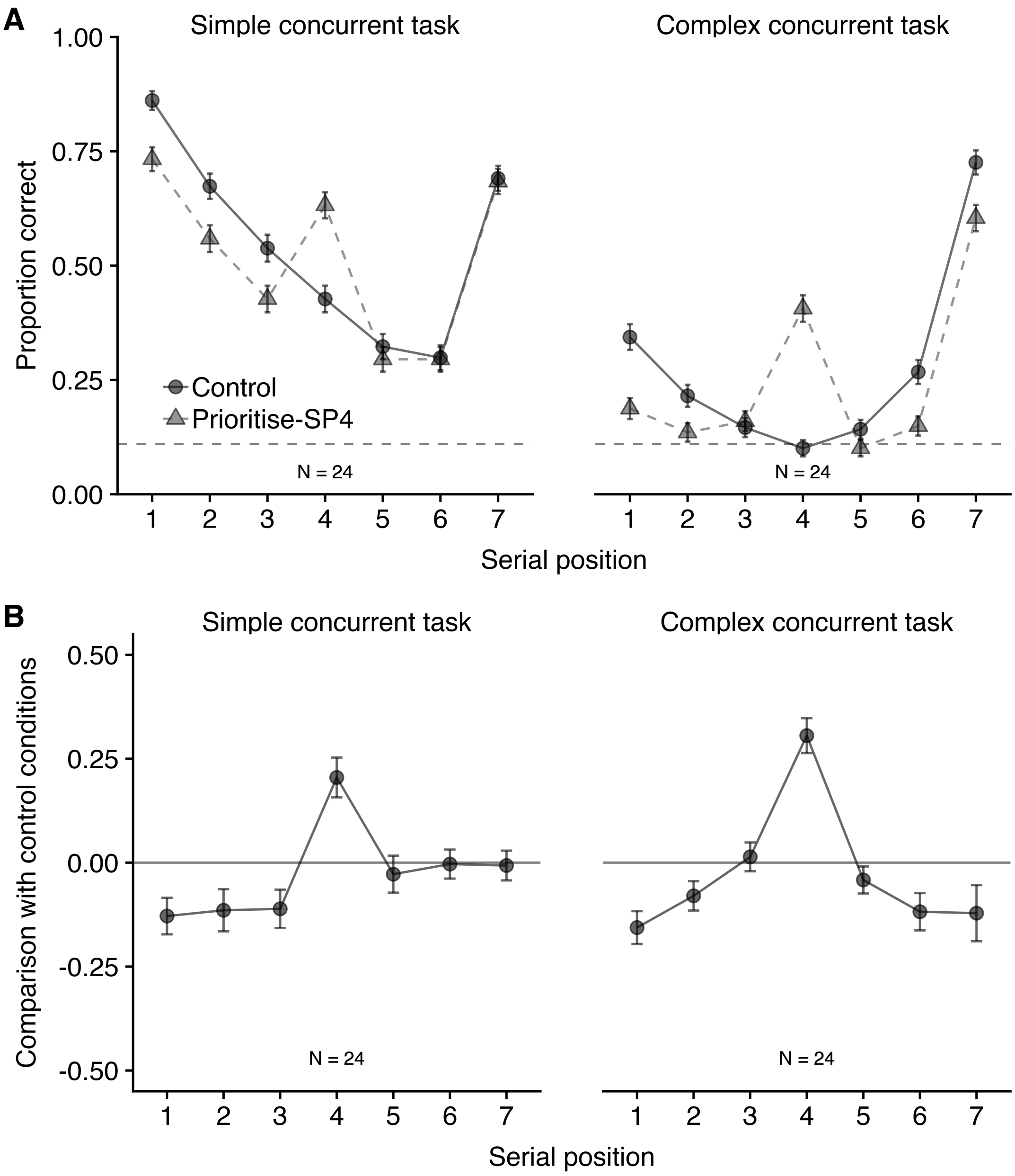


Figure 3.8. Performance in Experiment 5. A) Mean proportion correct (and SE) as a function of value, concurrent task and SP. The dotted grey line (at 0.11) indicates chance guessing rate, based on the pool of nine digits. B) Comparison between Prioritise-SP4 and Control condition (centred at 0) as a function of concurrent task and SP.

A 2 (Value: Prioritise-SP4 vs Control) x 2 (concurrent task: simple concurrent task vs complex concurrent task) x 7 (SP: 1-7) within-subjects ANOVA revealed no significant effect of value ($F(1, 23) = 2.46$, $MSE = .05$, $p = .131$, $\eta_p^2 = .10$; $BF_{10} = 0.83$, $BF_{01} = 1.20$). There was, however, a significant effect of SP (*GG corrected* $F(3.06, 70.43) = 72.80$, $MSE = 0.07$, $p < .001$, $\eta_p^2 = .76$; $BF_{10} > 10{,}000$), characterised by primacy and recency effects, as well as superior performance at SP4. There was also a significant effect of concurrent task ($F(1, 23) = 225.01$, $MSE = 0.05$, $p < .001$, $\eta_p^2 = .91$; $BF_{10} > 10{,}000$) with participants exhibiting higher accuracy in the simple concurrent task condition ($M = .53$, $SE = .02$) relative to the complex concurrent task condition ($M = .26$, $SE = .01$). No significant interaction emerged between value and concurrent task, ($F(1, 23) < .01$, $MSE = .02$, $p = .951$, η_p^2 <. 01; $BF_{10} = 0.12$, $BF_{01} = 8.33$). However, as in the previous experiments, there was a significant interaction between value and SP ($F(6, 138) = 20.17$, $MSE = 0.02$, $p < .001$, $\eta_p^2 = .47$; $BF_{10} > 10{,}000$). There was also a significant interaction between concurrent task and SP (*GG corrected* $F(3.45, 79.26) = 31.38$, $MSE = 0.05$, $p < .001$, $\eta_p^2 = .58$; $BF_{10} = 10{,}000$). A three-way interaction between value, concurrent task, and SP also emerged, although the BF analysis indicated evidence of no effect (*GG corrected* $F(3.92, 90.16) = 2.59$, $MSE = 0.03$, $p = .043$, $\eta_p^2 = .10$; $BF_{10} = 0.50$, $BF_{01} = 2.00$). The BF analysis indicated that the most likely model included main effects of concurrent task and SP, as well as two-way interactions between value and SP, and concurrent task and SP ($BF_{10} > 10{,}000$ relative to a model containing participant only). In order to explore these effects further, additional analyses were conducted at SP4 only (the targeted position) and a composite score of all other SPs (less valuable items).

3.4.2.2. Serial position 4

Mean performance at SP4 and the data for individual participants is displayed in Figure 3.9 as a function of value and concurrent task.

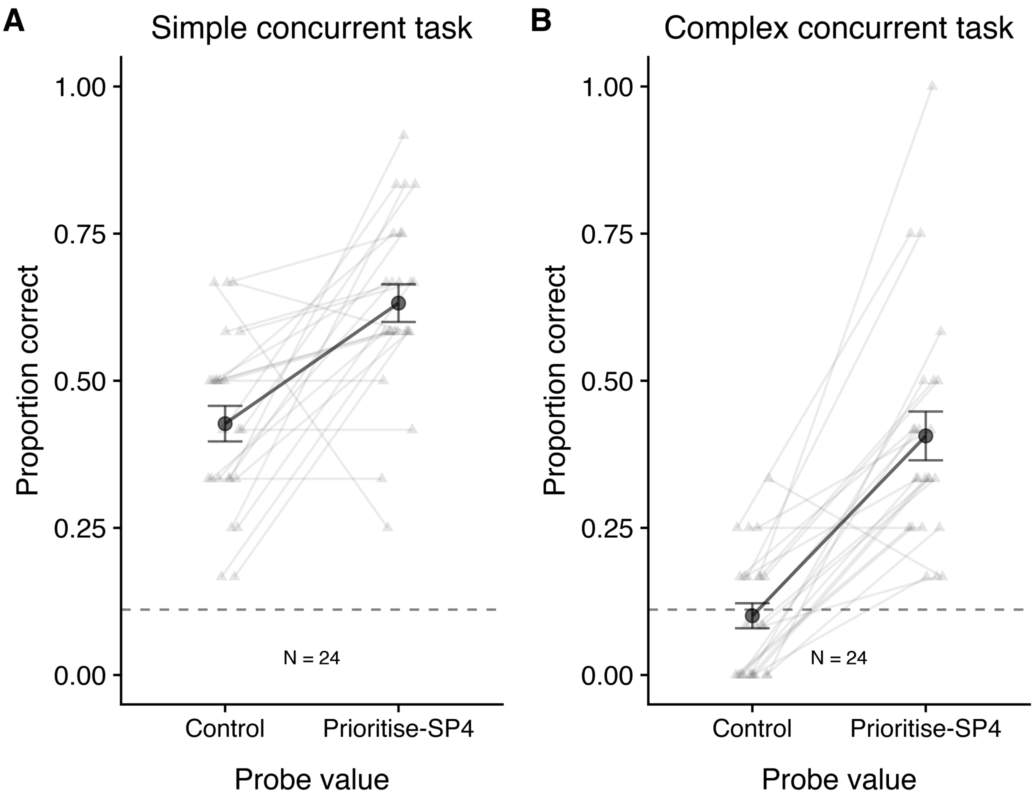


Figure 3.9. The data for individual participants at SP4 in Experiment 5 as a function of value and concurrent task. Part A displays the data for the simple concurrent task condition, whilst part B displays the data for the complex concurrent task conditions. The lighter, grey lines with triangular points display the data for individual participants, whilst the thicker, darker line with circular points reflects the mean across participants. The error bars display SE, and the dotted line at 0.11 reflects chance guessing level, based on the pool of nine possible digits.

A 2 (Value: Prioritise-SP4 vs Control) x 2 (Concurrent task: simple concurrent task vs complex concurrent task) within-subjects ANOVA revealed a significant main effect of value ($F(1, 23) = 70.29$, $MSE = .02$, $p < .001$, $\eta_p^2 = .75$; $BF_{10} > 10{,}000$), whereby participants exhibited greater accuracy at SP4 in the Prioritise-SP4 condition (*M* = .52, *SE* = .03) relative to the Control condition (*M* = .26, *SE* = .02). There was also a significant main effect of concurrent task (*F*(1, 23) = 94.04, *MSE* = .02, $p < .001$, $\eta_p^2 = .80$; $BF_{10} > 10,000$), with participants exhibiting greater accuracy in the simple concurrent task condition (*M* = .53, *SE* = .02) relative to the no concurrent task condition (*M* = .25, *SE* = .03). There was no interaction between value and concurrent task (*F*(1, 23) = 2.32, *MSE* = 0.03, $p = .142$, $\eta_p^2 = .09$), with the BF analysis indicating equivocal evidence ($BF_{10} = 0.92$, $BF_{01} = 1.09$). Refuting the prediction, the value effect was actually numerically larger in the complex task condition (see Figure 3.8 and Figure 3.9). The BF analysis indicated that the most likely model included main effects of value and concurrent task ($BF_{10} > 10,000$ relative to a model containing participant only).

3.4.2.3. Analysis of less valuable items

As in the previous experiments, composite scores were calculated for each condition by averaging performance across all SPs except SP4. These composite scores are displayed in Figure 3.10. A 2 (Value: Prioritise-SP4 vs Control) x 2 (concurrent task: simple concurrent task vs complex concurrent task) within-subjects ANOVA revealed a significant main effect of value (*F*(1, 23) = 16.41, *MSE* < .01, $p < .001$, $\eta_p^2 = .42$; $BF_{10} = 846.41$) with participants exhibiting higher accuracy at these positions in the Control condition (*M* = 0.44, *SE* = 0.02) relative to the Prioritise-SP4 condition (*M* = 0.36, *SE* = 0.02). There was a significant main effect of concurrent task (*F*(1, 23) = 235.33, *MSE* < .01, $p < .001$, $\eta_p^2 = .91$; $BF_{10} > 10,000$), with participants exhibiting reduced accuracy for the less valuable items in the complex condition (*M* = 0.27, *SE* = 0.01) relative to the simple condition (*M* = .53, *SE* =

0.02). The interaction between value and concurrent task was not significant ($F(1, 23) = 0.44$, $MSE < 0.01$, $p = .512$, $\eta_p^2 = .02$; $BF_{10} = 0.33$, $BF_{01} = 3.03$). The BF analysis indicated that the most likely model included main effects of value and concurrent task ($BF_{10} > 10,000$ relative to a model containing participant only).

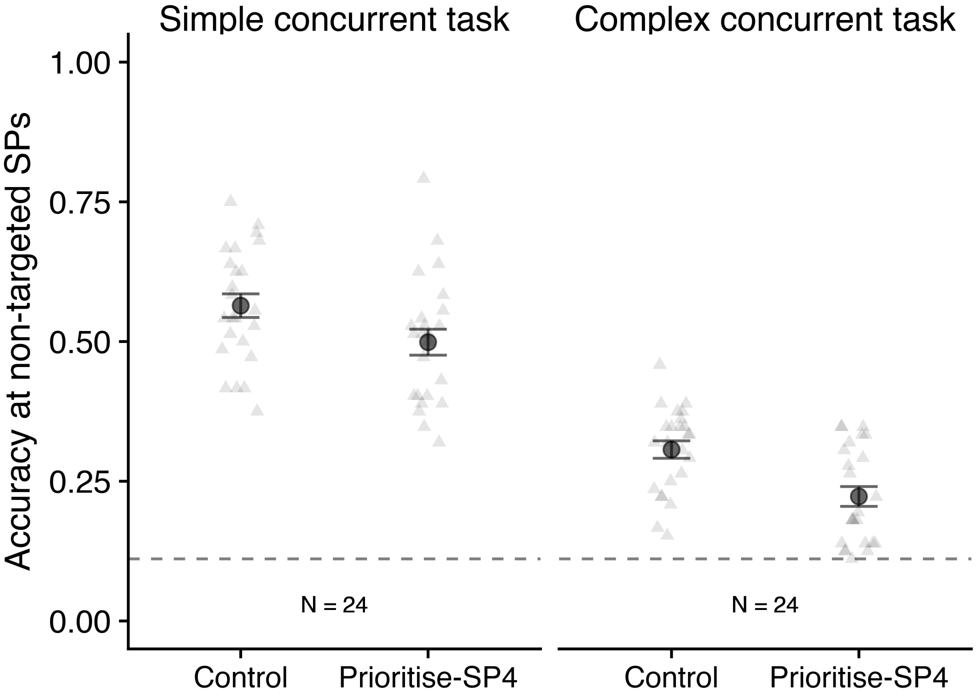


Figure 3.10. Mean composite costs in Experiment 5, as a function of value and SP. The lighter, triangular points display the mean composite cost scores for individual participants. The darker, circular points reflect the mean across participants, whilst the error bars denoting SE. The dashed line at 0.11 shows the chance guessing rate, based on a pool of nine digits.

Observation of Figure 3.8A suggests that performance at several of the less valuable items was near floor in the complex concurrent task condition. This was particularly apparent in the Prioritise-SP4 condition, which might suggest that participants abandoned the less valuable items in order to retain the more valuable digit. In order to investigate this further, a series of one-sample t-test were conducted to explore whether performance in the complex concurrent task conditions significantly differed from chance guessing rate (0.11). P-values were corrected using Bonferroni-Holm. In the Prioritise-SP4 condition, performance at SP4 ($t(23) = 7.15, p < .001, d = 1.46$; $BF_{10} > 10{,}000$) and SP7 ($t(23) = 9.51, p < .001, d = 1.94$; $BF_{10} > 10{,}000$) significantly differed from chance, whilst all other SPs did not ($t \geq -0.42$ and ≤ 2.12, $p \geq .226$; $BF_{10} \leq 1.41$). This suggests that performance was at floor for all positions except from the more valuable item and the final item. In the Control condition, accuracy at SP1 ($t(23) = 7.25, p < .001, d = 1.48$; $BF_{10} > 10{,}000$), SP2 ($t(23) = 3.66, p = .005, d = 0.75$; $BF_{10} = 27.71$), SP6 ($t(23) = 4.19, p = .002, d = 0.86$; $BF_{10} = 88.93$) and SP7 ($t(23) = 12.56, p < .001, d = 2.56$; $BF_{10} > 10{,}000$) was significantly greater than chance, whilst performance at SP3, SP4, and SP5 did not differ ($t \geq -0.44$ and $\leq 1.51, p \geq .431$; $BF_{10} \leq 0.58$; $BF_{10} \geq 1.72$).

3.4.2.4. Omissions

As performance at all of the less valuable SPs (except the final position) was at floor in the Prioritise-SP4 condition, this might suggest that participants abandoned these items in order to retain the more valuable digit. Within the task, participants were told that they could guess or say ‘blank’ if they were unsure of a digit. Further analysis was therefore conducted to investigate whether the number of omissions (i.e. the number of times participants said “blank” or did not give a response) differed as a function of value, concurrent task and SP within this experiment. The mean (and SE) proportion of responses that were classed as omissions is displayed in Figure 3.11. A 2 (Value: Prioritise-SP4 vs Control) x 2 (Concurrent task: simple concurrent task vs complex concurrent task) x 7 (SP: 1-7) within-subjects ANOVA revealed a main effect of value ($F(1, 23) = 16.79$, $MSE = .03$, $p < .001$, $\eta_p^2 = .42$; $BF_{10} > 10{,}000$), whereby the number of omissions was larger in the Prioritise-SP4 condition ($M = .23$, $SE = .04$) than the Control condition ($M = .18$, SE = .04). There was also a significant effect of concurrent task ($F(1, 23) = 31.29$, $MSE = .18$, $p < .001$, $\eta_p^2 = .58$; $BF_{10} > 10{,}000$), whereby participants omitted more responses in the complex condition ($M = .30$, $SE = .05$) than the simple condition ($M = .11$, $SE = .02$). A significant main effect of SP also emerged (*GG corrected* $F(2.90, 66.69) = 13.98$, $MSE = .09$, $p < .001$, $\eta_p^2 = .38$; $BF_{10} > 10{,}000$). There were significant two-way interactions between value and SP (*GG corrected* $F(3.55, 81.57) = 12.79$, $MSE = .04$, $p < .001$, $\eta_p^2 = .36$; $BF_{10} > 10{,}000$), and concurrent task and SP (*GG corrected* $F(2.31, 53.14) = 9.19$, $MSE = .08$, $p < .001$, $\eta_p^2 = .29$; $BF_{10} > 10{,}000$). There was also a significant three-way interaction between value, concurrent task, and SP (*GG corrected* $F(4.11, 94.44) = 7.39$, $MSE = .02$, $p < .001$, $\eta_p^2 = .24$; $BF_{10} > 10{,}000$). The interaction between value and concurrent task approached significance ($F(1, 23) = 4.06$, $MSE = .03$, $p = .056$, $\eta_p^2 = .15$), but was strongly supported by the BF analysis ($BF_{10} = 14.94$). The BF analysis revealed that the most likely model included all main effects and interactions ($BF_{10} > 10{,}000$ relative to a model containing participant only). In order to explore these effects further, ANOVAs were conducted at SP4 only and a composite score of all other SPs (i.e. less valuable items).

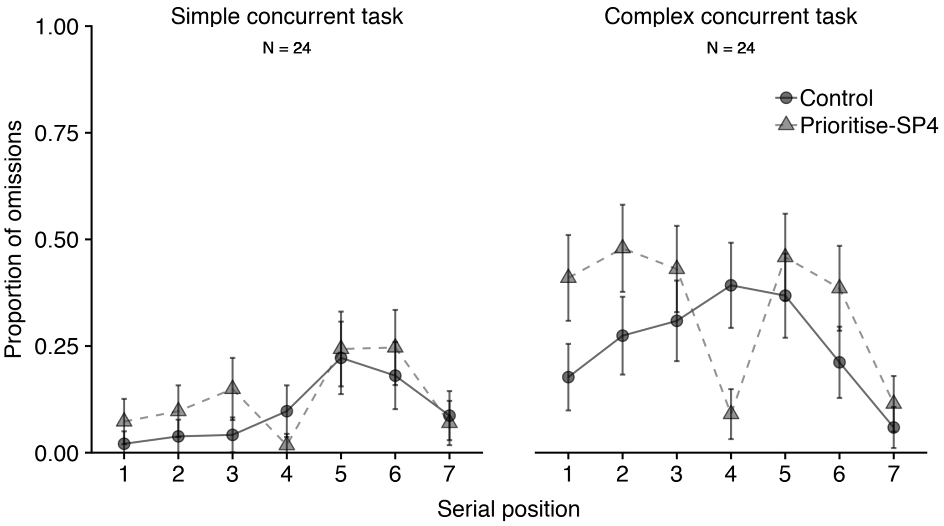


Figure 3.11. Mean proportion of responses classed as omissions (and SE) in Experiment 5, as a function of value, concurrent task, and SP.

3.4.2.4.1. Serial position 4

A 2 (Value: Prioritise-SP4 vs Control) x 2 (Concurrent task: simple concurrent task vs complex concurrent task) ANOVA revealed a main effect of value, whereby the mean proportion of omissions was higher in the Control condition (*M* = .25, *SE* = .05) than the Prioritise-SP4 condition (*M* = .05, *SE* = .02; *F*(1, 23) = 21.64, *MSE* = .04, *p* < .001, η_p^2 = .49; BF_{10} > 10,000). There was also a main effect of concurrent task, whereby participants omitted more responses in the complex condition (*M* = .24, *SE* = .05) relative to the simple condition (*M* = .06, *SE* = .02; *F*(1, 23) = 19.23, *MSE* = .04, *p* < .001, η_p^2 = .46; BF_{10} > 10,000). A significant interaction emerged between value and concurrent task (*F*(1, 23) = 14.57, *MSE* = .02, *p* = .001, η_p^2 = .39; BF_{10} > 10,000). The BF analysis indicated that the most likely model included main effects of value and concurrent task, and the interaction (BF_{10} > 10,000 relative to a model with participant only). Bonferroni-Holm corrected paired sample t-tests revealed that participants omitted more responses in the Control condition than the Prioritise-SP4 condition under both simple (*Prioritise-SP4 M* = .02, *SE* = .01; *Control M* = .10, *SE* = .03; *t*(23) = -3.22*, p* = .004; d = -0.66; BF_{10} = 10.98) and complex conditions (*Prioritise-SP4 M* = .09, *SE* = .04; *Control M* = .39, *SE* = .08; *t*(23) = -4.53*, p* < .001; *d* = -0.92; BF_{10} = 187.60).

3.4.2.4.2. Less valuable items

As with accuracy, a composite score was calculated by averaging the number of omissions at the less valuable items. A 2 (Value: Prioritise-SP4 vs Control) x 2 (Concurrent task: simple concurrent task vs complex concurrent task) ANOVA revealed a significant main effect of value (*F*(1, 23) = 35.68, *MSE* < .01, *p* < .001, η_p^2 = .61; BF_{10} = 342.00), whereby participants omitted more responses in the Prioritise-SP4 condition (*M* = .26, *SE* = .04) than the Control condition (*M* = .17, *SE* = .03). There was also a significant effect of concurrent task (*F*(1, 23) = 32.68, *MSE* = .03, *p* < .001, η_p^2 = .59; BF_{10} > 10,000), whereby participants omitted more responses in the complex condition (*M* = .31, *SE* = .05) than the simple condition (*M* = .12, *SE* = .03). A significant interaction emerged between value and concurrent task (*F*(1, 23) = 10.65, *MSE* < .01, *p* = .003, η_p^2 = .32; BF_{10} = 1.86). The BF analysis indicated that the most likely model included main effects of value and concurrent task and the interaction (BF_{10} > 10,000 relative to a model with participant only).

To explore the interaction, paired-sample t-tests were conducted and corrected using Bonferroni-Holm. There was a significant effect of value in both the simple concurrent task condition (*Prioritise-SP4 M* = .15, *SE* = .03; *Control M* = .10, *SE* = .02; *t*(23) = 3.31, *p* = .003, *d* = 0.68; BF_{10} = 13.33) and the complex concurrent task condition (*Prioritise-SP4* $M = .38$, $SE = .06$; *Control* $M = .23$, SE = .05; $t(23) = 5.26$, $p < .001$, $d = 1.07$; $BF_{10} = 968.93$), although this effect was larger in the latter condition.

3.4.2.5. Questionnaire

As in the previous experiments, ratings were adjusted so that 0 reflects no perceived effect, negative scores reflect perceived costs, and positive scores reflect perceived boosts. Mean perceived effects are displayed in Table 3.3. One sample t-tests, corrected using Bonferroni-Holm, were conducted to investigate whether participants perceived benefits or costs of prioritisation in the simple and complex concurrent task conditions. Participants perceived significant benefits of prioritisation to the more valuable digit in the simple ($t(23) = 4.84$, $p < .001$, $d = 0.99$; $BF_{10} = 379.60$) and complex concurrent task conditions ($t(23) = 4.78$, $p < .001$, $d = 0.98$; $BF_{10} = 330.56$). They also perceived significant costs to the less valuable items in both conditions (*Simple* $t(23) = -2.36$, $p = .027$, $d = -0.48$; $BF_{10} = 2.13$; *Complex* $t(23) = -3.55$, $p = .003$, $d = -0.72$; $BF_{10} = 22.02$).

Table 3.3.

Mean (and SE) perceived effects of value obtained from the questionnaire. As in the previous experiments, negative values demonstrate that participants perceived costs, whilst positive values reflect perceived benefits. Values could range from -4 to 4.

	Simple concurrent task	Complex concurrent task
Effect to the more valuable item (SP4)	1.67 (0.34)	1.38 (0.29)
Effect to the less valuable items	-0.79 (0.34)	-1.63 (0.46)

3.4.3. Discussion

In this experiment, participants attempted to prioritise a more valuable item whilst engaging in a concurrent task that disrupted verbal rehearsal (simple concurrent task), or verbal rehearsal and executive control (complex concurrent task). Overall accuracy was lower when participants performed the complex concurrent task (Calia et al., 2019), indicating that executive control resources are involved in immediate serial recall of digit sequences (St. Clair-Thompson & Allen, 2013). Moreover, replicating the previous experiments, value effects were observed, whereby accuracy at the targeted position (SP4) was significantly higher when this item was more valuable. This effect was numerically larger in the complex concurrent task conditions, demonstrating that individuals are able to prioritise information in the current paradigm when executive control resources are diminished.

Costs to the less valuable items emerged, although this effect did not appear to vary as a function of concurrent task. It is, however, worth noting that mean accuracy at the less valuable positions (except the final item) was at floor in the complex concurrent task condition (see Figure 3.8A). This suggests that disruption to both verbal rehearsal and executive control renders individuals only able to recall the more valuable item and the most recently encountered in the sequence.

Evidence that value boosts emerged under high cognitive load might suggest that these effects do not appear crucially to rely on executive resources in a serial recall, verbal WM task. This would be in line with studies using verbal material in the directed-remembering literature (Middlebrooks et al. 2017) and the cueing literature (Krefeld-Schwalb, 2018), but would contrast with research employing a similar type of prioritisation in visual WM (Hu et al., 2016). It is, however, important to review this outcome within the broader set of findings. Although prioritisation costs did not vary as a function of concurrent task, performance at all of the less valuable positions except the final item was at chance level in the complex concurrent task condition. This suggests that individuals may have abandoned the less valuable digits and only attempted to retain the more valuable item. In order to explore this possibility further, the mean rate of omissions was analysed. At SP4, the omission rate was higher in the Control condition, thus providing further evidence that participants were able to prioritise the more valuable item for recall. However, at the less valuable SPs, participants were significantly more likely to omit responses in the Prioritise-SP4 condition. This might therefore provide evidence that participants did indeed abandon the less valuable items in order to retain the more valuable digit.

As noted, a clear recency advantage emerged. This can be seen even in the condition where participants are encouraged to direct their attention to another item and perform a cognitively demanding concurrent task. Such findings are consistent with claims that the final item is retained automatically in verbal WM (Baddeley &

Hitch, 1993) and mirrors effects reported in the visual domain (Allen, Baddeley, & Hitch, 2014; Hu et al., 2016).

Analysis of the questionnaire revealed that participants were aware of value boosts at the targeted position (SP4) and costs to the less valuable items in both concurrent task conditions. This is in line with the behavioural pattern of benefits and costs. However, evidence that perceived boosts emerged in the simple concurrent task condition is not consistent with Experiment 4, where participants did not perceive benefits of prioritisation with the condition involving a simple concurrent task. However, it is worth noting that the evidence of no perceived effect in Experiment 4 was not particularly strong (BF_{10} = 0.49, BF_{01} = 2.04), and therefore the differences between the two experiments may be due to variability in subjective opinions.

3.5. Predicting the size of value boosts and costs

As shown in Figure 3.3, Figure 3.6, and Figure 3.9, there were larger individual differences in the size of the value effects, with some participants obtaining a large boost, whilst others experienced no effect or even a cost at the targeted SP. In Experiments 3 and 4, participants completed a FDR task before the main experimental task. This was implemented in order to exclude participants who would likely perform at ceiling in the main task. However, this also allows us to conduct exploratory analysis to investigate whether STM capacity is associated with the size of value boosts and costs. Participants in Experiment 3 and Experiment 4 both completed a condition in which SP5 was more valuable and no concurrent task was performed. Analysis was therefore conducted to examine whether performance on the FDR task was associated with the size of the performance boost at SP5 and the cost at the other SPs. This condition was selected in order to maximise the possible sample size. The size of the boost was calculated by subtracting performance in the Control condition at SP5 from performance at the same position in the Prioritise-SP5 condition. Similarly, the size of the cost was calculated by subtracting the average performance at all SPs (except SP5) in the Control condition from the average performance at the same SPs in the Prioritise-SP5 condition.

One possibility is that individuals with a lower STM capacity might be worse at directing their attention as they struggle to accurately retain a representation of the goal whilst also remembering the memoranda. Conversely, it may be that individuals with poorer STM abilities recognise that they are performing sub-optimally and thus apply the prioritisation strategy more readily than individuals with a higher capacity.

The correlation between FDR and boost at SP5 is displayed in Figure 3.12A, whilst the correlation between FDR and the composite cost is displayed in Figure 3.12B. Two frequentist Pearson's product-moment correlations (corrected using Bonferroni-Holm) were conducted to investigate statistical significance. The data was also analysed using Bayesian correlations, using the BayesFactor package (Morey et al., 2018) in R. The default priors were used. The correlation between FDR and the value boost was not significant ($r(39) = -.22$, $p = .345$; $BF_{10} = 0.79$, $BF_{01} = 1.27$). There was also no significant correlation between FDR and the composite cost to less valuable items ($r(39) = .09$, $p = .582$; $BF_{10} = 0.40$, $BF_{01} = 2.50$).

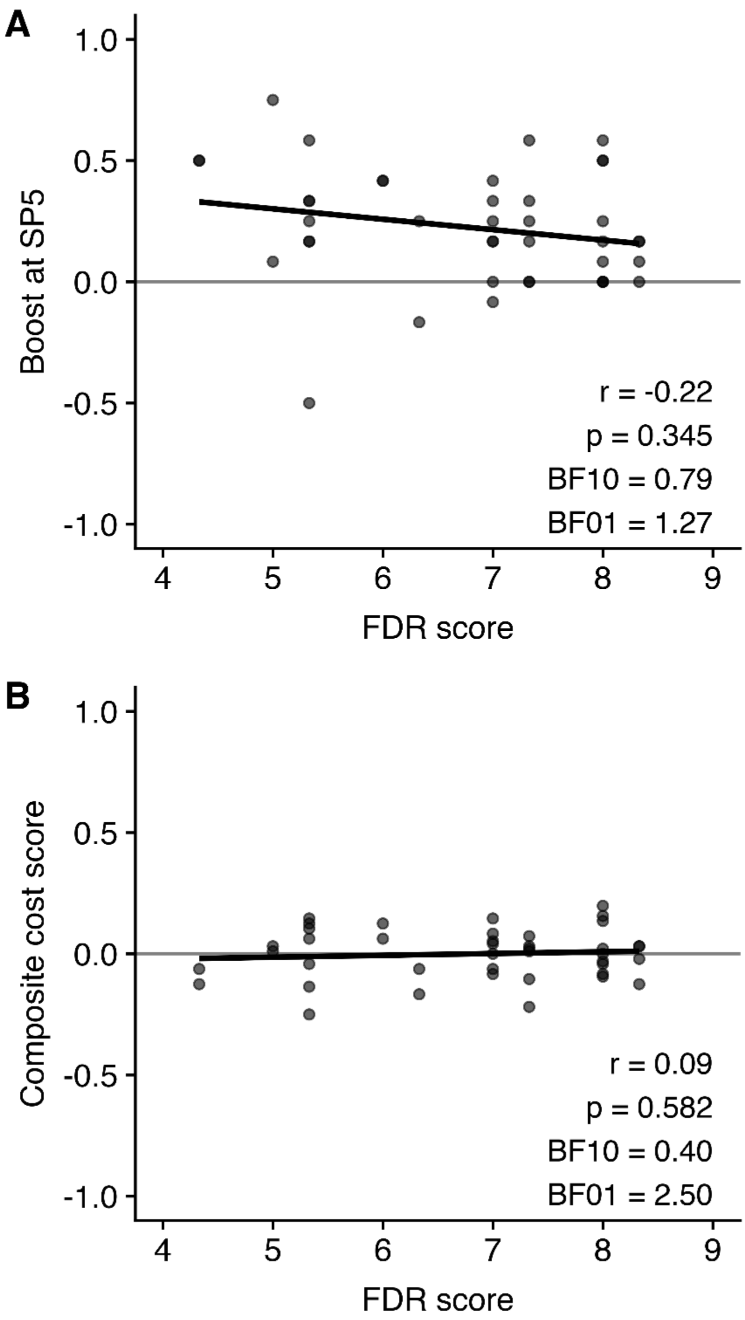


Figure 3.12. The correlations between FDR score and the boost at SP5 (A) and the composite cost to the less valuable items (B). The circular points reflect individual data points. When the circle is darker than the rest, there is more than one datapoint at that position. The thick, dark line reflects the line of best fit, whilst the lighter grey line at 0 reflects no relationship.

3.6. General discussion

Previous research has demonstrated that individuals can prioritise more valuable information presented visually (Chapter 2; Allen & Ueno, 2018; Hu et al., 2014; 2016; Hitch et al., 2018; Sandry et al., 2014). However, this series of experiments was the first to investigate whether attention can be directed to particular items in an auditory-verbal WM task. Experiment 3 revealed that individuals are able to prioritise more valuable information, regardless of whether this item appears near the beginning, middle, or end of a verbal sequence. This effect was larger when participants prioritised an item towards the end of the sequence (i.e. SP7) relative to an item near the beginning of the sequence (i.e. SP3). Value boosts were accompanied by costs only when participants directed their attention to SP7. Two follow-up experiments then explored how these effects are moderated by concurrent performance of a simple (Experiment 4) or complex (Experiment 5) verbal task. The simple task was implemented to disrupt verbal rehearsal, whilst the complex task was designed to disrupt verbal rehearsal and executive control. In Experiment 4, value effects emerged which were, somewhat unexpectedly, larger in the simple concurrent task condition. Significant value effects were also observed when participants completed a more complex verbal task (Experiment 5), although these boosts were accompanied by chance level performance at the majority of the less valuable positions.

Evidence that value effects emerged in all three experiments provides clear evidence that individuals can direct their attention to particularly valuable information in purely verbal WM tasks. But how might such effects be obtained? The most straightforward explanation of the current findings is that value effects in verbal WM do not critically rely on verbal rehearsal or attentional refreshing, with effects instead reflecting another process, such as enhanced encoding or consolidation. However, as the concurrent task was performed during encoding and maintenance, this account would assume that the critical process is independent of articulation and executive control. An alternative possibility is that the concurrent tasks disrupted, rather than completely blocked, maintenance mechanisms. For instance, individuals might have been able to rehearse a very limited amount of information between articulations of the days of the week and months of the year. Similarly, the complex verbal task might have reduced, rather than completely abolished, executive control resources. If so, these residual rehearsal abilities and/or executive resources could have been targeted towards the more valuable item. This would explain the pattern of findings in Experiments 4 and 5, whereby value effects emerged, alongside costs to other items. This is particularly clear in the final experiment, whereby value boosts were accompanied by chance level recall at the majority of the less valuable positions. Participants also omitted significantly more of the responses at the less valuable positions in the Prioritise-SP4 condition, indicating that they may have abandoned these items in order to retain the more valuable digit.

The observation of intact value effects in Experiment 5 would appear to contrast with recent work in the visual domain, which has indicated a reduced probe value effect under attentional load (Hu et al., 2016). These differences in findings might indicate that the cognitive mechanisms underlying value effects in WM differ depending on modality. Visuospatial WM appears to be more vulnerable to dual-task interference (e.g. Morey, 2018; Morey et al., 2013) and more closely associated with central executive control (Gray et al., 2017). Therefore, it may be that at least temporarily prioritising an item within a visual sequence is somewhat more demanding, compared to the verbal domain.

However, it is worth noting that the fate of low value items differed in Experiment 5 and Hu et al. (2016). As discussed above, there is evidence that participants in Experiment 5 may have abandoned the less valuable items in order to retain the more valuable digit. Although significant probe value costs were observed in Hu et al. (2016), memory for the low value items remained above chance for most, if not all, items. Differences in findings between these studies might then partly reflect changes in strategic approach motivated by methodological differences between experiments. The current study used serial recall, in which all items were assessed on every trial. Under these circumstances, a successful approach might be to focus remaining resources on preserving the more valuable digit, as this will be assessed on every trial. In contrast, Hu et al. (2016) employed a cued-recall paradigm, in which each serial position was assessed 25% of the time. In this case, where the more valuable item is unlikely to be tested, participants might abandon the prioritisation strategy under high cognitive load, and instead focus on remembering as many items as possible. This approach would reflect the flexible ways in which individuals apply limited-capacity WM and attentional control systems, dynamically shifting strategic approaches to optimise performance. Evidence of this would be consistent with findings that strategy use within a given task varies substantially depending on the retrieval method used (Morrison, Rosenbaum, Fair, & Chein, 2016). It would also be in line with the views of Macken et al. (2015), who suggest that performance on a memory task depends on factors relating to the individual, the material, and the task.

The distinct outcomes might also be explained by another methodological difference between the current study and previous work in the visual domain (e.g. Chapter 2; Allen & Ueno, 2018; Berry et al., 2018; Hitch et al., 2018; Hu et al., 2014; 2016). As participants were presented with nine digits in the present study, it was anticipated that they would struggle to keep track of the SPs whilst also retaining the digits presented. A star was therefore presented just before, and during, presentation of the more valuable digit to alert participants which item they should direct their attention towards. This differs from research exploring probe value effects in the visual domain, where no such cue has been presented (e.g. Chapter 2; Allen & Ueno, 2018; Hitch et al., 2018; Hu et al., 2014; 2016). One possibility is that the effects observed within this chapter do not entirely reflect the value manipulation, but at least partially reflect the use of the star cue. This may have increased the distinctiveness of this item during the encoding phase, thus potentially enhancing memory for it. The effects of distinctiveness on memory are well documented, with items that differ in a single feature being remembered better than items that are homogenous in nature (Hunt, 1995; Hunt & Lamb, 2001; Von Restorff, 1933). Moreover, when the item is physically distinct, there is evidence that distinctiveness does not require attention and occurs relatively automatically (Bireta & Mazzei, 2016). As such, it is possible that the effects observed in these experiments were not reliant on executive resources because they partially reflected a distinctiveness effect. However, if the effects reflects a combination of value and distinctiveness, and value boosts consistently rely on executive resources (Hu et al., 2016), one would expect the boost to be reduced in size when participants completed an attentionally-demanding concurrent task. This was not observed in Experiment 5, with the value boost being numerically larger when executive resources were reduced.

A related, but more extreme, possibility is that the effects described do not reflect value effects at all, with boosts entirely driven by distinctiveness. If distinctiveness captures attention involuntarily (Bireta & Mazzei, 2016), this account might predict the pattern of findings observed in Experiment 5. However, given the large effect sizes observed in visual WM where such cues have not been presented (e.g. Chapter 2), it is unlikely that participants ignored the value-based instructions within these experiments. It is also unlikely that participants were completely unable to prioritise the more valuable information, as value effects are well documented in visual WM (Chapter 2; Allen & Ueno, 2018; Hitch et al., 2018; Hu et al., 2014; 2016) and across domains in LTM (Adcock et al., 2006; Castel, Humphreys, et al. 2011; Gruber and Otten, 2010; Middlebrooks, Kerr, & Castel, 2017; Robison & Unsworth, 2017). Nevertheless, it would be beneficial for further research to investigate whether value effects emerge in verbal WM when the more valuable item is not perceptually distinct from the other items.

There were large individual differences in the size of the value effects, with some participants experiencing a large value boost at the targeted position, whilst others exhibited no effect or even a cost. Given that participants completed a FDR task in Experiments 3 and 4, further exploratory analysis was conducted to investigate whether forward digit span was related to the size of the value boosts and costs. No significant correlations were observed, with the BF analysis also slightly in favour of no effect for both correlations. This suggests that STM capacity might not be associated with the size of value effects. Caution should, however, be taken here as the BF analysis does not provide strong evidence to support this claim. Nevertheless, it would be beneficial for further research to investigate how other measures, such as complex span tasks that require both storage and processing, relate to value effects.

The questionnaires conducted at the end of each experiment revealed that participants had some awareness of the benefits and costs associated with directing attention to a particular item. Generally, individuals were aware that directing attention towards a digit might enhance memory for that item, but negatively affect others within the same trial. However, participants sometimes overestimated the costs and underestimated the benefits of prioritisation. Evidence that participants were generally aware of the value boosts is in line with findings from Berryhill et al. (2012), who found that participants believed that cueing enhanced memory. However, in this study, the cued item was tested on 100% of the trials, meaning that awareness of the effects to uncued items could not be tested. The current set of experiments therefore yielded a novel finding, by demonstrating that participants are also aware that directing attention towards a particular item may negatively impact others presented within the same trial.

A potential limitation of this set of studies is that the simple and concurrent tasks used (Experiments 4 and 5) involved participants articulating information whilst listening to the series of digits. This was implemented to ensure that the simple concurrent task disrupted verbal rehearsal, and the complex concurrent task disrupted both verbal rehearsal and executive resources. However, it may have potentially caused participants to miss some of the digits presented. Steps were, however, taken to minimise this. For instance, participants were presented with the digits through headphones and able to adjust the volume of the audio to ensure they could hear them. Furthermore, as participants were required to complete the concurrent tasks alongside all value conditions in Experiment 5, it is likely that any negative effect was equivalent across conditions.

The findings reported within this chapter may have theoretical implications. Value effects are at least partially considered to reflect the more valuable item being retained in an accessible, privileged state, such as the FoA (Chapter 2; Hu et al., 2014; 2016; Hitch et al., 2018). Taken together with findings from the visual domain (e.g. Chapter 2; Hitch et al., 2018; Hu et al., 2014; 2016), these experiments therefore suggest that the FoA is modality-general (Cowan, 2005; Hitch et al., 2018; Hu et al., 2014; Oberauer, 2013), and can hold verbal and visual information. Furthermore, these experiments might open up several new avenues of research. Within the literature, it has been suggested that probe value effects in WM are reliant on executive resources (Hitch et al., 2018; Hu et al., 2016). However, the current experiments indicate that this is unlikely to be true in all tasks. Although the effects reported here might partially or fully reflect distinctiveness, a more likely possibility is that the difference in findings reflect the domain assessed (i.e. visual vs verbal) or the retrieval method used (i.e. cued or serial recall). It would therefore be useful for research to examine which of these factors is likely to have resulted in different outcomes than those observed in previous research (e.g. Hu et al., 2016). More broadly, these findings also highlight the need for research to investigate the characteristics that define value effects and the FoA, and those that differ depending on task factors (Hitch et al., 2018).

These findings might also have practical implications. In the visual world, individuals can avert their gaze away from less relevant information in order to avoid encoding. However, such mechanisms do not exist with acoustic processing

(Macken, Phelps, & Jones, 2009), meaning that verbal information considered irrelevant or less relevant to current goals may be encoded and processed within WM (Macken et al., 2009). The ability to prioritise information in verbal WM is therefore of critical importance. The current experiments demonstrate that individuals can orient attention towards particularly valuable representations in this domain, even under extreme conditions where verbal rehearsal and executive control are disrupted. Importantly, these boosts do not always appear to come at a significant cost to less valuable items, particularly under normal circumstances where individuals are able to refresh or verbally rehearse information. However, before this manipulation is added to tasks that rely on verbal WM, research would first be required to investigate how prioritisation affects performance in everyday tasks.

3.7. Conclusions

The experiments within this chapter indicate that individuals can direct their attention to more valuable information in a purely verbal WM task, even when rehearsal and attentional mechanisms are reduced. It was also demonstrated that individuals have a fair (but not complete) understanding of the effects of value effects. Evidence of significant value effects extends previous findings which have demonstrated that individuals can prioritise valuable information that is presented visually (Chapter 2; Allen & Ueno, 2018; Hitch et al., 2018; Hu et al., 2014; 2016; Sandry et al., 2014). Conversely, evidence that value boosts emerged under complex concurrent task conditions contrasts with findings from the visual domain (Hu et al., 2016). This does not necessarily suggest that value effects in verbal WM do not rely on executive resources, however. Instead, the differences in findings could reflect other task factors, such as the nature of the retrieval test (i.e. serial vs cued-recall) or the presence of a visual cue accompanying the more valuable item. Further research is needed to investigate this. Nevertheless, regardless of the precise reason, it is clear that individuals can prioritise more valuable items when executive control resources are reduced under certain conditions. Doing so might have catastrophic effects on the less valuable items, however, which are either remembered poorly or completely abandoned.

CHAPTER 4

CAN CHILDREN PRIORITISE MORE VALUABLE INFORMATION IN WORKING MEMORY?

The experiments reported within this chapter are published as: Atkinson, A. L., Waterman, A. H., & Allen, R. J. (2019). Can children prioritize more valuable information in working memory? An exploration into the effects of motivation and memory load. *Developmental Psychology, 55*(5), 967–980.

4.1. Introduction

Chapters 2 and 3 demonstrated that adults can direct their attention to important items in visual and verbal WM. The current chapter builds on this by examining whether children aged 7-10 years old can also prioritise more valuable information in WM when they are sufficiently motivated to do so. Also of interest was whether these effects are observed across various memory loads, and modes of presentation (i.e. sequential and simultaneous presentation).

As discussed in Chapter 1, WM abilities do not reach adult-like levels until approximately 14-15 years of age (Gathercole, Pickering, Ambridge et al., 2004). At every stage of development, there are large individual differences in WM abilities, with some children exhibiting an impressive capacity whilst others show substantial impairments (Alloway, 2006). These individual differences in WM capacity have important implications in childhood, with many studies reporting a close relationship between WM and academic attainment (Alloway & Alloway, 2010; Alloway,

Alloway, & Wootan, 2014; Gathercole, Pickering, Knight et al. 2004; Holmes & Adam, 2006). For instance, Gathercole, Pickering, Knight et al. (2004) found that scores on WM tests were significantly associated with performance in English and mathematics at 7 years of age, and mathematics and science at 14 years of age. Similar findings were reported by Alloway and Alloway (2010), who found that WM was a better predictor of later performance in literacy and numeracy than IQ. This is perhaps unsurprising when one considers the vast number of classroom activities that require individuals to simultaneously store and process information. For instance, within classroom settings, instructions given to children are often lengthy, comprising several separate elements (Gathercole, Lamont, & Alloway, 2006), e.g., "finish what you are doing, come sit on the carpet, cross your legs and put your finger on your lips". These complex instructional sequences place substantial demands on WM, as children must successfully store the sequence whilst performing the actions (Gathercole et al., 2008; Jaroslawska et al., 2015; Waterman et al., 2017). Many subject specific activities, such as mental arithmetic and reading, are also thought to rely heavily on WM (Alloway, 2006; Fürst & Hitch, 1997; Seigneuric & Ehrlich, 2005; Seigneuric, Ehrlich, Oakhill, & Yuill, 2000).

Given that WM is required for many classroom activities, but that abilities are not fully developed in childhood, applying the limited capacity system as effectively as possible is particularly important for this group. As discussed previously, this could involve directing attention to particularly important or valuable information. Several studies have examined whether children can use cues to orient their attention within WM (Astle, Nobre, & Scerif, 2012; Shimi et al., 2014; Shimi & Scerif, 2015; 2017). For example, Shimi et al. (2014; Experiment 1) presented 7-year-olds, 11-year-olds, and adults with simultaneous arrays of four items, and tested memory for one of them following a brief delay. In the cueing conditions, the array was either preceded (pre-cue) or followed (retro-cue) by a cue informing participants which item would be tested at retrieval. These cues were 100% valid and always identified the item that would be later assessed. Performance in these conditions was compared to a neutral condition, in which no cue was presented. All groups significantly benefited from the pre-cues, although retro-cues only enhanced performance in the 11-year-olds and the adults. This demonstrates that, under certain conditions, children as young as 7-years-old can direct their attention to more-goal relevant information in WM.

Further work has revealed that cueing boosts may be modulated by similar factors in children and adults. For instance, Shimi and Scerif (2017) found no pre-cues or retro-cue benefits in 7-year old children and adults when two items were presented. In contrast, significant pre-cue and retro-cue boosts were observed in both groups when arrays contained four items. This demonstrates that the effects of cues differ depending on the number of items presented in both children and adults. This is line with several studies from the broader adult literature, which have reported that cueing boosts increase as memory load becomes larger (Astle, Sumemrfield et al. 2012; Kuo et al., 2012; Nobre et al., 2008; Souza et al., 2014; van Moorselaar et al., 2015). This is thought to reflect the strategic use of cues, whereby individuals attend to them more when the number of to-be-remembered items is increased (Shimi & Scerif, 2017). It is, however, important to note that the emergence of retro-cue effects in 7-year olds when four items were presented contrasts with Shimi et al. (2014), whereby no retro-cues emerged in this age group. It is unclear why this difference in findings might have emerged between the studies, although it may indicate that the presence or absence of a retro-cue boost is not particularly reliable in younger children. Nevertheless, taken together, these studies provide some evidence that children can direct their attention to information in WM based on visual cues, although younger children may not always be able to do this as accurately as older children or adults.

There is also evidence to suggest that children can benefit from increases in probe frequency, whereby one or more item is more likely to be tested than the rest. Evidence for this was provided by Cowan et al. (2010). In this study, 7-8-year olds, 12-13-year olds, and adults completed a change detection task that involved remembering the colour and location of circles and triangles for a brief period of time. Participants completed various attentional conditions, in which one of the shapes was tested 100% of the time, 80% of the time, 50% of the time, or 20% of the time. When four items were presented, all three groups were able to allocate their attention based on probe frequency. However, when the arrays contained six items, an interaction emerged between attention condition and age group, with the younger children less able to successfully distribute their attention. In line with the cueing literature, this indicates that 7-8-year-old children can direct their attention to more goal-relevant information in WM, although not always as effectively as older children or adults.

Relatively few studies have examined the effects of reward on memory in children (Berry et al., 2018; Castel, Humphreys, et al. 2011; Castel, Lee, Humphreys, & Moore, 2011). Castel, Humphreys et al. (2011) employed the value-directed remembering paradigm to investigate how the ability to selectively remember valuable items in episodic memory changes across the lifespan. Within the task, participants were shown series of words sequentially that were associated with a point value ranging from 1-12. They were then given 30 seconds to verbally recall as many words as possible in order to maximise their score. In order to investigate how selectivity changed with age, children (*M*. age = 8.14), adolescents (*M*. age = 14.52), young adults (*M*. age = 20.34), middle-aged adults (*M*. age = 56.66), young-old adults (*M*. age = 71.42) and old-old adults (*M*. age = 84.70) were recruited. A main effect of age group was found, with the children and adolescents being less selective than the young adults, the middle-aged adults, and the young-old adults. All groups showed some evidence of some selectivity, however, with the children and adolescents being approximately twice as likely to recall a word worth 12 points than a word worth 1 point.

To date, only one set of experiments have examined whether children can strategically prioritise more valuable information in WM (Berry et al., 2018). Within this series of experiments, 7-8, 8-9, and 9-10-year olds were presented with three items sequentially and asked to recall the colour of one probed item following a brief delay. Prior to encoding, participants were either told that all items were equally valuable (control condition) or that either the first item (Experiments 1 and 2) or the final item (Experiment 3) was relatively more valuable. As with previous experiments in adults (Hu et al., 2014; 2016; Hitch et al., 2018), this points system was notional, although children were informed that they would be given a sticker if they collected enough points. Across all experiments, performance at the more valuable SP did not differ in the priority and control conditions, with BF analysis providing evidence of no effects. This was taken as evidence that children cannot prioritise more valuable information in WM. As previous research in adults has indicated that prioritisation relies on executive resources (Hu et al., 2016), these findings suggest that children might not have the necessary cognitive resources to strategically prioritise more valuable information within WM (Berry et al., 2018). In contrast, significant recency effects were observed in all experiments, suggesting that children can benefit from more automatic forms of attentional direction (Allen et al., 2014; Hu et al., 2014).

There are, however, a few methodological features of Berry et al. (2018) which might explain why children were not able to prioritise the more valuable item. Firstly, in line with previous studies in adults (e.g. Chapters 2 and 3; Allen & Ueno, 2018; Hitch et al., 2018; Hu et al., 2014; 2016), the points system used was purely notional. Whilst this approach appears to motivate adults to prioritise information, it might not be sufficiently motivating for children. Indeed, it has been suggested that children may need more motivation to engage fully in psychological experiments compared to adults (Brewer et al., 2013). There is also evidence that tasks may underestimate children's abilities when they do not sufficiently engage the child or are not presented in an age-appropriate context (Borke, 1975; McGarrigle & Donaldson, 1974; Rose & Blank, 1974; Watanabe, 2017). Evidence of this has been provided by work investigating egocentrism in children. In 1956, Piaget and Inhelder developed the three mountains task, in which participants are asked to identify a doll's viewpoint of a scene whilst ignoring their own. From this, it was concluded that children younger than seven are egocentric, and therefore unable to understand that others have a different view of the world (Piaget & Inhelder, 1956). To explore this ability further, Borke (1975) developed a more child-friendly version of this task which used a toy model village and introduced a narrative to give the task meaning and context. Using this version, it was found that children as young as 3 years of age were not classified as egocentric. Similar findings have been reported for other cognitive constructs, such as conservation of liquid or number (Light, Buckingham, & Robbins, 1979; McGarrigle & Donaldson, 1974; Rose & Blank, 1974; Watanabe, 2017). As such, it is possible that children might be able to prioritise more valuable information in WM if the reward system is more meaningful and the task is more age appropriate.

Secondly, in Berry et al (2018) children were only ever presented with three item sequences. However, cueing effects vary as a function of memory load presented, with large effects observed when more items are presented (Astle, Summerfield et al., 2012; Kuo et al., 2012; Nobre et al., 2008; Shimi & Scerif, 2017; Souza et al., 2014; van Moorselaar et al., 2015). It is therefore possible that memory load will also influence whether children prioritise more valuable information using the probe value paradigm, and that the 3-item sequences used in Berry et al (2018) might not have tapped into this ability.

These issues were examined in the current experiments. The task was made more age-appropriate by placing it in the context of a story, with children able to use the points collected in a specially designed game at the end of the session. Children were also told they would win a prize if they collected enough points. Memory load was also manipulated, with either three or four items presented per sequence. As in Chapter 2 and most previous research employing the probe value paradigm (Berry et al., 2018; Hitch et al., 2018; Hu et al., 2014; 2016), Experiment 6 displayed items sequentially. This was implemented in order to closely mirror Berry et al. (2018). Experiment 7 then examined effects using simultaneous arrays. This was investigated for a number of reasons. Firstly, as Allen and Ueno (2018) recently demonstrated that adults can prioritise valuable information presented simultaneously, it would be useful to investigate whether children also possess this ability. Converging outcomes from sequential and simultaneous modes of presentation would also provide more convincing evidence that children can (or cannot) prioritise more valuable information in WM. In addition to this, such findings would bridge findings with other paradigms which have typically presented items simultaneously, such as cueing and probe frequency (Astle et al., 2012; Cowan et al., 2010; Shimi et al., 2014, Shimi & Scerif, 2017). Finally, these findings are likely to have practical importance, as information might be encountered sequentially (e.g. reading a book; Berry et al., 2018) or simultaneously (e.g. wall displays) in educational settings.

In both experiments, children aged 7-8 and 9-10-years old were recruited. Two age groups were tested in order to investigate whether the ability to prioritise valuable information increases with age. Proactive control strategies, in which individuals plan ahead for future responses and maintain task-relevant information before it is required, appear to develop throughout childhood (Chevalier, James, Wiebe, Nelson & Espy, 2014). As such, older children (aged 9-10 years) might be able to prioritise valuable information more effectively than younger children (7-8 years). The ages selected are also in line with the upper and lower age groups tested in Berry et al. (2018) and is similar to those used in other experiments that have investigated cueing effects in children (e.g. Shimi et al., 2014; Shimi & Scerif, 2017).

4.2. Experiment 6

Children completed a visual WM task, in which series of coloured shapes were presented sequentially. After a brief delay, the outline of one shape was presented, and participants were asked to recall the colour. Before encoding, participants were either told that all of the items were equal valuable (equal probe value) or that the first item was worth more points than the rest (differential probe value). The probe value manipulation was targeted at SP1 in order to retain consistency between these experiments and Chapter 2, which examined adults' ability to prioritise information in a similar task. The main comparison of interest was performance at SP1 in the differential and equal probe value conditions. Any primacy effect observed should therefore be present in both conditions and thus not act as a confound. The points system was made meaningful and age-appropriate through the incorporation of a story and game. At the start of the session, children were shown a 'friendly alien' whose planet had been invaded by 'evil aliens'. They were told they would collect energy points during the memory task which they could use to 'zap' the evil aliens in a specially designed game at the end of the session. They were also told that they would be given a prize if they collected enough points. To explore whether the probe value effects differed depending on the number of to-be-remembered items, memory load was manipulated, with either three or four items presented per trial.

Evidence of probe value boosts, whereby performance at the first position is higher in the differential probe value condition than the equal probe value condition, would indicate that children can prioritise more valuable information when they are sufficiently motivated to do so. In addition to this, if memory load is important within the current paradigm, an interaction between probe value and memory load should emerge. Based on the cueing literature (e.g. Shimi & Scerif, 2017), one might expect larger probe value effects when four items are presented. Given that proactive control strategies develop throughout childhood (Chevalier et al., 2014), it was also predicted that the older children would show larger probe value effects than the younger children. More generally, it was expected that there would be a main effect of memory load, with higher accuracy for the three item sequences than the four item sequences. It was also anticipated that a main effect of age group would emerge, with older children exhibiting higher accuracy than the younger children. However, as in Chapters 2 and 3, it was predicted that no main effect of probe value would be observed, as any boost at the more valuable item is likely to be offset by decreases in accuracy at the less valuable items.

4.2.1. Method

4.2.1.1. Participants

A primary school agreed to participate. The school is in a moderate socioeconomic status (SES) neighbourhood, with the index of multiple deprivation indicating that the area is amongst the 50% least deprived neighbourhoods in the country (Ministry of Housing, Communities & Local Governments, 2015). Children at the school are predominantly White-British and native English speakers. All children who participated spoke fluent English and had no known learning difficulties. Thirty-four younger children completed the experiment (aged 7-8 years; *M. age* = 7.92*, SD* = 0.30*;* 21 males). All children in this group were in Year 3 (UK). From this group, two children were removed for failing to engage with the articulatory suppression concurrent task and one child was removed as their performance was below chance level. The final analysis was therefore run on data from 31 younger children (*M. age* = 7.94*, SD* = 0.30*,* 18 males). Thirty-three older children also participated (aged 9-10 years; *M. age* = 9.93, *SD* = 0.30; 12 males). All children in this group were in Year 5 (UK). One child was absent on the second day of testing and therefore only participated in the 3-item condition. The final analysis for the 4-item conditions was therefore run on 32 older children (*M. age* = 9.93, *SD* = 0.31; 12 males), whilst the analysis for the 3-item conditions was run on all participants in this age group. Ethical approval for both experiments reported in this chapter was granted by the School of Psychology Ethics Committee at the University of Leeds (Ethics reference number: PSC-210).

4.2.1.2. Design, materials, and procedure

A 2 (Probe value: differential vs equal) x 2 (Memory load: 3-item vs 4-item) x 2 (Age group: younger children vs older children) mixed design was employed. Probe value and memory load were manipulated within-subject, whilst age group was a between-subject variable. The probe value and memory load conditions were blocked. Participants completed the experiment in two sessions, blocked by memory load. The sessions were completed on different days. The order of the memory load sessions, and the order of the probe value blocks within the memory load sessions, was counterbalanced. Within each probe value-memory load block, each SP was as likely to be tested. In the 3-item conditions, participants completed three practise trials and 30 experimental trials per block. In the 4-item conditions, participants completed four practise trials and 40 experimental trials per block. Within each block, each SP was tested 10 times.

The experimental paradigm used is displayed in Figure 4.1. Participants were presented with either three or four coloured shapes sequentially, with each presented for 500ms. There was a 250ms ISI between shapes. Stimuli were created by randomly pairing one of six colours (red, yellow, green, blue, purple, black) with one of six shapes (circle, triangle, cross, arch, flag, arrow). No colour or shape was repeated within the same trial. The shapes were presented on a white background at one of eight positions around a 2° imaginary circle located at the screen centre. All stimuli measured approximately 1.5°, based on a viewing distance of 50cm. After a delay of 1000ms, the outline of one shape was displayed in the centre of the screen and participants were asked verbally to recall the colour. Responses were recorded by the experimenter using a keyboard. This was implemented to ensure that the task was not too demanding for the children and that the correct keys were pressed. This is also in line with Chapter 2, in which the experimenter recorded the response for participants. As in Chapter 2, participants whispered the word ‘la’ to disrupt verbal recoding (Baddeley, 1986). This was implemented as research suggests that the age groups tested verbalise to-be-remembered visual information (Flavell, Beach, & Chinsky, 1966; Conrad, 1971; Hitch & Halliday, 1983; Hitch, Halliday, Dodd, & Littler, 1989). Participants were informed of the probe value manipulation during the instructions. In the differential probe value condition, they were told that correct recall of the first shape would earn them four ‘energy points’, and that correct recall of any other shape would earn them one ‘energy point’. In the equal probe value condition, they were told that correct recall of any item would earn them one ‘energy point’.

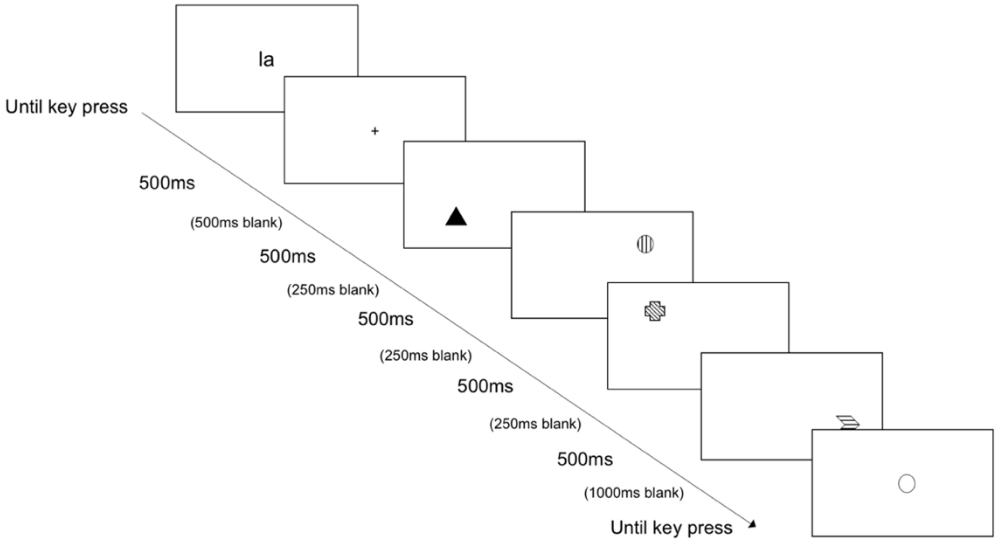
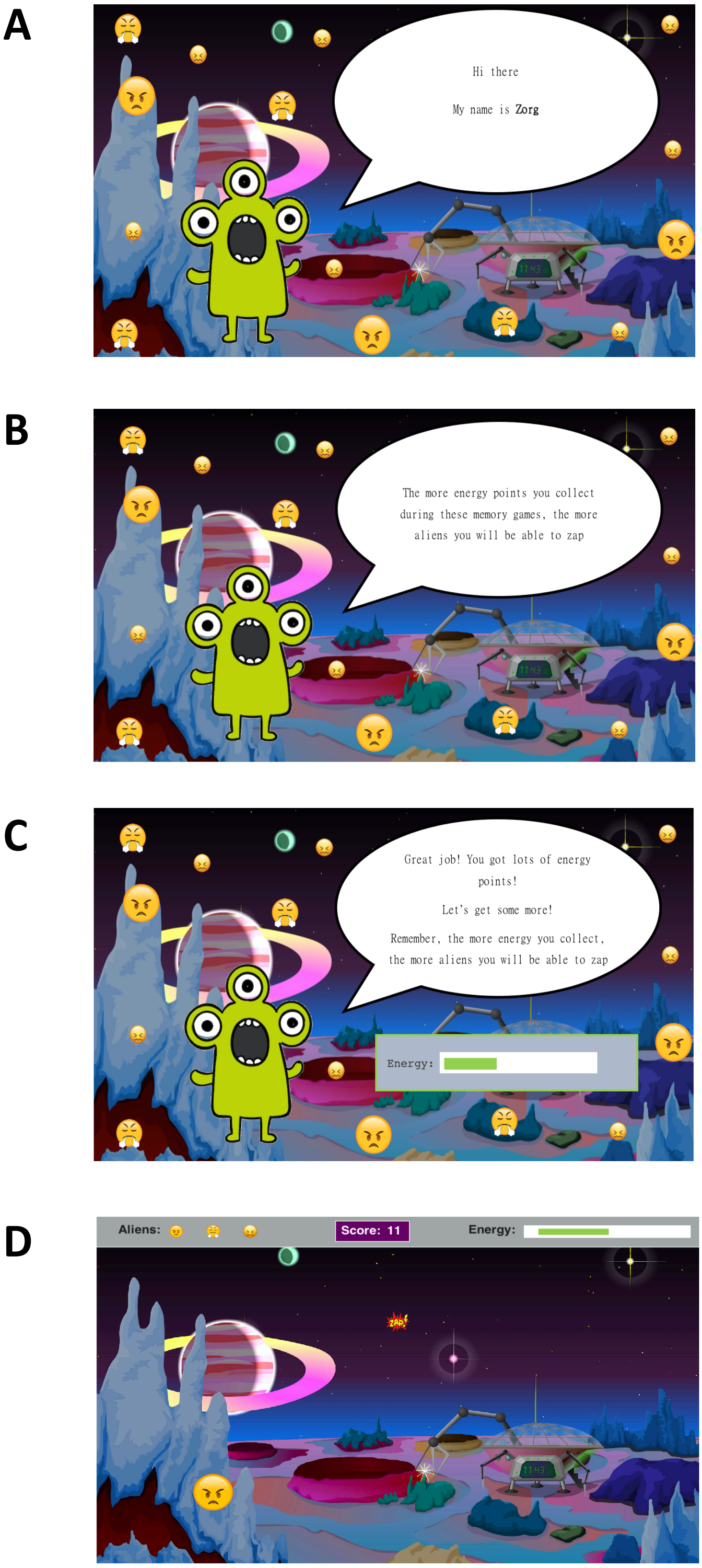


Figure 4.1. The experimental paradigm used in Experiment 6, with a 4-item trial as an illustrative example. The different types of shading reflect different colours (e.g. red, green, blue). Figure not to scale.

Before the instructions of the memory task, children were introduced to a friendly alien named 'Zorg' and were told a short story about him (see Figure 4.2A and B, and Appendix D for the full text). They were told that his planet had been invaded by evil aliens, and asked if they would help 'zap' them. To zap the aliens, they needed energy points, which could be collected by playing memory games (i.e. the experimental task). They were also told that they would get a prize if they collected enough points.

After every 10 trials, children were shown an energy bar that slowly increased throughout the session. The increase in energy was not linked to their true performance and increased by the same amount for each child (see Figure 4.2C). This was implemented to ensure that motivation was not affected by prior performance and that children were not discouraged if they performed poorly.

Participants were also reminded of the probe value instructions directly after these screens.



Hi there

My name is Zorg

The more energy points you collect during these memory games, the more aliens you will be able to zap

Great job! You got lots of energy points!

Let's get some more!

Remember, the more energy you collect, the more aliens you will be able to zap

Figure 4.2.

Sample images from the alien story and game used in each experiment. Before the start of the task, children were told a story about an alien (e.g. A and B; see Appendix D for the full story). After every 10 experimental trials, the alien would appear and tell the child how much energy they had collected (C). They then played the 'zap an alien' game at the end of each session (D).

At the end of the session, children were told they had accumulated enough energy points from the memory games to play the 'zap an alien' game (see Figure 4.2D). In this game, the 'evil aliens' appeared on screen and participants had to click on them before they disappeared in order to 'zap' them. Children were told the same story and played the same game in both sessions. At the end of each session, children were told they had collected enough energy points to receive a prize and chose a piece of stationery as a reward. All children played the game and received the prize regardless of how well they performed on the task.

4.2.1.3. Data analysis

The dependent variable was accuracy, determined by proportion of trials answered correctly. This was calculated by dividing the number of trials responded correctly by the number of trials responded correctly plus the number of trials responded incorrectly. Separate analysis was conducted for the 3-item and 4-item conditions as the number of SPs differed between the conditions. As some previous research has suggested that older children can direct their attention in WM more effectively than younger children (Cowan et al., 2010; Shimi et al., 2014), age group was also added as a factor.

4.2.2. Results

4.2.2.1. Three items

Mean accuracy (and SE) as a function of probe value, SP and age group is displayed in Figure 4.3A. Means and SE are also presented in Table 4.1. A 2 (Probe value: differential vs equal; within-subject) x 3 (SP: 1-4; within-subject) x 2 (Age group: Younger children vs Older children; between-subject) mixed ANOVA was conducted. This revealed no significant effect of probe value ($F(1, 62) = .42$, $MSE = .02$, $p = .521$, $\eta_p^2 < .01$; $BF_{10} = 0.13$, $BF_{01} = 7.69$), demonstrating that this manipulation did not affect overall performance on the task. There was, however, a significant effect of age group ($F(1, 62) = 4.43$, $MSE = .08$, $p = .039$, $\eta_p^2 = .07$; $BF_{10} = 1.17$), with older children ($M = .60$, $SE = .02$) exhibiting higher accuracy than younger children ($M = .54$, $SE = .02$). There was also a significant effect of SP (*GG-corrected* $F(1.52, 94.33)= 35.64$, $MSE = .06$, $p < .001$, $\eta_p^2 = .37$; $BF_{10} > 10{,}000$). Bonferroni-Holm pairwise comparisons revealed significant differences between SP1 ($M = .53$, $SE = .02$) and SP2 ($M = .48$, $SE = .02$; $p = .022$), SP1 and SP3 ($M = .70$, $SE = .02$; $p < .001$), SP2 and SP3 ($p < .001$). There was a significant interaction between probe value and SP (*GG corrected* $F(1.71, 106.27) = 13.13$, $MSE = .03$, $p < .001$, $\eta_p^2 = .18$; $BF_{10} = 186.72$), but no other interactions ($F \leq 2.21$, $p \geq .123$, $BF_{10} \leq 0.34$, $BF_{01} \geq 2.94$). These findings were corroborated by BF analysis, which indicated that the most likely model included main effects of SP and age group, and an interaction between probe value and SP ($BF > 10{,}000$ relative to the null model containing only participant).

To investigate the interaction between probe value and SP, a series of paired sample t-tests (corrected using Bonferroni-Holm) were conducted to compare performance in the differential and equal probe value conditions at the various SPs.

At SP1, accuracy in the differential probe value condition (*M* = .57, *SE* = .03) was significantly higher than accuracy in the equal probe value condition (*M* = 50, *SE* = .03; *t*(63) = 3.00, *p* = .008, *d* = 0.37; BF_{10} = 7.74). This outcome was also observed at SP2 (*Differential M* = .51, *SE* = .02; *Equal M* = .45, *SE* = .02; *t*(63) = 2.34, *p* = .022, *d* = 0.29, BF_{10} = 1.72). At SP3, the opposite pattern of results was observed, with participants performing significantly better in the equal probe value condition (*M* = .75, *SE* = .02) than the differential probe value condition (*M* = .65, *SE* = .03; *t*(63) = -3.36, *p* = .003; *d* = -0.42, BF_{10} = 20.09).

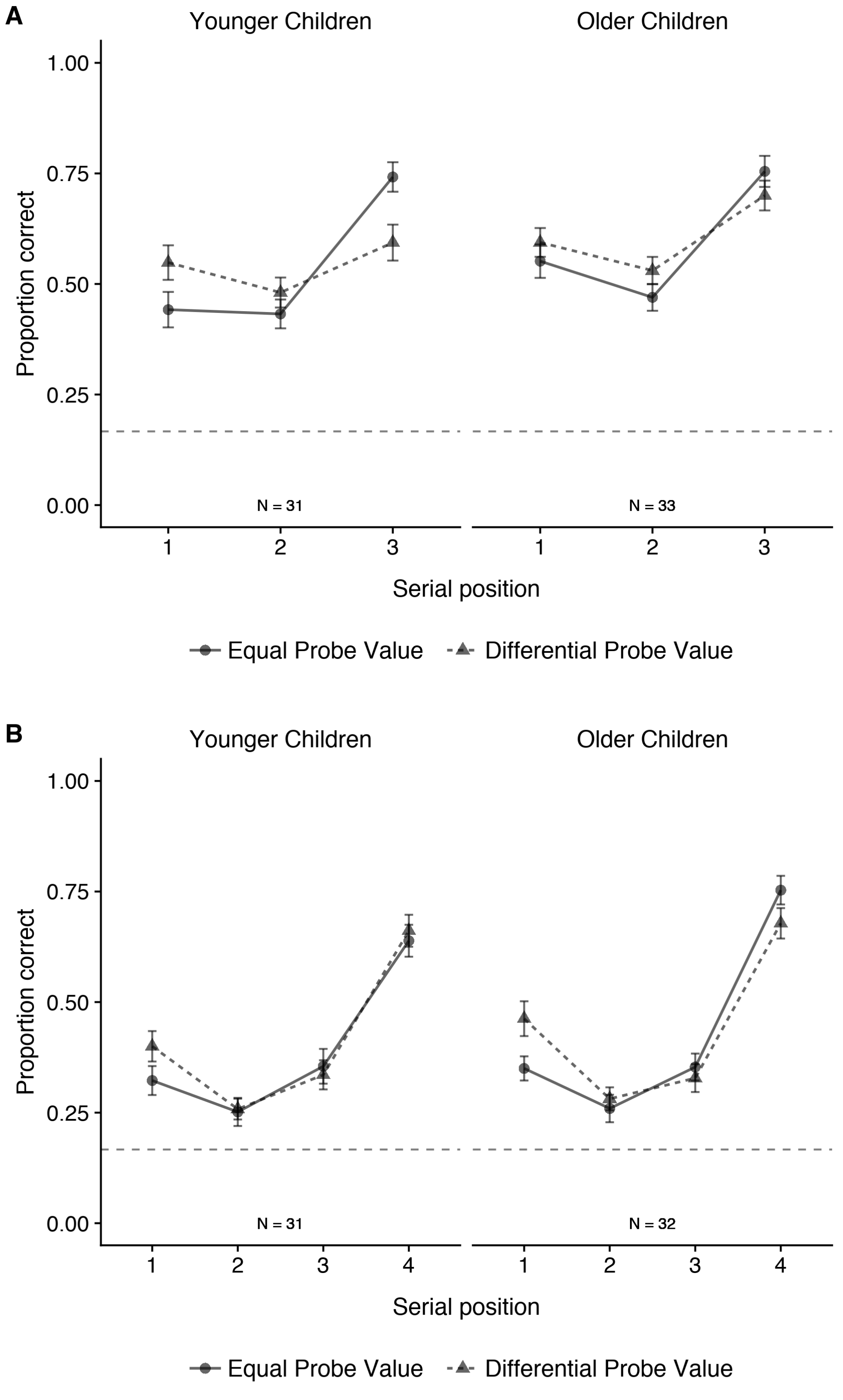


Figure 4.3. Mean proportion correct in Experiment 6 as a function of probe value, serial position, and age group in the 3-item (A) and 4-item conditions (B). Error bars denote standard error, and the dotted line indicates chance performance.

4.2.2.2. Four items

Mean accuracy (and SE) as a function of probe value, SP and age group is displayed in Figure 4.3B. Means and SE are also presented in Table 4.1, collapsed across age groups. A 2 (Probe value: differential vs equal; within-subject) x 4 (SP: 1-4; within-subject) x 2 (Age group: Younger children vs Older children; between-subject) mixed ANOVA revealed no effect of probe value ($F(1, 61) = 1.62$, $MSE = .02$, $p = .208$, $\eta_p^2 = .03$; $BF_{10} = 0.16$, $BF_{01} = 6.25$), demonstrating that this manipulation had no effect on overall performance. There was also no main effect of age group ($F(1, 61) = 2.16$, $MSE = .05$, $p = .147$, $\eta_p^2 = .03$; $BF_{10} = 0.33$, $BF_{01} = 3.03$). There was, however, a significant effect of SP (*GG corrected* $F(2.57, 156.84) = 102.64$, $MSE = .05$, $p < .001$, $\eta_p^2 = .63$; $BF_{10} > 10{,}000$). Bonferroni-Holm pairwise comparisons revealed significant differences between SP1 ($M = .38$, $SE = .02$) and SP2 ($M = .26$, $SE = .02$; $p = .001$), SP1 and SP4 ($M = .68$, $SE = .02$; $p = .001$), SP2 and SP3 ($M = .34$, $SE = .02$; $p = .002$), SP3 and SP4 ($p = .001$), and SP3 and SP4 ($p = .001$) An interaction between probe value and SP emerged ($F(3, 183) = 4.06$, $MSE = .03$, $p = .008$, $\eta_p^2 = .06$; $BF_{10} = 1.32$), although no other interactions were observed ($F \leq 1.11$, $p \geq .348$; $BF_{10} \leq 0.15$, $BF_{01} \geq 6.67$). The BF analysis indicated that the most likely model included main effects of SP and an interaction between probe value and SP ($BF_{10} > 10{,}000$ relative to the null model containing only participant).

To investigate the interaction between probe value and SP, a series of paired sample t-tests were conducted to compare performance in the differential and equal probe value conditions at the various SPs (corrected using Bonferroni-Holm). At SP1, performance in the differential probe value condition ($M = .43$, $SE = .03$) was significantly better than performance in the equal probe value condition (*M* = .34, *SE* = .02; *t*(62) = 3.31, *p* = .008; *d* = 0.42, BF_{10} = 17.83). There were no significant differences at the other SPs (*t* ≥ -1.04 and ≤ 0.54, *p* ≥ .906, *d* ≥ -0.13 and ≤ 0.07*;* BF_{10} ≤ 0.23, BF_{01} ≥ 4.35).

Table 4.1.

Mean accuracy (and SE) in Experiment 6 as a function of probe value, memory load and SP, collapsed across age group. N = 64 for the 3-item conditions and N = 63 for the 4-item conditions.

		SP1	SP2	SP3	SP4
3 items	Differential probe value	.57 (.03)	.51 (.02)	.65 (.03)	-
	Equal probe value	.50 (.03)	.45 (.02)	.75 (.02)	-
4 items	Differential probe value	.43 (.03)	.27 (.02)	.33 (.02)	.67 (.02)
	Equal probe value	.34 (.02)	.26 (.02)	.35 (.02)	.70 (.03)

4.2.2.3. Across memory loads

Accuracy at SP1 is displayed in Figure 4.4A as a function of probe value and memory load. Aggregated data for individual participants is displayed, as well as mean accuracy across participants (and SE). To investigate whether the probe value boosts differed across memory loads, a 2 (Probe value: differential vs equal: within-subject) x 2 (Memory load: 3-item vs 4-item; within-subject) x 2 (Age group: Younger children vs Older children' between-subject) mixed ANOVA was conducted at SP1. One participant in the older children's group completed only the 3-item condition and was therefore excluded from this analysis. The analysis was therefore run on data from 31 younger children and 32 older children. This revealed a main effect of probe value ($F(1, 61) = 22.42, MSE = .02, p < .001, \eta_p^2 = .27$; $BF_{10} = 753.71$), whereby performance was higher in the differential probe value condition ($M = .50, SE = .02$) than the equal probe value condition ($M = .41, SE = .02$). There was also a main effect of memory load ($F(1, 61) = 40.91, MSE = .03, p < .001, \eta_p^2 = .40$; $BF_{10} > 10{,}000$), whereby performance was higher in the 3-item condition ($M = .53, SE =. 02$) than the 4-item condition ($M = .38, SE = .02$). There was no significant effect of age group ($F(1, 61) = 2.51, MSE = .08, p = .118, \eta_p^2 = .04$; $BF_{10} = 0.69, BF_{01} = 1.45$), and no interactions ($F \leq 1.33$, p ≥ .254, $BF_{10} \leq 0.42, BF_{01} \geq 2.38$). BF analysis revealed that the most likely model included main effects of probe value and memory load ($BF_{10} > 10{,}000$ relative to the null model containing only participant).

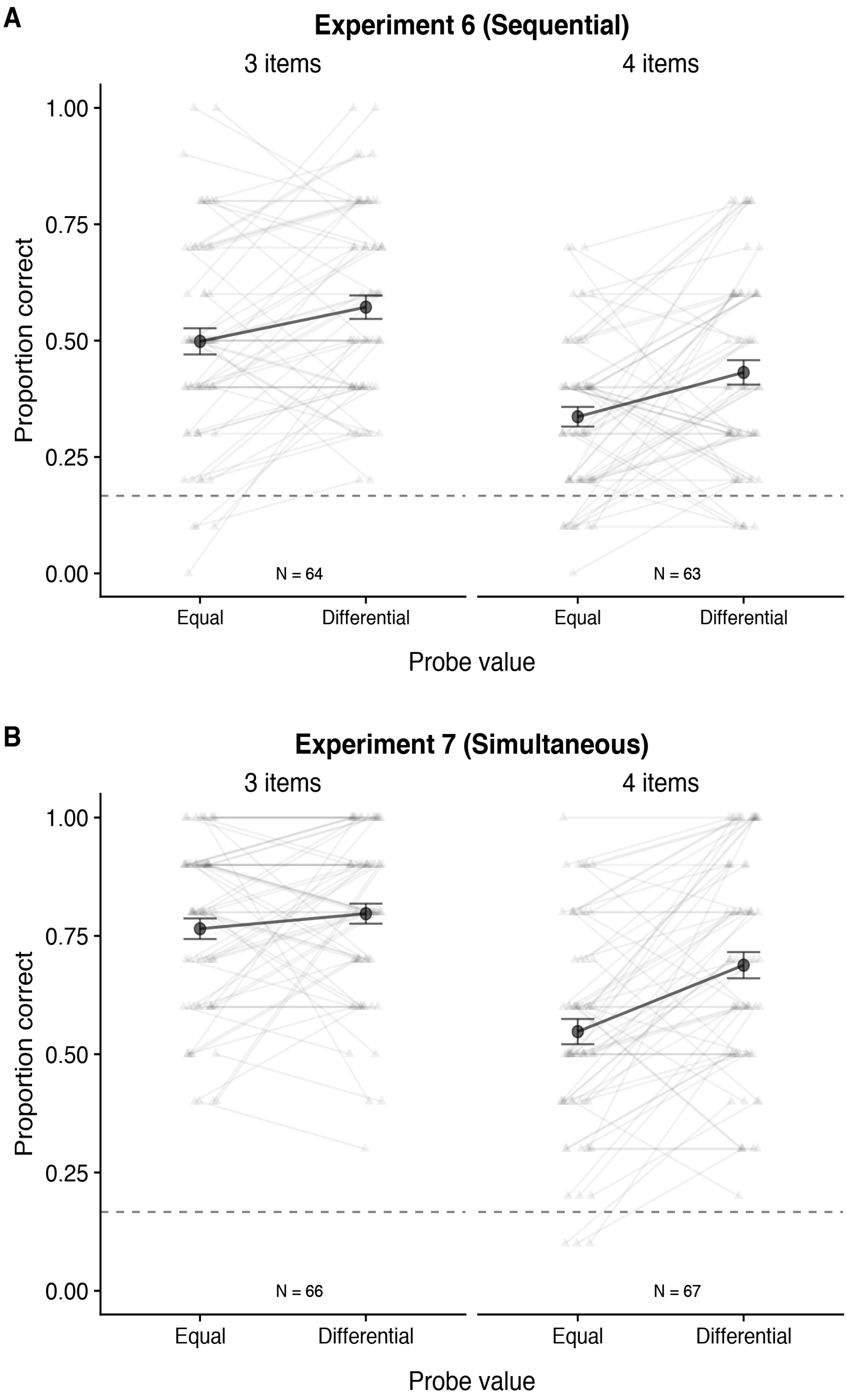


Figure 4.4. Accuracy at SP1 in Experiment 6 (A) and the top-left position in Experiment 7 (B) as a function of probe value and memory load. Performance is collapsed across age groups, as there were no interactions containing this factor in

either experiments. The lighter lines with triangular points display the mean accuracy in each condition for individual participants. The darker, bolder line with circle points display the mean across participants, whilst the error bars denoting SE. The dotted horizontal line reflects chance guessing rate. An upward trend is indicative of a prioritisation boost, whereby participants performed better at the more valuable item in the differential probe value condition.

4.2.2.4. Comparison with Berry et al. (2018)

Experiment 2 of Berry et al. (2018) employed a similar design to the 3-item condition in the current experiment, except that task was not placed within the context of a child-friendly story. As the data from Berry et al. (2018) is publicly available on the Open Science Framework (https://osf.io/xgrnc), additional analysis was conducted to directly compare the data from these experiments. Such findings would reveal whether the difference in findings can be attributed to the inclusion of the story and game in the current experiment. It would also allow us to explore whether gamification significantly enhances performance on WM tasks in children, or whether it simply encourages the child to comply with task instructions and complete the task optimally.

In Berry et al. (2018), children aged 7-8 years (Year 3), 8-9 years (Year 4) and 9-10 years (Year 5) were tested. The 8-9-year-old group (Year 4) were omitted from this analysis, as this age group were not tested in the current experiment. Data were available for 29 Year 3s (referred to as younger children hereafter) and 30 Year 5s (referred to as older children hereafter). Participants were removed if they had special educational needs (11 participants), data was not available for both probe value conditions (4 participants), or their age fell well outside of the expected range for the year group (e.g. one participant in the younger children's group was removed as their age was recorded as 9.55 years). The final analysis was therefore run on 107 participants. This included 64 participants who were included in the 3-item analysis of the current experiment (31 younger children and 33 older children), and 43 participants from Experiment 2 of Berry et al., (2018; 21 younger children and 22 older children).

Performance as a function of probe value, SP, and experiment is displayed in Figure 4.5A. Accuracy at SP1 as a function of probe value and experiment is presented in Figure 4.5B. As no significant interaction were observed between age group and any other variable in either the current experiment or Berry et al. (2018), this variable was omitted from the analysis. A 2 (Probe value; differential vs equal) x 3 (SP; 1-3) x 2 (Experiment; current vs Berry et al., 2018) mixed ANOVA was conducted, as well as a BF equivalent. Probe value and SP were within-subject variables whilst experiment was manipulated between-subject. Evidence of an interaction between probe value, SP, and experiment would be expected based on the differing outcomes when the experiments are analysed separately. This would provide further evidence that the additional motivational elements encouraged children to comply with the task instructions and apply the probe value strategy. Furthermore, evidence of a main effect of experiment, whereby performance was higher in the current experiment, would suggest that overall performance was enhanced by the story and game.

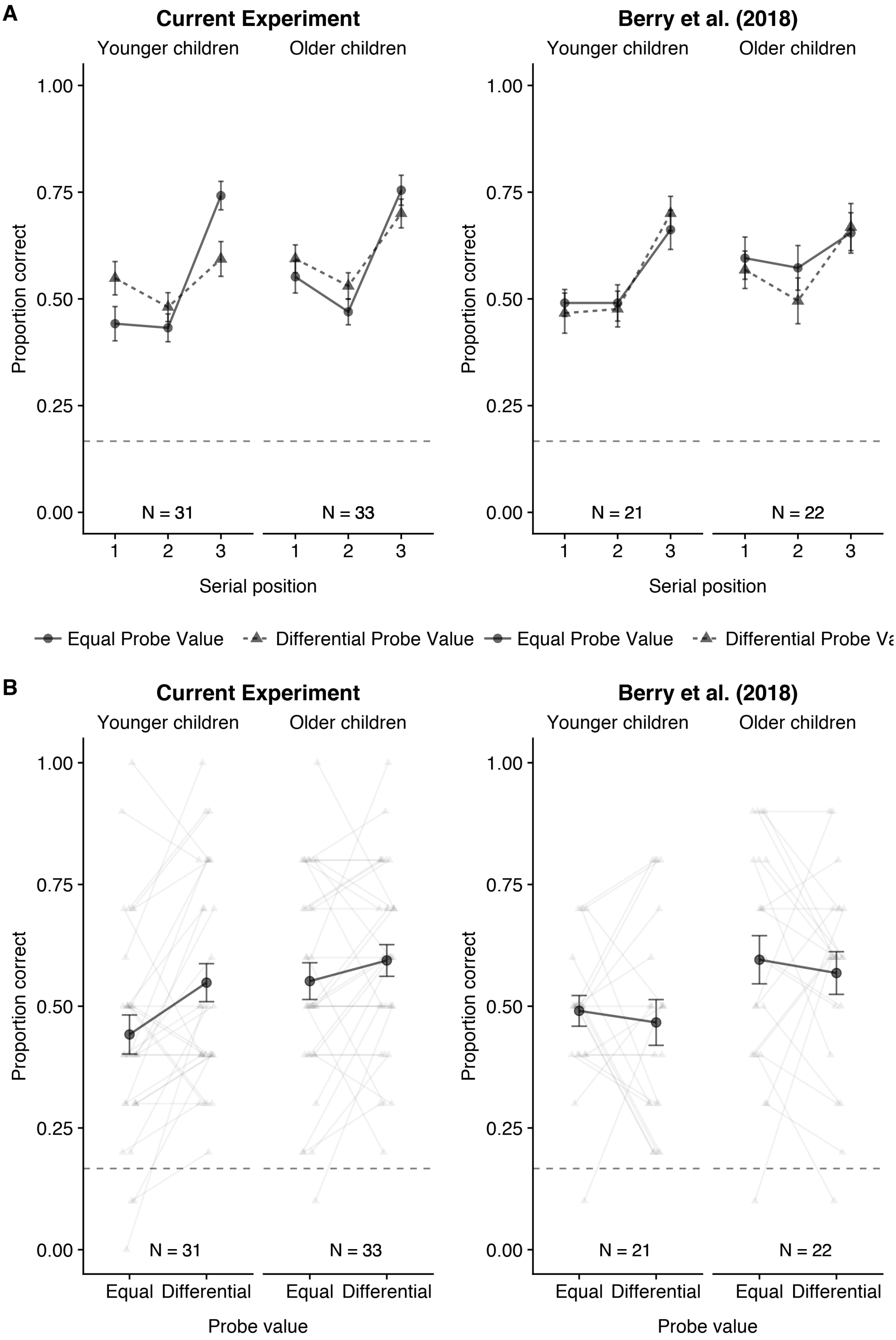


Figure 4.5. A) Mean accuracy as a function of probe value, SP, and experiment. Error bars reflect SE. B) Mean accuracy at SP1 as a function of probe value, and experiment. The light grey lines with triangular points display the mean accuracy for

each participant. The darker line with circular points reflects mean accuracy across as participants, with the error bars displaying SE.

The analysis revealed a main effect of SP (*GG corrected* $F(1.61, 168.89) = 40.27$, $MSE = .06$, $p < .001$, $\eta_p^2 = .28$, $BF_{10} > 10{,}000$), whereby accuracy at SP3 ($M = 0.69$; $SE = .02$) was higher than SP1 ($M = 0.53$; $SE = .02$; $p < .001$) and SP2 ($M = 0.49$; $SE = .02$; $p < .001$). There was no overall main effect of probe value ($F(1, 105) = .07$, $MSE = .02$, $p = .798$, $\eta_p^2 < .01$; $BF_{10} = 0.09$, $BF_{01} = 11.11$) or experiment ($F(1, 105) < .01$, $MSE = .08$, $p = .990$, $\eta_p^2 < .01$; $BF_{10} = 0.14$, $BF_{01} = 7.14$). As predicted, there was a significant interaction between probe value, SP, and experiment ($F(1.89, 198.44) = 8.52$, $MSE = .03$, $p < .001$, $\eta_p^2 = .08$, $BF_{10} = 48.71$). No other significant interactions emerged ($F \leq 1.96$, $p \geq .147$, $BF_{10} \leq 0.17$, $BF_{01} \geq 5.88$). Supporting the frequentist findings, the BF analysis indicated that the most likely model included a main effect of SP and an interaction between probe value, SP, and experiment ($BF_{10} > 10{,}000$ relative to the null model with random effects of participant only).

In order to investigate the interaction further, repeated measures ANOVAs were conducted to investigate whether an interaction emerged between probe value and SP in the different experiments. In the current experiment, there was a main effect of SP (*GG corrected* $F(1.53, 96.09) = 36.17$, $MSE = .06$, $p < .001$, $\eta_p^2 = .37$, $BF_{10} > 10{,}000$), whereby accuracy at SP3 ($M = 0.70$; $SE = .02$) was higher than SP1 ($M = 0.54$; $SE = .02$; $p < .001$) and SP2 ($M = .48$; $SE = .02$; $p < .001$). No main effect of probe value emerged ($F(1, 63) = .44$, $MSE = .02$, $p = .508$, $\eta_p^2 < .01$; $BF_{10} = 0.13$, $BF_{01} = 7.69$), although there was a significant interaction between probe value and SP (*GG corrected* $F(1.69, 106.55) = 12.60$, $MSE = .03$, $p < .001$, $\eta_p^2 = .17$, $BF_{10} = 184.64$). Paired sample t-tests revealed that performance was better in the differential probe value condition (M = .57, SE = .03) than the equal probe value condition at SP1 (M = 50, SE = .03; $t(63) = 3.00$, p = .008, d = 0.37; BF_{10} = 7.74) and SP2 (*Differential* M = .51, SE = .02; *Equal* M = .45, SE = .02; $t(63) = 2.34$, p = .022, d = 0.29, BF_{10} = 1.72). This pattern was reversed at SP3, where accuracy was significantly higher in the equal probe value condition (M = .75, SE = .02) than the differential probe value condition (M = .65, SE = .03; $t(63) = -3.36$, p = .003; d = -0.42, BF_{10} = 20.09).

In Berry et al. (2018; Experiment 2), there was a main effect of SP (*GG corrected* $F(1.70, 71.37) = 10.94$, MSE = .07, $p < .001$, $\eta_p^2 = .21$; $BF_{10} > 10{,}000$), whereby accuracy at SP3 (M = 0.67; SE = .03) was higher than SP1 (M = 0.53; SE = .03; p = .020) and SP2 (M = .51; SE = .03; $p < .001$). There was, however, no main effect of probe value ($F(11, 42) = .58$, MSE = .03, p = .449, $\eta_p^2 = .01$; BF_{10} = 0.16, BF_{01} = 6.25). Importantly, there was also no significant interaction between probe value and SP (*GG corrected* $F(1.71, 71.90) = 1.00$, MSE = .04, p = .362, $\eta_p^2 = .02$; BF_{10} = 0.14, BF_{01} = 7.14).

4.2.3. Discussion

This experiment examined whether children can prioritise more valuable information in WM by increasing the motivation and age-appropriate nature of the reward system surrounding the task. Also of interest was whether probe value effects would vary as a function of memory load. In contrast to Berry et al (2018), significant probe value effects emerged, whereby participants exhibited higher performance at SP1 in the differential probe value condition relative to the equal probe value condition. This suggests that children are able to prioritise a more valuable item in order to facilitate later recall, provided that they are sufficiently motivated to do so. No interaction between probe value and memory load was observed, indicating that the increased motivational aspects of the task enabled children to prioritise with both 3- and 4-item sequences. As such, the absence of a probe value effect in Berry et al. (2018), where only 3-item sequences were used, cannot simply be attributed to an explanation based on memory load.

Moreover, additional analysis comparing the data from the current experiment and that obtained in Experiment 2 of Berry et al. (2018) revealed a significant interaction between probe value, SP, and experiment. Ultimately, this interaction was driven by probe value effects at SP1 and SP2 in the current experiment, but no such boosts in Berry et al. (2018). This therefore provides further evidence that the increased motivational aspects of the current experiment resulted in the emergence of probe value effects. Interestingly, no overall effect of experiment emerged in this analysis, indicating that the additional motivational elements in the current experiment did not enhance overall performance on the task.

There were few effects involving age. The older children exhibited better overall performance when 3-items were presented, but there were no age differences for the 4-item task. The main effect of age group was also absent in the analysis that collapsed across memory loads. More importantly, there were no interactions involving age in any of the analyses, indicating that the ability to prioritise valuable information in WM does not undergo substantial developmental changes between the ages of 7-10 years.

Evidence that the probe value boosts did not differ across memory load contrasts with the cueing literature, in which effects of cues are larger when more items are presented (Shimi & Scerif, 2017). However, one key distinction between these experiments is the mode of presentation used; Experiment 6 presented items sequentially, whilst Shimi and Scerif (2017) used simultaneous arrays. Experiment 7 therefore examined the impacts of probe value and memory load on visual WM for simultaneously presented arrays. This exploration will not only address the claims of Shimi and Scerif (2017), but will also connect research on probe value with the broader developmental literature on cueing and probe frequency effects, which have used simultaneous rather than sequential presentation (Cowan et al., 2010; Astle et al., 2012; Shimi et al., 2014; Shimi & Scerif, 2017).

4.3. Experiment 7

Experiment 7 examined whether children can direct their attention to valuable information in simultaneous arrays, and whether such effects vary as a function of memory load. Previous research has revealed important distinctions between sequential and simultaneous modes of presentation. For example, accuracy on WM tasks is considerably higher when information is presented simultaneously (Allen et al., 2006; Gorgoraptis et al., 2011; Morales, Calvo, & Bialystok, 2013). Moreover, research in adults has suggested that it may be easier to direct attention in WM when information is presented simultaneously (Gorgoraptis et al., 2011). This is thought to occur as one does not need to protect the item from further incoming stimuli when items during simultaneous arrays, a process that is essential when information is presented sequentially (Gorgoraptis et al., 2011).

Allen and Ueno (2018) recently demonstrated that adults are able to prioritise more valuable information in WM using simultaneous arrays. However, to date, research has not investigated whether children are also able to prioritise valuable information in WM when a simultaneous mode of presentation is used. Such a skill would be beneficial for children, as the visual environment often contains multiple items that vary in value or importance. Experiment 7 therefore investigated this.

Evidence of probe value effects would replicate Experiment 6, providing further evidence that children are able to prioritise more valuable information in WM. Also of interest was whether an interaction would emerge between probe value and memory load. Evidence of an interaction, whereby large probe value effects are observed in the 4-item condition, would be in line with Shimi and Scerif (2017). Such findings would, however, contrast with Experiment 6, suggesting that the mode of presentation is important when considering the effect of memory load on the ability to direct attention. Conversely, equivalent effects across memory loads would replicate Experiment 6, potentially highlighting an important distinction between probe value and cueing effects. It was unclear whether the probe value effects would increase with age, as this interaction was not observed in Experiment 6. However, such findings might be predicted based on previous findings suggesting that proactive control increase with age (Chevalier et al., 2014). More generally, a significant effect of memory load was expected, with participants exhibiting higher accuracy when three items were presented. This would be in line with Experiment 6. A significant effect of age group was also predicted, although this was only observed in the 3-item condition of Experiment 6.

4.3.1. Method

4.3.1.1. Participants

Children were recruited from the same primary school as in Experiment 6, although no participants had taken part in the previous experiment. Thirty-five younger children took part (7-8 years; Year 3; *M. age* = 8.01, *SD* = 0.29; 15 males). Two children were excluded for not properly engaging in the articulatory suppression task. Due to absence, one child from this group only completed the 4-item conditions. The final analysis for the 3-item conditions was therefore run on 32 younger children (*M. age* = 8.01, *SD* = 0.29; 13 males), whilst the analysis for the 4-item conditions was run on 33 younger children (*M. age* = 8.00, *SD* = 0.29; 13 males). Thirty-four older children (aged 9-10 years; Year 5; *M. age* = 9.82, *SD* = 0.25; 20 males) also participated in both sessions.

4.3.1.2. Design, materials, and procedure

With the exception of a few minor details relating to the presentation mode, the materials, design, and procedure were identical to Experiment 6. The experiment employed a 2 (Probe value: differential vs equal) x 3 (Memory load: 3-item vs 4-item) x 4 (Spatial location (SL): top-left, top-right, bottom-left, bottom-right) x 2 (Age group: younger children, older children) mixed design. Probe value, memory load, and SL were within-subject variables, whilst age group was a between-subject variable. In the differential probe value condition, participants were told that the top-left item was worth 4 points and the other items were worth 1 point. This SL was selected as it was considered to have properties that are most similar to SP1 (i.e. it may be the item participants are most likely to look at first). In the equal probe value condition, all of the items were worth 1 point. As in Experiment 6, trials were blocked by probe value and memory load. Children completed two sessions on separate days, blocked by memory load. The order of the memory load blocks, and the order of the probe value blocks within the memory load blocks, was counterbalanced. In each probe value-memory load block, there were 40 trials, with each SL being assessed 10 times. The experimental paradigm used is displayed in Figure 4.6. Participants were shown arrays of three or four coloured shapes simultaneously. Shapes appeared at one of four SLs positioned at the corners of a 2° imaginary circle, located at the centre of the screen. The arrays were displayed for 1500ms in the 3-item blocks and 2000ms in the 4-item blocks. In the 3-item conditions, an item was always presented in the top-left location as this is the SL at which the probe value manipulation was targeted. The other SLs were selected randomly. In the 4-item conditions, all SLs were occupied on every trial. The retention interval and the suppression task were identical to Experiment 6.

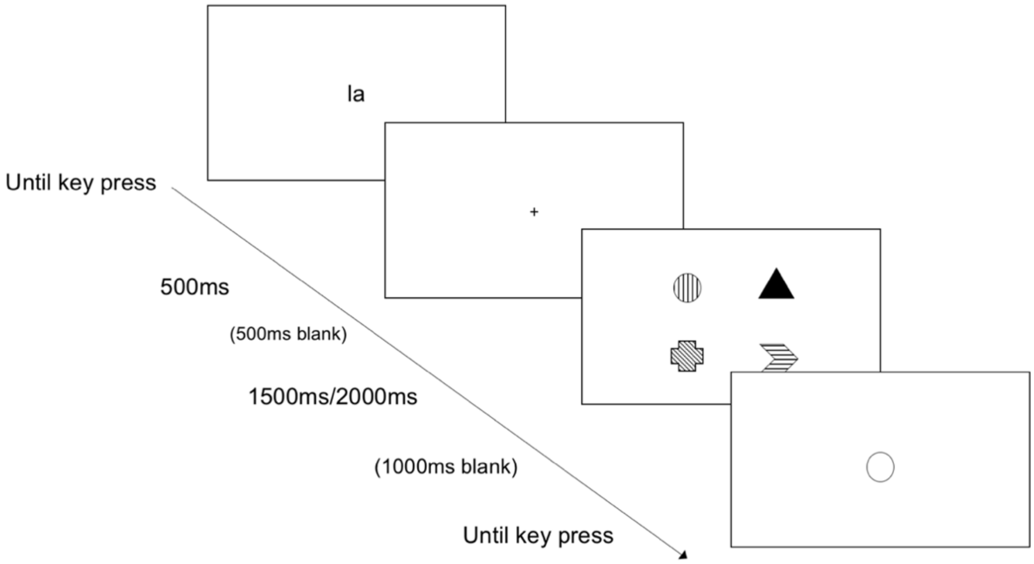


Figure 4.6. The experimental paradigm used in Experiment 7, with a 4-item trial as an illustrative example. The different types of shading reflect different colours (e.g. red, green, blue). The array was presented for 1500ms in the 3-item blocks and 2000ms in the 4-item blocks. Figure not to scale.

Before the start of each session, children completed a brief paper-based activity to ensure they understood the meaning of 'top-left'. In each trial, they were presented with four pictures of related objects (e.g. fruit, furniture, stationary) arranged in a 2 x 2 grid and asked to point to the top-left picture (see for an example). If they responded correctly on three consecutive trials, they immediately progressed onto the main experimental task. If they responded incorrectly on any of the three trials, the experimenter pointed to the top-left picture of three novel sets. The participant was then presented with three new sets, and asked to identify to the top-left picture in each. This was repeated a maximum of three times until children correctly selected the top-left picture on three consecutive trials. The same set of images was used in both sessions. 73% of younger children and 94% of older children responded correctly on the first attempt in session 1, whilst all children in both groups answered correctly on the first attempt in session 2. No child required more than two attempts in either session.

4.3.1.3. Data analysis

Frequentist and BF analysis was conducted. To mirror the analysis conducted in Experiment 6, separate analysis was conducted for the 3-item and 4-item conditions and age group was included as a factor.

4.3.2. Results

4.3.2.1. Three items

Mean accuracy (and SE) is displayed in Figure 4.6A as a function of probe value, SL, and age group. Means and SE are also presented in Table 2, collapsed across age groups. A 2 (Probe value: differential vs equal; within-subject) x 4 (SL: 1-4; within-subject) x 2 (Age group: Younger children vs Older children; between subject) mixed ANOVA was conducted. This revealed no significant main effect of probe value, $F(1, 64) = 0.27$, $MSE = .02$, $p = .604$, $\eta_p^2 < .01$; $BF_{10} = 0.10$, $BF_{01} = 10.00$), but a significant main effect of SL, $F(3, 192) = 7.30$, $MSE = .03$, $p < .001$, $\eta_p^2 = .10$; $BF_{10} = 1094.83$). Bonferroni-Holm post-hoc comparisons revealed significantly lower accuracy in the bottom-left position ($M = .68$, $SE = .03$), compared to the top-left position ($M = .78$, $SE = .02$; $p = .001$), the top-right position ($M = .76$, $SE = .02$; $p = .005$), and the bottom-right position ($M = .76$, $SE = .02$; $p = .005$). A significant effect of age group also emerged, $F(1, 64) = 9.32$, $MSE = .12$, $p = .003$, $\eta_p^2 = .13$; $BF_{10} = 10.21$, with older children ($M = .79$, $SE = .02$) exhibiting higher accuracy than the younger children ($M = .70$, $SE = .02$). There was no significant interaction between probe value and SL, $F(3, 192) = 0.99$, $MSE = .02$, $p = 0.400$, $\eta_p^2 = .02$; $BF_{10} = 0.05$, $BF_{01} = 20.00$). No other significant interactions emerged ($F \leq 1.51$, $p \geq .214$, $BF_{10} \leq 0.19$, $BF_{01} \geq 5.26$). The BF analysis indicated that the most likely model included main effects of SL and age group ($BF_{10} > 10{,}000$ relative to the null model with random effects of participant only).

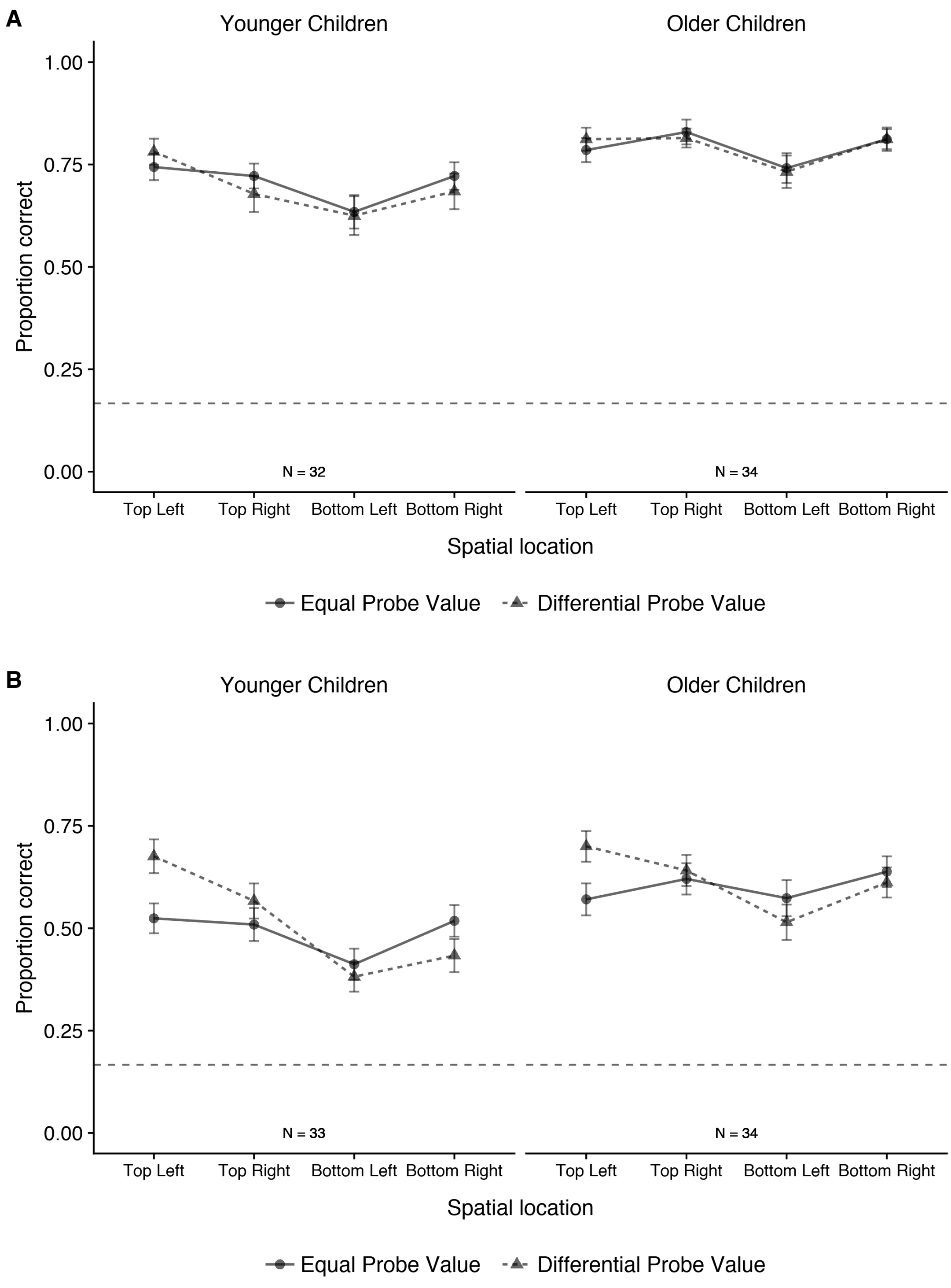


Figure 4.7. Mean proportion correct (and SE) in Experiment 7 as a function of probe value, spatial location and age group in the 3-item (A) and 4-item (B) conditions.

4.3.2.2. Four items

Mean accuracy (and SE) is displayed in Figure 4.6B as a function of probe value, SL and age group. Means and SE are also presented in Table 4.2, collapsed across age groups. A 2 (Probe value: differential vs equal; within-subject) x 4 (SL: 1-4; within-subject) x 2 (Age group: Younger children vs Older children; between-subject) mixed ANOVA revealed no significant effect of probe value $F(1, 65) = 1.66$, $MSE = .03$, $p = .202$, $\eta_p^2 = .03$; $BF_{10} = 0.23$, $BF_{01} = 4.35$). There was, however, a significant main effect of SL, *GG corrected* $F(2.61, 169.44) = 10.70$, $MSE = .06$, $p < .001$, $\eta_p^2 = .14$; $BF_{10} > 10{,}000$), with Bonferroni-Holm post-hoc comparisons revealing significantly lower accuracy at the bottom-left position ($M = .47$, $SE = .03$), relative to the top-left ($M = .62$, $SE = .02$; $p = .001$), top-right ($M = .58$, $SE = .02$; $p = .005$), and bottom-right positions ($M = .55$, $SE = .02$; $p = .020$). There was also an effect of age group $F(1, 65) = 9.26$, $MSE = .16$, $p = .003$, $\eta_p^2 = .13$; $BF_{10} = 9.94$), with older children ($M = .61$, $SE = .02$) performing better than the younger children ($M = .50$, $SE = .03$). Crucially, a significant interaction between probe value and SL emerged $F(3, 195) = 11.61$, $MSE = .03$, $p < .001$, $\eta_p^2 = .15$; $BF_{10} = 244.09$). No other interactions emerged ($F \leq 1.96$, $p \geq .131$, $BF_{10} \leq 0.59$, $BF_{01} \geq 1.69$). These findings were corroborated by BF factor analysis which revealed the most likely model contained main effects of SL and age group, and an interaction between probe value and SL ($BF_{10} > 10{,}000$ relative to the null model with random effects of participant only).

To investigate the key interaction between probe value and SL, Bonferroni-Holm corrected paired sample t-tests were conducted. There was a significant effect of probe value at the top-left position ($t(66) = 5.30$, $p < .001$; $d = 0.65$; $BF_{10} > 10{,}000$), with participants exhibiting higher accuracy in the differential probe value condition (*M* = .69, *SE* = .03) relative to the equal probe value condition (*M* = .55 *SE* = .03). No significant differences emerged at the other SLs ($t \geq -2.11$ and ≤ 1.26, $p \geq .117$, $d \geq -0.26$ and ≤ 0.15, $BF_{10} \leq 1.07$).

Table 4.2.

Mean accuracy (and SE) in Experiment 7 as a function of probe value, memory load and SL, collapsed across age group. N = 66 for the 3-item conditions and N = 67 for the 4-item conditions.

		TL	TR	BL	BR
3 items	Differential probe value	.80 (.02)	.75 (.03)	.68 (.03)	.75 (.03)
	Equal probe value	.77 (.02)	.78 (.02)	.69 (.03)	.77 (.02)
4 items	Differential probe value	.69 (.03)	.60 (.03)	.45 (.03)	.52 (.03)
	Equal probe value	.55 (.03)	.57 (.03)	.49 (.03)	.58 (.03)

TL = top-left, TR = top-right, BL = bottom-left, BR = bottom-right

4.3.2.3. Across memory loads

Accuracy at the top-left position is presented in Figure 4.4B as a function of probe value and memory load. As in Experiment 6, mean accuracy is presented for each participant and across participants. A 2 (Probe value: differential vs equal; within-subject) x 2 (Memory load: 3-item vs 4-item; within-subject) x 2 (Age group: Younger children vs Older children; between-subject) mixed ANOVA was conducted at the top-left position to explore whether probe value boosts vary across memory loads. One participant in the younger children's group completed only the 4-item condition and was therefore excluded from this analysis. The analysis was therefore run on data from 32 younger children and 34 older children. A significant effect of probe value emerged $F(1, 64) = 21.22, MSE = .02, p < .001, \eta_p^2 = .25$; $BF_{10} = 3900.50$), with accuracy in the differential probe value condition ($M = .74, SE = .02$) higher than the equal probe value condition ($M = .66, SE = .02$). There was also a significant effect of memory load $F(1, 64) = 76.19, MSE = .02, p < .001, \eta_p^2 = .54$; $BF_{10} > 10{,}000$) with higher accuracy in the 3-item conditions ($M = .78, SE = .02$) relative to the 4-item conditions ($M = .62, SE = .02$). No significant effect of age group emerged $F(1, 64) = 0.83, MSE = .10, p = .365, \eta_p^2 = .01$; $BF_{10} = 0.38, BF_{01} = 2.63$). There was a significant interaction between probe value and memory load $F(1, 64) = 9.82, MSE = .02, p = .003, \eta_p^2 = .13$; $BF_{10} = 10.25$), but no other interactions ($F \leq .213, p \geq .646, BF_{10} \leq 0.25, BF_{01} \geq 4.00$). The BF analysis yielded similar outcomes, indicating that the most likely model included main effects of probe value and memory load, as well as an interaction between probe value and memory load ($BF_{10} > 10{,}000$ relative to the null model with random effects of participant only).

To investigate the interaction between probe value and memory load, paired sample t-tests were conducted (corrected using Bonferroni-Holm). A significant difference between the probe value conditions emerged in the 4-item conditions, $t(65) = 5.24, p < .001, d = 0.65$; $BF_{10} = 8673.41$) with participants exhibiting higher accuracy in the differential probe value condition relative to the equal probe value condition. There was, however, no effect of probe value in the 3-item conditions, $t(65) = 1.33, p = .189$; $d = 0.16, BF_{10} = 0.31$; $BF_{01} = 3.23$).

4.3.3. Discussion

Experiment 7 examined whether children can direct their attention to valuable item information during simultaneous presentation, and whether such effects vary as a function of memory load. It was found that children do prioritise more valuable information, but only under certain memory load conditions. Probe value effects were observed when four items were presented, with accuracy at the top-left location significantly higher when that location was associated with more points (the differential probe value condition) than when all locations were associated with the same number of points (the equal probe value condition). However, when arrays contained three items, no significant probe value effects emerged, with BF analysis providing evidence of no effect. Such findings are in line with Shimi & Scerif (2017), who found significant cueing effects when four items were presented, but not when arrays contained two items.

In both the 3- and the 4-item analyses, there was a significant effect of age group, with the older children showing higher accuracy overall relative to the younger children. This age effect did not, however, emerge in the analysis across SLs. There were also no significant interactions containing age.

Finally, an unexpected finding emerged with children's accuracy, with children generally less accurate at the bottom-left item in the array. This was observed in both the 3- and 4-item conditions. These findings are in line with previous research suggesting that accuracy on visual WM tasks may vary by spatial location (Della Sala, Darling, & Logie, 2010). However, the precise pattern reported in the current study is not consistent with that reported by Della Sala et al. (2010), who found that adults remember items on the left better. One possibility is that the spatial locations that are remembered most accurately might differ depending on age.

It is unclear why such effects would occur, however, warranting further research to investigate this in both developmental and adult populations.

4.4. General discussion

The experiments reported within this chapter examined whether children aged 7-10 years can prioritise more valuable information in WM if the reward system underpinning the task is child-friendly and engaging. Effects were examined using sequential (Experiment 6) and simultaneous (Experiment 7) modes of presentation. Significant probe value effects were observed in both experiments, demonstrating that children can direct their attention to more valuable information provided that they are sufficiently motivated to do so. Further supporting this conclusion, cross-experimental analysis found a significant effect of probe value in Experiment 6, but no effect in Berry et al. (2018; Experiment 2), which had fewer motivational aspects. Importantly, the probe value effects were observed without memory for other items dropping to floor, indicating that children did not simply abandon the other representations in order to retain the more important item. This demonstrates that children as young as 7-8-year olds can distribute their attention across items in a sophisticated manner. Such outcomes are consistent with previous findings which have shown that children are able to direct their attention in WM based on cues (Astle et al., 2012; Shimi et al., 2014; Shimi & Scerif, 2017) and probe frequency (Cowan et al., 2010). However, these experiments are the first to demonstrate that children can prioritise more valuable information in WM, and that they are able to orient their attention when a sequential mode of presentation is used (Experiment 6).

Berry et al. (2018) suggested that 7-10-year-old children are not able to direct their attention to more valuable information in WM due to under-developed executive resources. However, the current outcomes demonstrate that children are able to prioritise valuable items when they are motivated to do so. This suggests that children are likely to assess the cognitive effort associated with strategies and use this to determine whether to apply them (Chevalier, 2018). Individuals might only apply cognitively demanding strategies (like prioritising more valuable items in cued-recall WM tasks; Hu et al., 2016) when they are motivated to perform optimally and believe that the reward for doing so is worth the cognitive effort involved.

The differences in findings between the current experiments and Berry et al. (2018) are in line with previous research which has suggested that children may show cognitive abilities earlier in development if the task and its context are engaging and age appropriate. They are also in line with Brewer et al. (2013), who suggested that children may need more motivation to complete experimental tasks to the best of their ability and that researchers should exert caution when extending paradigms used in adults to a developmental context. Interestingly, however, the analysis comparing the data from Experiment 6 and Berry et al. (2018; Experiment 7) revealed no main effect of experiment. This therefore indicates that the additional motivational elements did not enhance overall performance on the task. This is in line with previous findings, which have indicated that gamification does not significantly increase visual WM capacity in typically developing children (Dovis, Van der Oord, Wiers, & Prins, 2012). Taken together, this suggests that gamification elements encourage children to comply with task instructions, employ appropriate strategies (such as prioritisation), and complete experimental tasks properly, though it does not increase overall performance on WM tasks. Such findings may have broad implications for psychologists, educators, and other professionals who wish to understand how cognitive abilities develop during childhood.

A secondary aim of this set of studies was to investigate whether probe value effects vary as a function of memory load, an effect that has been reported in the cueing literature in both adults (Astle et al., 2012; Kuo et al., 2012; Nobre et al., 2008; Souza et al., 2014; van Moorselaar et al., 2015) and children (Shimi & Scerif, 2017). When information was encountered sequentially, there was no significant interaction between probe value and memory load (Experiment 6). However, when items were displayed simultaneously, significant effects of probe value were observed when four items were presented, but not when three items were presented (Experiment 7). Why might this pattern of findings have emerged? One possibility is that whether probe value effects vary as a function of memory load depends on the presentation mode used. If items are presented sequentially, children might direct their attention to the more valuable item regardless of how many items are presented. Conversely, if simultaneous arrays are used, children might not be able to prioritise a more valuable item if the set size is low. However, it is not clear why this pattern of results would emerge, as it is considered easier to direct attention in WM when information is presented simultaneously (Gorgoraptis et al., 2011). Indeed, the effect size for the probe value comparison was larger at the target item when information was presented simultaneously in the 4-item condition (*Sequential (Experiment 6) d* = 0.42, *Simultaneous (Experiment 7) d* = 0.65).

An alternative, and more likely possibility, is that children selectively direct their attention in WM when the amount of information presented is at, or above, capacity limits (Shimi & Scerif, 2017). As shown in Figure 4.4, mean accuracy was highest across both experiments when three items were presented simultaneously. As such, participants may have decided that additional strategies were not necessary to maximise performance in this condition (Shimi & Scerif, 2017), or not worth the cognitive effort they would involve (Chevalier, 2018). This may be particularly true, as children tend to overestimate their memory abilities, and may therefore have believed they were performing even better than they were (Shin, Bjorklund, & Beck, 2007; Yussen and Levy, 1975), In contrast, when four items were presented simultaneously (Experiment 7), or three or four items were presented sequentially (Experiment 6), children may have found the task more difficult and therefore prioritised the more valuable item in order to maximise their point score.

Although evidence that probe value effects increased with memory load (Experiment 7) is consistent with cueing studies (Astle et al., 2012; Kuo et al., 2012; Nobre et al., 2008; Shimi & Scerif, 2017; Souza et al., 2014; van Moorselaar et al., 2015), it somewhat contrasts with outcomes from the probe frequency literature (Cowan et al., 2010). In Cowan et al. (2010), it was found that the youngest children (aged 7-8 years) were able to direct their attention as effectively as adults when four items were presented, but that this ability was reduced when arrays contained six items. There are several possible explanations for these differing findings. Firstly, in the current experiment, participants were presented with either three or four items, whereas in Cowan et al. (2010), arrays contained four or six items. As 7-8-year-old children in Cowan et al. (2010) were able to direct their attention as effectively as adults when four items were presented, it may be that children have difficulty directing attention to particular items in WM when memory load is further increased (i.e. to six items). Alternatively, the differences in findings might reflect the number of items participants were asked to direct their attention towards. Within the current experiment and the cueing literature (Shimi et al., 2014; Shimi & Scerif, 2017), attention was orientated to one particular item. In Cowan et al. (2010), participants were presented with arrays containing equal numbers of circles and triangles, and told that one of the shapes was more likely to be assessed. Participants therefore needed to direct their attention towards multiple items within the array (i.e. two items in the four-load condition and three items in the six-load condition). One possibility is that younger children struggle to direct their attention towards multiple items simultaneously. However, if so, it is likely that children would have shown some impairments in the load-4 condition where they needed to orient their attention towards two items. This effect was not observed, with the children exhibiting similar attentional effects to adults at this memory load. Relatedly, it is possible that the way in which participants approached the task differed between experiments. As half of the array was more likely to be tested in Cowan et al. (2010), participants may have attempted to filter out the items that were unlikely to be assessed, rather than prioritising those that were. This contrasts with the current paradigm, in which participants are considered to prioritise the more valuable item, whilst retaining the less valuable information. These approaches might be differently affected by age and memory load, thus potentially explaining the differences in results. Finally, given that Chapter 2 revealed differences between probe value and probe frequency effects, it is possible that these manipulations differ in the extent to which they are affected by age and memory load. It would be beneficial for further research to investigate these potential explanations. This would provide further insights into the ways in which children are able to direct their attention in WM, and how the manipulations might differ from each other.

No age group interactions emerged in either experiment, demonstrating that the ability to prioritise valuable information in WM does not substantially increase between 7-10 years of age. It is, however, worth nothing that the effect sizes observed were considerably smaller than those reported in previous studies employing similar paradigms in adults (e.g. Chapter 2 and Hu et al., 2016). Further developmental changes must therefore occur after 10 years of age. As such, it would be beneficial for research to explore how the ability to prioritise more valuable information in WM develops across late childhood and adolescence. Given that probe value effects are thought to rely on executive control (Hu et al., 2016), increases in the size of probe value effects beyond 10 years of age might reflect the development and maturation of executive resources in late childhood and adolescence (Jurado & Rosselli, 2007; Waszak, Li, & Hommel, 2010). One might therefore expect larger probe value effects in adolescence than those observed in the current experiments. However, Castel, Humphreys, et al. (2011) found that children (*M*. age = 8.14) and adolescents (*M*. age = 14.52) showed similar levels of selectivity in a reward-based memory task, with both being less selective than adults. Based on this, one might expect adolescents to show smaller probe value boosts than adults, and effects that are of a similar magnitude to children. Evidence of this would demonstrate that the ability to prioritise more valuable information in WM is relatively stable throughout development, before maturating during late adolescence or early adulthood.

Relatedly, it would be useful for research to investigate whether probe value effects are observed in younger children. It has been suggested that children begin to use more proactive control strategies at approximately 6-8 years of age (Blackwell & Munakata, 2014; Chatham, Frank, & Munakata, 2009; Chevalier et al., 2014; Elke &

Wiebe, 2017). Before this, they are thought to rely on more reactive strategies, responding to events only when they occur and retrieving information only when it is required (Chatham et al., 2009; Chevalier et al., 2014). It is therefore unclear whether younger children (e.g. 5-year olds) would be able to prioritise more valuable information, or indeed direct their attention in response to any information that identifies a particular item as being more important or goal-relevant (e.g. visual cues or probe frequency; Chevalier et al., 2014).

A potential limitation of these experiments is that children were presented with feedback that did not accurately reflect their true performance on the task. This was implemented to ensure that all children received the same feedback and that this did not affect their behaviour in subsequent trials. For instance, if a child was told multiple times that they were performing poorly, they may give up and simply guess. Similarly, if a child was told that they were answering correctly on all of the trials, they may decide that it is not worth trying to prioritise the more valuable item. Either of these responses would have introduced additional error variance, making it difficult to assess whether children can prioritise valuable information. However, as feedback would provide an external sensory indicator of the reward (Hammer et al., 2015), further research is warranted to investigate how probe value effects are affected by feedback.

Moreover, a potential confound when comparing Experiment 6 to Berry et al. (2018) is that the participants differed in SES. In Berry et al. (2018), the majority of children were from low SES areas, whilst participants in the current experiments were recruited from a school in a moderate SES area. Some research has identified a link between SES, WM, and executive functions, with children from higher SES backgrounds exhibiting better WM and executive functions (e.g. Arán-Filippetti, 2013; Hackman, Gallop, Evans, & Farah, 2015; Noble, McCandliss, & Farah, 2007; Noble, Norman, & Farah, 2005). Disparities in SES between studies might therefore at least partially explain the differences in findings between the current experiments and Berry et al. (2018). However, if this was the case, one would expect higher overall performance in Experiment 6 relative to Berry et al. (2018), a finding that was not observed in the cross-experimental analysis. It is therefore unlikely that differences in SES can fully account for the differences between findings, though it would be useful for research to explore whether personal characteristics such as SES affect the ability to direct attention in WM.

4.5. Conclusions

The experiments presented within this chapter are the first to demonstrate that children can prioritise more valuable information in WM when they are sufficiently motivated to do so. When information was presented sequentially, this ability was robust, with effects observed across memory loads and age groups. When simultaneous arrays were used, probe value effects were observed when four items presented, indicating that the children's ability to prioritise information is not limited to tasks using sequential displays. However, when three items were presented, no probe value effects emerged. This is likely to reflect participants deciding not to apply this strategy because they were already performing well on the task. Taken together, this suggests that children as young as seven years old can selectively apply encoding strategies when such approaches are likely to enhance performance on WM tasks.

CHAPTER 5

INVESTIGATING THE DURABILITY OF PROBE VALUE BOOSTS

5.1. Introduction

The previous empirical chapters demonstrated that adults can direct their attention to more valuable visual (Chapter 2) and verbal information in WM (Chapter 3). Following this, Chapter 4 showed that children can also prioritise particularly valuable information in WM when they are sufficiently motivated to do so. The current chapter developed this work further by examining the durability of this effect in adults. More specifically, the study investigated whether prioritising a valuable item for a WM task yields a boost that can be observed at LTM. Also of interest was whether the emergence of an effect at LTM depends on whether the more valuable item was tested at WM. A secondary aim of the research was to examine whether probe value affects response time (RT) at both the WM and the LTM phases. This was not examined in the previous chapters as all experiments thus far asked participants to respond verbally (as opposed to a mouse click or key press).

As discussed in Chapter 1, several paradigms have been developed to examine the effects of value on episodic memory, including value-directed remembering (e.g. Castel et al., 2002; Robison & Unsworth, 2017) and monetary incentives (e.g. Adcock et al., 2006; Gruber & Otten, 2010; Shigemune et al., 2010). From these studies, it has been concluded that participants remember items associated with higher values better than those associated with lower values. This has been found across various time delays and using a variety of material, including visual (e.g. Adcock et al., 2006; Gruber & Otten, 2010; Shigemune et al., 2010) and verbal stimuli (e.g. Castel et al., 2002; Robison & Unsworth, 2017).

Further research has built on these findings by examining how directing participants to, or away from, particular items at WM affects longer-term retention. Festini and Reuter-Lorenz (2013; Experiment 2) examined how attempting to forget particular information during a WM task affects performance on a surprise LTM test. Participants were presented with a set of three words on either side of the screen. A cue was then displayed that informed participants which set of words they should forget. After a short delay, a test word was displayed in the centre of the screen and participants had to identify whether this formed part of the to-be-remembered set. Where the probe was *not* one of the to-be-remembered words, it could either be unrelated to any of the encoded items, semantically related to the items in the remember list, semantically related to the items in the forget list, or one of the items presented in the forget list. Of particular interest was the number of false alarms in each of these conditions. Following this, participants completed a surprise LTM test in which they had to identify whether a series of words had been studied during the encoding phase. At WM, the number of false alarms was significantly higher for the lures that were semantically similar to the remember list relative to the lures that were related to the forget list. This provides evidence that individuals could forget the list they were told to, or at least prioritise the words they were instructed to remember. At LTM, accuracy was higher for the items participants were encouraged to remember during the WM phase relative to the items they were told to forget. This study therefore demonstrates that participants can selectively focus on particular items at WM, and that doing so affects longer-term retention.

Cueing has also been used to examine how directing attention at WM affects LTM (LaRocque et al., 2015; Reaves, Strunk, Phillips, Verhaeghen, & Duarte, 2016; Strunk, Morgan, Reaves, Verhaeghen, & Duarte, 2018). Reaves et al. (2016) presented participants with three coloured images of everyday objects on either side of the screen. Before item onset, a pre-cue was presented, indicating which side of the screen the tested item would appear on. After the items had been displayed, a 100% valid retro-cue appeared on some trials, which informed participants of the item that would be tested at retrieval. In other trials, a neutral cue was presented which provided no information to participants. After a short delay, an item was presented in the centre of the screen and participants had to indicate whether it had been presented during encoding. After a delay of approximately 25 minutes, participants completed a LTM test, where all of the items were tested again. Retro-cueing improved performance in the WM and the LTM tasks, although the boost was smaller in the latter test. The boost at LTM was observed even when the analysis was restricted to trials where participants responded correctly at WM. This suggests that the boost at LTM was not purely driven by participants being more likely to respond correctly to the cued items at WM. RTs were also significantly faster for the cued items relative to the neutral items at WM. No significant effect was observed at LTM, however, suggesting that the boost to RT was short-lived.

Strunk et al. (2018) replicated and extended these findings by examining the short and long-term effects of cueing in young and older adults. Both groups exhibited higher accuracy in the retro-cue trials relative to the neutral trials at WM and LTM. In addition, both groups responded faster to the cued items at WM relative to the neutral-cue trials, although no RT effect was observed at LTM. The older adults exhibited poorer performance and slower responses overall, but the magnitude of the accuracy and RT boosts did not differ across groups. Taken together, these studies suggest that cueing a particular item at WM enhances longer-term retention of such representations. It also appears to quicken RTs, but only at WM.

Although these studies have investigated how directing attention at WM affects LTM, research to date has not examined the durability of probe value effects. As discussed previously within this thesis, there is growing evidence to suggest that the probe value effect differs from other attentional manipulations, such as cueing and probe frequency (Chapter 2; Hitch et al., 2018). It would therefore be beneficial for research to specifically investigate whether prioritising an item for a WM test enhances longer-term retention. Such findings would have important implications, as it would help to delineate whether items held in a privileged state within WM (such as the FoA) are retained more effectively at longer delays than items that are not associated with a privileged state.

Furthermore, relatively little research has examined how prioritising a more valuable item affects RTs. This was examined by Sandry et al. (2014), who presented participants with letters and then assessed memory after a short delay using a 2-Alternative forced choice (AFC) test. During encoding, one of the letters was displayed in red, indicating that it was relatively more valuable to the rest. At retrieval, participants were faster to respond to the more valuable item (relative to the same SP when all items were equally valuable) and the final item. However, as Chapter 3 highlighted some potential differences between the mechanisms involved in prioritising verbal and visuo-spatial information, it would be useful to investigate whether such effects emerge when visual stimuli are used. It would also be beneficial to investigate whether prioritising a more valuable item for a WM test affects RTs at LTM.

The current study examined these research questions. Participants were presented with series of four images of everyday objects sequentially. After a brief delay, participants completed a 4-AFC test, in which they had to identify the item that had been presented during the encoding phase. This retrieval method was implemented (as opposed to cued recall; Allen & Ueno, 2018; Berry et al., 2018; Hitch et al., 2018; Hu et al., 2014; 2016) to ensure that performance was above chance on the surprise LTM test. Everyday objects were used in order to create a large pool of stimuli and to ensure that items were not repeated across trials. However, this also provided an opportunity to examine whether probe value effects emerge when stimuli more analogous to everyday objects are used. As in the previous experiments (Chapters 2-4), participants were either told that the first item was more valuable than the rest (differential probe value), or that all of the items were equally valuable (equal probe value). The manipulation was targeted at SP1 in order to closely match Chapters 2 and 4. As in these chapters, the main comparison of interest was between SP1 in the differential and equal probe value conditions. Any primacy effect observed is therefore unlikely to act as a confound as this should be similar across conditions. Approximately ten minutes after the end of the WM task, participants completed a surprise LTM test. Half of the items tested during this phase had been assessed at WM, whilst half had not been tested. This allowed us to explore whether the emergence of probe value effects at LTM differs depending on whether the item was tested at WM. RTs were also measured at both the WM and the LTM phases. Following the experiment, participants completed two short questionnaires. The first asked participants whether they predicted the LTM test, whilst the second assessed awareness about probe value effects at WM and LTM.

Based on previous published work and the consistent findings presented in this thesis thus far (i.e. Chapters 2-4; Allen & Ueno, 2018; Hitch et al., 2018; Hu et al., 2014; 2016), it was predicted that participants would exhibit higher accuracy in the differential probe value condition relative to the equal probe value condition at SP1. Such findings would extend previous research by demonstrating that probe value effects emerge when everyday objects are used and memory is assessed using recognition. At LTM, it was predicted that a probe value boost would occur regardless of whether the item was tested at WM. This was based on evidence that increasing the value of an item yields LTM boosts in the absence of an intervening WM test (e.g. Adcock et al., 2006; Gruber & Otten, 2010), and studies which have shown that items that are retro-cued and tested at WM are remembered better at LTM (LaRocque et al., 2015; Reaves et al., 2016; Strunk et al., 2018). It was, however, predicted that the effect would be larger when the item was tested at WM, as assessing the item at this phase may reinforce the probe value manipulation. Furthermore, based on findings from Sandry et al. (2014) and the cueing literature (Reaves et al., 2016; Strunk et al., 2018), it was predicted that participants would respond faster in the differential probe value condition at SP1, relative to the equal probe value condition. However, it was predicted that RTs would not differ as a function of probe value at LTM (Reaves et al., 2016; Strunk et al., 2018). Finally, in line with Chapter 3, it was anticipated that participants' awareness would broadly reflect the pattern of behavioural results observed.

5.2. Experiment 8

5.2.1. Method

5.2.1.1. Participants

Thirty participants completed the experiment (*M. age* = 20.27, *SD* = 2.42; Range = 18.32-29.42; 3 males). Participants were fluent English speakers, had normal or corrected-to-normal vision, and had no known learning difficulties. They were reimbursed with course credit or cash. The study was approved by the School of Psychology Ethics Committee at the University of Leeds (Ethics reference number 17-0017).

5.2.1.2. Design, materials, and procedure

The study comprised two main parts: a WM phase and a LTM phase. At the WM phase, a 2 (Probe value: differential vs equal) x 4 (SP: 1-4) within-subject design was employed. At LTM, a 2 (Probe value: differential vs equal) x 4 (SP: 1-4) x 2 (Tested at WM: tested vs not tested) within-subjects design was employed. At both the WM and LTM test phases, the dependent variables were accuracy (proportion correct) and RT. The experiment was completed as one session, taking approximately 60 minutes. The structure of the session is displayed in Figure 5.1.

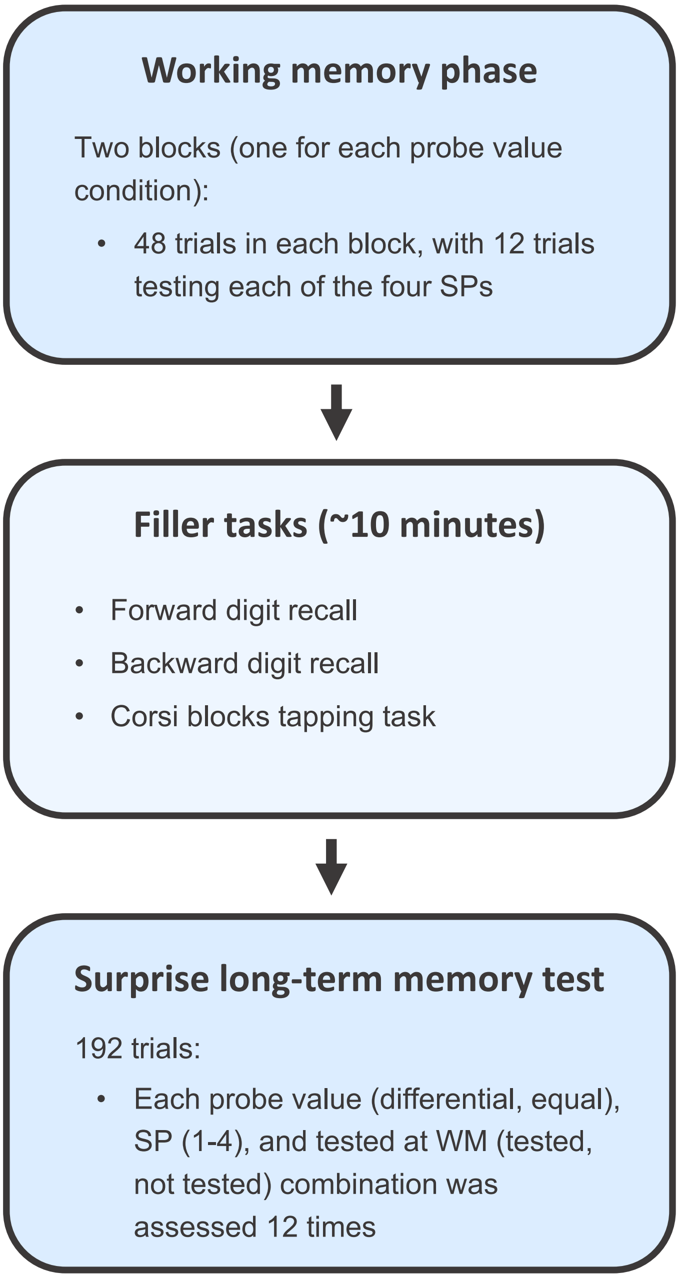


Figure 5.1. The structure of the experiment. Participants first completed the WM phase Following this, participants completed a series of filler tasks, which took approximately 10 minutes. These comprised tests commonly used to assess WM. Finally, participants completed a surprise LTM test of the items presented during the WM phase. Half of the items had been tested during the WM phase, whilst half had not.

Participants first completed a WM phase. This comprised two blocks (one for each of the probe value conditions), with each containing 48 trials. The order of these blocks was fully counterbalanced across participants. Within the blocks, each of the four SPs was assessed 12 times. The order of the SP trials within each block was randomised such that participants could not predict which item would be tested.

The paradigm used in the WM phase is displayed in Figure 5.2A. Participants were first presented with a blank screen for 1000ms, followed by a randomly generated number between 20-99 for 1000ms. Participants repeated this number aloud until the retrieval phase in order to disrupt verbal rehearsal (Baddeley, 1986). A fixation cross was then displayed for 1000ms, followed by a blank screen for 500ms. Next, participants were presented with four images of everyday objects for 500ms, separated by an ISI of 250ms. These images were taken from two large datasets: the Bank of Standardized Stimuli (BOSS; version 2; Brodeur, Dionne-Dostie, Montreuil, & Lepage, 2010; Brodeur, Guerard, & Bouras, 2014) and from Brady, Konkle, Alvarez, & Oliva (2008). The images in the BOSS database were photographed by the authors (Brodeur et al., 2010; 2014), whilst the images in Brady et al's (2008) set was collated using commercially available image sets (Hemera Photo-Objects, Volumes I and II) and Google Image searches. Where the same object appeared in both databases, one of these was removed to ensure that each object presented was distinct. The images were presented in greyscale as pilot work revealed at-ceiling performance when coloured images were used. The images presented were selected pseudo-randomly for each participant, with the constraint that each image could only be presented once during the entire experiment. Each image appeared at one of eight equally spaced locations positioned around an imaginary circle of radius 5.66°, located at the centre of the screen. The positions used in each trial were selected pseudo-randomly, with the constraint that no position could be used more than once within a trial. The images measured approximately 4°, based on a viewing distance of 50cm. Following item presentation, there was a retention interval of 1000ms. After this, participants were presented with one item from the encoding phase and three lures, which had not been displayed in the experiment thus far. The items were presented at corners of an imaginary 8° wide square, located at the centre of the screen. Participants had to click on the item that had been presented during the encoding phase using a computer mouse. The images remained on screen until the participant responded. Participants were told that accuracy was more important, but that they should respond as quickly as possible.

Before the encoding phase, participants were told the probe value instructions. In the differential probe value condition, they were informed that they would get 4 points if they were asked about the first item and they responded correctly. If they were asked about any other item and they responded correctly, they would get 1 point. In the equal probe value condition, all of the items were worth the same number of points (1 point). At the start of each probe value block, participants completed two practise trials to familiarise themselves with the task. Reminders of the probe value manipulation were presented after 12 trials.

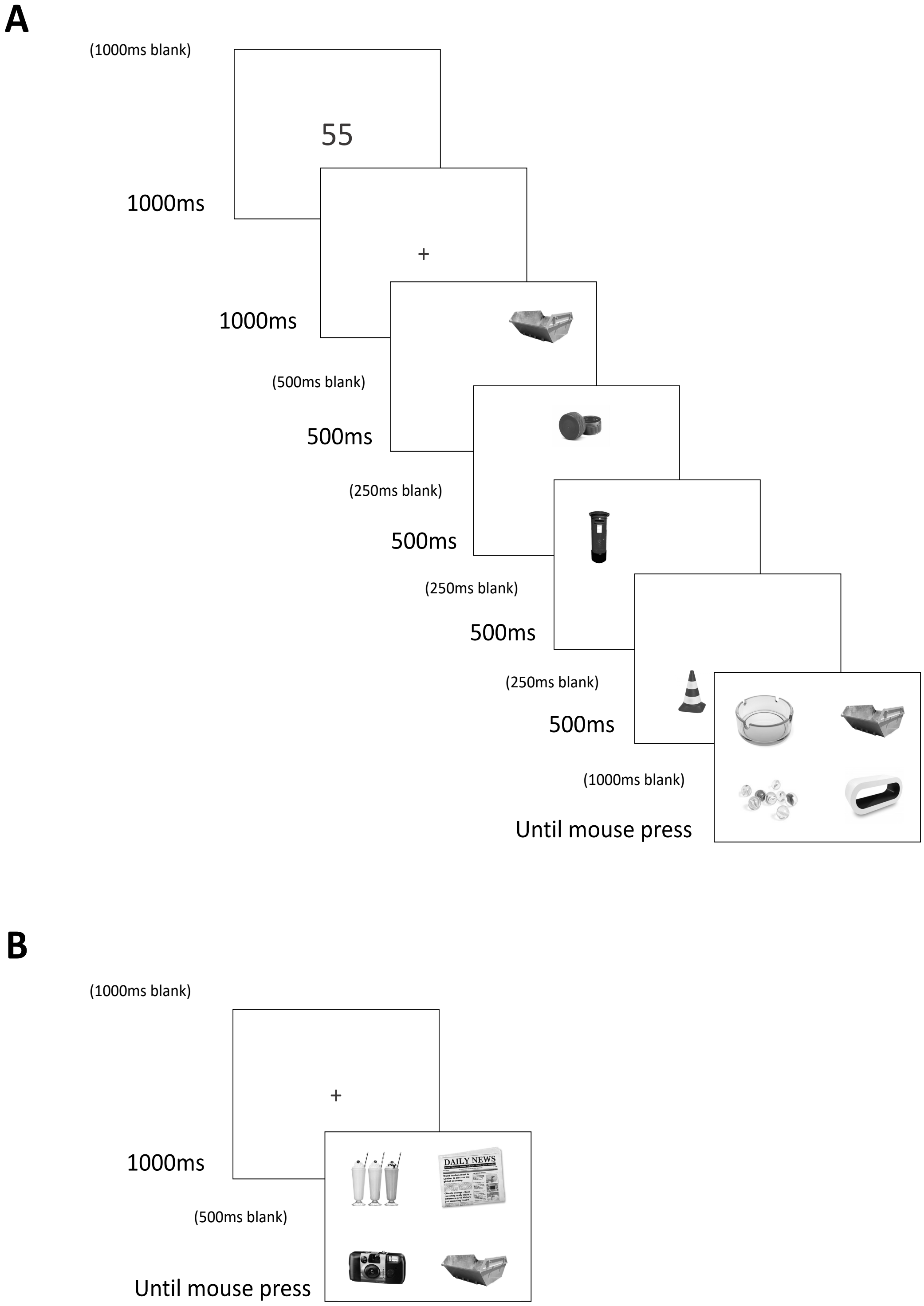


Figure 5.2. The schematic used in the working memory (A) and long-term memory (B) trials. Figure not to scale.

Following the WM phase, participants completed a series of filler tasks. Three WM tests were used in order to reduce the likelihood that participants would anticipate the LTM test: FDR, BDR, and the Corsi blocks tapping task. In the FDR task, participants were read out series of numbers by the experimenter, which they had to repeat back in the same order. Each sequence contained series of digits from 0 to 9. Participants first completed two practise trials, which contained two digits per sequence (e.g. 3-5). Following this, participants were presented with sequences that slowly increased in length from two digits at the start of the task to nine digits at the end of the task. Three trials were given at each length. If participants successfully recalled two or three of the sequences at a given length, they progressed onto the next length. If they answered 0 or 1 correct, the task ended. The BDR task used similar sequences, with participants asked to recall the digits in a backwards order (e.g. 3-5-1 should be recalled as 1-5-3). In the Corsi blocks tapping task, participants were presented with a Lego board containing nine squares arranged in a random order. These squares contained numbers on them in order to make it easier to administer the task, but these were only visible to the experimenter. The experimenter tapped a pattern, which the participant was asked to reproduce in the same order. As with the digit recall tasks, participants first completed two practise trials. The sequence length gradually increased from two to nine, although the task ended if participants responded correctly on zero or one trials at a given length. The order of these tasks was fully counterbalanced across participants. The tasks took approximately 10 minutes.

After the series of filler tasks, participants completed a surprise LTM test where memory for the items presented during the WM task was assessed. A delay of 10 minutes was selected so that the experiment could be completed in a single session. This is also broadly in line with some previous studies investigating the effects of cueing (Reaves et al., 2016; Strunk et al., 2018) and value (e.g. Gruber & Otten, 2010), where LTM tests have been completed after delays of 15-25 minutes. This comprised 192 trials. A schematic displaying the procedure and timings for this phase is displayed in Figure 5.2B. Within each trial, a blank screen was presented for 1000ms, followed by a fixation cross for 1000ms, and then a blank screen for a further 500ms. After this, participants were presented with four images of everyday objects. The items were presented at corners of an imaginary 8° wide square, located at the centre of the screen. One of the images had been presented during the encoding phase of the WM task and three were new lures that had not been presented at any other point during the experiment. Participants had to select which item had been presented during the encoding phase using a computer mouse. In the "tested at WM" trials, the target item had already been the target item during the WM phase. In the "not tested at WM" trials, the item tested had not been assessed at WM. As with the WM phase, participants were told that accuracy was more important, but that they should try to respond as quickly as possible. Each probe value (differential, equal), SP (1-4), and tested at WM (tested, not tested) combination was assessed 12 times. The order of these trials was randomised to ensure that order of trials in the WM and LTM phases were unrelated. Participants were asked to take a short break after every 40 trials to ensure they were able to concentrate fully on the task.

After the LTM phase, participants completed a short questionnaire (see Appendix E). This asked them whether they predicted the LTM task (yes/no). They were also asked the extent to which they thought about the images between the WM and LTM tests. This was measured on a 7-point Likert scale, where 1 reflected 'not at all' and 7 reflected 'all of the time'. Participants were also asked whether they believed prioritisation helped or harmed their memory for the more valuable item and the less valuable items in both the WM and LTM phases. This was similar to the questionnaires employed in Chapter 3, with participants asked to give a response on a 9-point Likert scale (where 1 = large negative effect, 5 = no perceived effect, and 9 = large positive effect). As in Chapter 3, these responses were recoded by subtracting five from each value, such that -4 reflects a large negative effect, 0 indicates no effect, and 4 reflects a large positive effect.

5.2.1.3. Data analysis

The experiment primarily aimed to assess how prioritising an item for a WM test would affect performance in a surprise LTM task. As such, participants who anticipated the second memory task were excluded from all analysis. This is in line with some previous research which has removed participants who anticipated the surprise memory test (Murayama & Kuhbandner, 2011). This resulted in 6 out of 30 participants (20%) being excluded from the analysis. The analysis was therefore conducted on the data for 24 participants. Generally, these participants did not report thinking about the objects much in the interval between the WM and LTM tests (M = 2.04, SE = 0.34, where 1 = not at all and 7 = all of the time). Across all analysis, proportion correct was used as the outcome measure as opposed to d' in order to aid comparisons with the previous chapters (Chapters 2-4). Proportion correct was calculated by dividing the number of trials responded correctly by the number of trials responded correctly plus the number of trials responded incorrectly.

RTs were measured from the onset of the test stimuli until the participants responded using the mouse in ms. RTs for incorrect responses were discarded. This removed 385/2304 trials in the WM phase (16.71%) and 2088/4608 in the LTM phase (45.31%). Visual inspection of the individual data revealed that the RT for one trial (differential probe value, SP2) in the LTM phase was over 100,000ms (100 seconds). This trial was removed to ensure that the SD of this participant's condition was not artificially inflated, which would have affected the trimming procedure. Finally, RT trimming was conducted. RTs that fell 2.5 SDs above or below the mean for each condition for each participant were excluded. This trimming procedure removed 20/1919 data points in the WM phase (1.04%) and 7/2521 data points in the LTM phase (0.28%).

5.2.2. Results

5.2.2.1. Accuracy (proportion correct)

5.2.2.1.1. Working memory

Proportion correct in the WM task is displayed in Figure 5.3A as a function of probe value and SP. The means for individual participants at SP1 is presented in Figure 5.4A, as a function of probe value. A 2 (Probe value: differential vs equal) x 4 (SP) within-subjects ANOVA was conducted. No significant effect of probe value emerged ($F(1, 23) = 1.28$, $MSE = .02$, $p = .270$, $\eta_p^2 = .05$; $BF_{10} = 0.29$; $BF_{01} = 3.45$), although there was a main effect of SP ($F(3, 69) = 7.02$, $MSE = .02$, $p < .001$, $\eta_p^2 = .23$; $BF_{10} = 140.44$). Bonferroni-Holm post-hoc comparisons revealed significant differences between SP1 ($M = .88$, $SE = .02$) and SP2 ($M = .80$, $SE = .02$; $p = .010$), SP1 and SP3 (M = .79, SE = .03; p = .006), SP2 and SP4 (M = .88, SE = .02; p = .036), and SP3 and SP4 (p = .036). A significant interaction between probe value and SP also emerged ($F(3, 69)$ = 5.85, MSE = .02, p = .001, η_p^2 = .20; BF_{10} = 43.49. BF analysis revealed that the best model included a main effect of SP, as well as an interaction between probe value and SP (BF_{10} = 3145.70 relative to a null model containing participant only).

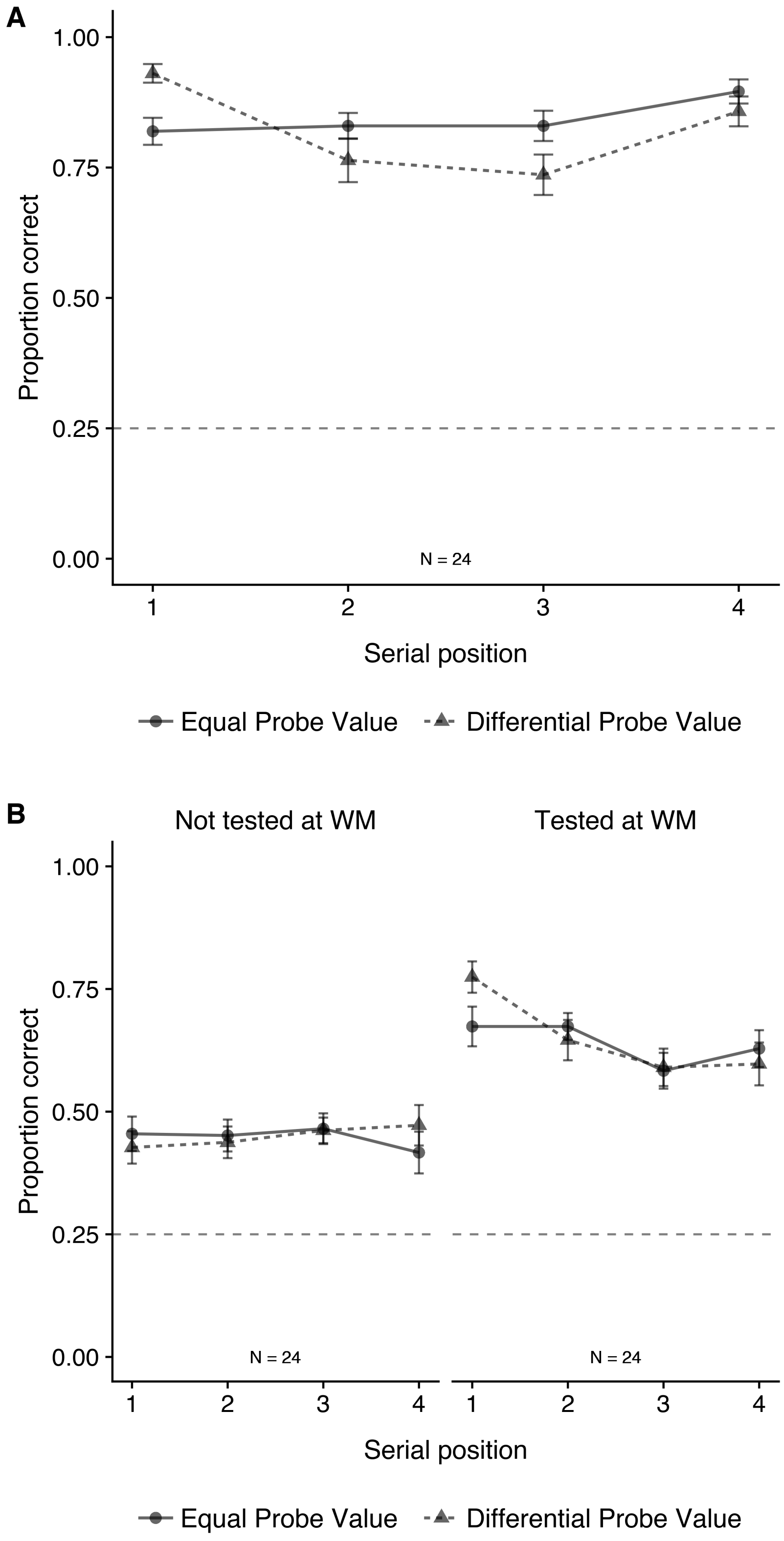


Figure 5.3. Mean proportion correct. Part A displays accuracy in the WM phase as a function of probe value and SP, whilst part B displays accuracy in the LTM phase as

a function of probe value, SP and tested at WM. Error bars denote SE. The horizontal dotted line at 0.25 reflects chance rate.

To investigate the interaction further, a series of paired sample t-tests were conducted to investigate if an effect of probe value was observed at each SP. Bonferroni-Holm correction was applied to correct for multiple comparisons. There was a significant effect at SP1 ($t(23) = 3.52$, $p = .012$, $d = 0.72$; $BF_{10} = 20.51$), with participants recognising more items correctly in the differential probe value condition (M = .93, SE = .02) relative to the equal probe value condition (M = .82, SE = .03). There were no significant effects at the other SPs ($t \geq -2.21$ and ≤ -1.35, $p \geq .185$, $d \geq -0.45$ and ≤ -0.28; $BF_{10} \leq 1.65$).

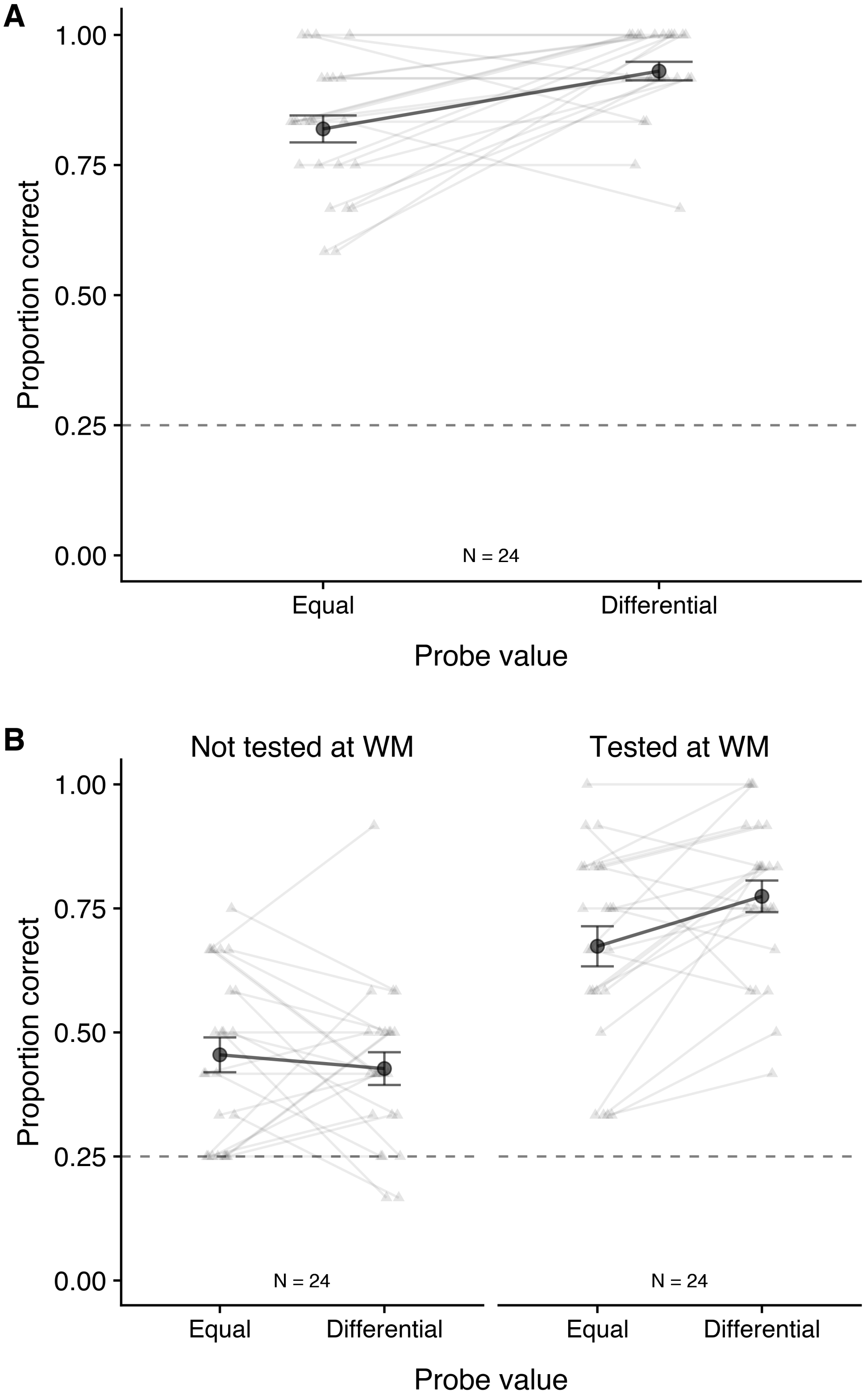


Figure 5.4. Mean accuracy (proportion correct) at SP1 for individual participants. Figure 5.4A displays performance in the WM phase, as a function of probe value. Figure 5.4B displays performance in the. LTM phase as a function of probe value

and tested at WM. The lighter lines with triangular points reflect the mean performance of individual participants, whilst the darker line with circular points reflects mean accuracy across participants. Error bars denote SE. The dotted grey line at 0.25 reflects chance performance.

5.2.2.1.2. Long-term memory

Proportion correct as a function of probe value, SP, and tested at WM is displayed in Figure 5.3B. Means for individual participants at SP1 is presented in Figure 5.4B, as a function of probe value and tested at WM. A 2 (probe value: differential vs equal) x 4 (SP) x 2 (tested at WM: tested vs not tested) within-subjects ANOVA was conducted. This revealed no main effect of probe value ($F(1, 23) = .22$, $MSE = .02$, $p = .646$, $\eta_p^2 < .01$; $BF_{10} = 0.13$, $BF_{01} = 7.69$). There was also no significant main effect of SP, although this approached significance ($F(3, 69) = 2.49$, $MSE = .03$, $p = .068$, $\eta_p^2 = .10$; $BF_{10} = 0.83$, $BF_{01} = 1.20$). There was, however, a main effect of tested at WM ($F(1, 23) = 122.25$, $MSE = .03$, $p < .001$, $\eta_p^2 = .84$; $BF_{10} > 10{,}000$), with participants exhibiting higher accuracy for items that were tested ($M = .65$, $SE = .03$) relative to items that were not ($M = .45$, $SE = .02$). There was no interaction between probe value and SP (*GG corrected* $F(2.06, 47.25) = .70$, $MSE = .03$, $p = .504$, $\eta_p^2 = .03$; $BF_{10} = 0.06$, $BF_{01} = 16.67$), or probe value and tested at WM ($F(1, 23) = .23$, $MSE = .01$, $p = .635$, $\eta_p^2 = .01$; $BF_{10} = 0.16$, $BF_{01} = 6.25$). However, a significant interaction emerged between SP and tested at WM ($F(3, 69) = 5.66$, $MSE = .02$, $p = .002$, $\eta_p^2 = .20$; $BF_{10} = 19.30$). There was also a significant three-way interaction between probe value, SP, and tested at WM ($F(3, 69) = 3.22$, $MSE = .02$, $p = .028$, $\eta_p^2 = .12$), although the BF analysis was slightly in favour of no effect ($BF_{10} = 0.72$,

BF_{01} = 1.39). The BF analysis indicated that the most likely model included a main effect of tested at WM, as well as an interaction between SP and tested at WM (BF_{10} > 10,000 relative to null model containing participant only).

In order to explore the significant three-way interaction between probe value, SP, and tested at WM, two separate 2 (Probe value: differential vs equal) x 4 (SP: 1-4) within-subjects ANOVAs were conducted on the data from the tested at WM condition and the not tested at WM condition. When the items were not tested at WM, there were no main effects and no interaction ($F \leq .82$, $p \geq .487$, $BF_{10} \leq 0.16$). When the items were tested at WM, there was no main effect of probe value ($F(1, 23) = .50$, $MSE = .01$, $p = .488$, $\eta_p^2 = .02$; $BF_{10} = 0.19$, $BF_{01} = 5.26$), although there was a main effect of SP ($F(3, 69) = 8.19$, $MSE = .02$, $p < .001$, $\eta_p^2 = .26$; BF_{10} = 3115.39). Bonferroni-Holm pairwise comparisons revealed that SP1 ($M = .72$, SE = .03) significantly differed from SP3 ($M = .59$, $SE = .03$; $p < .001$) and SP4 ($M = .61$, $SE = .04$; $p = .005$). A significant difference also emerged between SP2 ($M = .66$, SE = .03) and SP3 ($p = .032$), whilst the difference between SP1 and SP2 approached significance ($p = .057$). Additionally, there was a marginally significant interaction between probe value and SP (*GG corrected* $F(2.32, 53.29) = 3.13$, $MSE = .02$, p = .045, $\eta_p^2 = .12$, although the BF was entirely equivocal ($BF_{10} = 1.00$). To investigate this interaction, Bonferroni-Holm corrected paired sample t-tests were conducted to investigate whether a probe value effect emerged at each SP. A significant effect emerged at SP1 ($t(23) = 2.84$, $p = .036$, $d = 0.58$; $BF_{10} = 5.15$), with participants exhibiting higher accuracy in the differential probe value condition ($M = .77$, SE = .03) than the equal probe value condition ($M = .67$, $SE = .04$). No significant effects emerged at the other SPs ($t \geq -1.16$ and $\leq .17$, $p \geq .774$, $d = \geq -0.24$ and ≤ 0.03; BF_{10} ≤ 0.39).

5.2.2.1.3. Trials correct at working memory

As shown in the previous section, a significant probe value effect was observed at LTM, but only when the item was tested at WM. This analysis included all of the data, regardless of whether participants responded correctly or not at WM. The LTM boost in the tested at WM condition might therefore have not resulted from a durable probe value boost *per se* (Reaves et al., 2016). Instead it may have been driven by a carry-over effect from WM, with participants recognising more of the items at SP1 in the differential probe value condition relative to the equal probe value condition at this phase. To investigate this, the LTM data was re-analysed including only trials that were tested at WM and that participants responded correctly on (Reaves et al., 2016). This was conducted to investigate whether items initially associated with a higher value were still more likely to be recognised at LTM when performance at WM was controlled for.

Only the condition where items were tested at WM were included in this analysis. Within this condition, three hundred and eighty-five out of 2304 trials were removed (16.71%) due to participants responding inaccurately at WM. Proportion correct at LTM for trials that were tested at WM and participants responded correctly is displayed in Figure 5.5, as a function of probe value and SP. A 2 (probe value: differential vs equal) x 4 (SP) repeated measures ANOVA revealed no significant main effect of probe value ($F(1, 23) = .02$, $MSE = .02$, $p = .882$, $\eta_p^2 < .01$; $BF_{10} = 0.16$, $BF_{01} = 6.25$), although there was a main effect of SP ($F(3, 69) = 7.33$, $MSE = .03$, $p < .001$, $\eta_p^2 = .24$; $BF_{10} = 858.68$). Pairwise comparisons (corrected using Bonferroni-Holm) revealed that SP1 ($M = .76$, $SE = .03$) significantly differed from

SP3 (M = .62, SE = .03; p = .002) and SP4 (M = .64, SE = .04; p = .005). The difference between SP2 (M = .71, SE = .03) and SP3 also approached significance (p = .064). Importantly, the interaction between probe value and SP $F(3, 69) = 1.63$, MSE = .02, p = .192, η_p^2 = .01) was not significant, with the BF analysis also providing some evidence of no effect (BF_{10} = 0.25, BF_{01} = 4.00). The BF analysis revealed that the best model included a main effect of SP (BF_{10} = 858.68 relative to a null model containing participant only). From this, it can be concluded that there was no significant probe value boost at LTM when only the trials participants answered correctly at WM were considered. This suggests that the probe value boost at LTM in the tested-at-WM condition may simply reflect a carry-over effect from participants being more likely to respond correctly at WM.

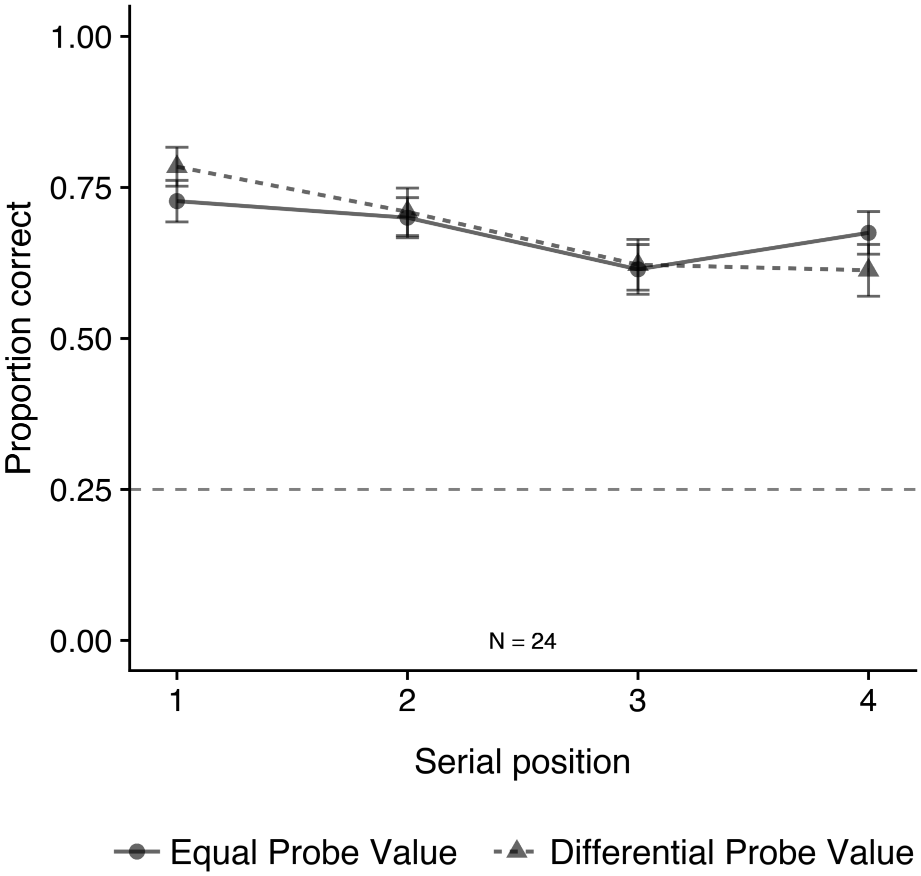


Figure 5.5. Proportion correct at LTM for trials participants were tested on at WM and responded correctly as a function of probe value and SP. Error bars denote SE. The dashed grey line at 0.25 displays chance rate.

5.2.2.2. Response times

5.2.2.2.1. Working memory

Mean RT in the WM phase (and SE) is displayed in Figure 5.6A as a function of probe value and SP. Figure 5.7A displays the mean RT for individual subjects at SP1. A 2 (Probe value: differential vs equal) x 4 (SP) within-subjects ANOVA revealed no main effect of probe value ($F(1, 23) < .01$, $MSE = 157{,}734.57$, $p = .993$,

$\eta_p^2 < .01$; $BF_{10} = 0.05$, $BF_{01} = 20.00$). A main effect of SP emerged ($F(3, 69) = 9.46$, $MSE = 76{,}733.88$, $p < .001$, $\eta_p^2 = .29$; $BF_{10} > 10{,}000$). Bonferroni-Holm corrected pairwise comparisons revealed a significant difference between SP1 ($M = 1864.46$, $SE = 80.94$) and SP2 ($M = 2049.29$, $SE = 92.62$; $p = .005$), SP2 and SP3 ($M = 1916.34$, $SE = 88.48$; $p = .048$), SP2 and SP4 ($M = 1752.90$, $SE = 79.46$; $p < .001$), and SP3 and SP4 ($M = .59$, $SE = .03$; $p = .008$). There was an interaction between probe value and SP ($F(3, 69) = 11.33$, $MSE = 84074.13$, $p < .001$, $\eta_p^2 = .33$; $BF_{10} > 10{,}000$). The BF analysis revealed that the best model included a main effect of SP, and probe value and SP interaction ($BF_{10} > 10{,}000$ relative to a null model containing participant only).

To investigate the probe value and SP interaction, Bonferroni-Holm corrected paired sample t-tests were conducted to examine whether a significant effect of probe value emerged at each SP. There was a significant effect at SP1 ($t(23) = -3.90$, $p = .004$, $d = -0.80$; $BF_{10} > 10{,}000$), with participants responding faster in the differential value condition ($M = 1658.98$, $SE = 58.31$) relative to the equal probe value condition ($M = 2069.94$, $SE = 123.47$). The difference also approached significance at SP4 ($t(23) = 2.30$, $p = .093$, $d = 0.47$; $BF_{10} = 8.94$), with participants exhibiting better performance in the equal probe value condition than the differential probe value condition. No significant difference emerged at the other SPs ($t \leq 1.90$, $p \geq .142$, $d \leq 0.39$; $BF_{10} \leq 0.74$; $BF_{01} \geq 1.35$).

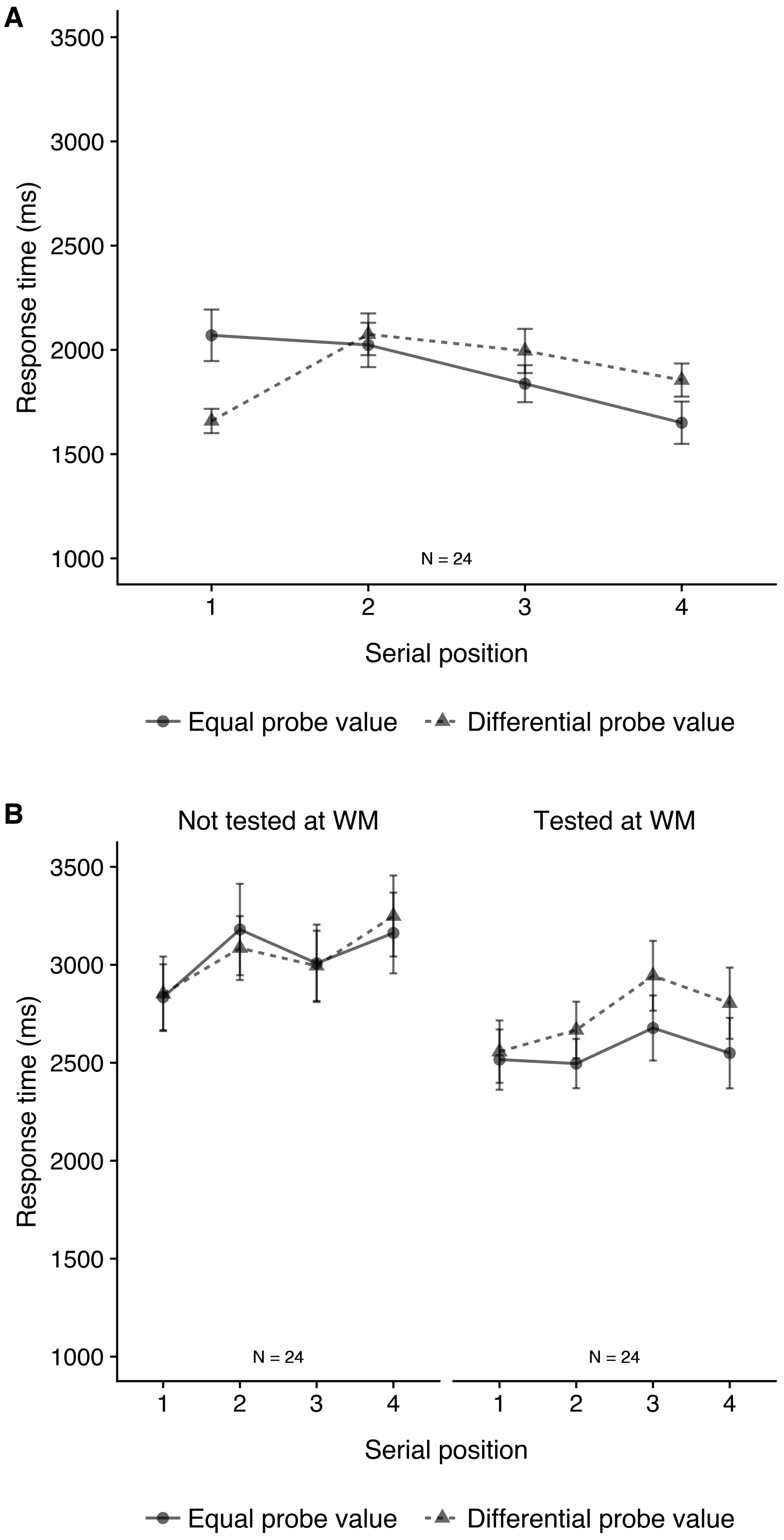


Figure 5.6. Trimmed RTs. Mean RTs for the WM task is displayed in part A as a function of probe value and SP. Mean RTs for the LTM task is presented in part B as a function of probe value, SP, and tested at WM. Error bars denote SE.

5.2.2.2.2. Long-term memory

Mean RT (and SE) in the LTM is displayed in Figure 5.6B, as a function of probe value, SP, and tested at WM. Mean RT at the SP targeted (SP1) for individual participants is displayed in Figure 5.7B. A 2 (Probe value: differential vs equal) x 4 (SP: 1-4) x 2 (Tested at WM: tested vs not tested) repeated-measures ANOVA revealed no main effect of probe value ($F(1, 23) = 0.68$, $MSE = 630{,}552.25$, $p = .419$, $\eta_p^2 = .03$; $BF_{10} = 0.27$, $BF_{01} = 3.70$) or SP ($F(3, 69) = 2.16$, $MSE = 298{,}793.14$, $p = .101$, $\eta_p^2 = .09$; $BF_{10} = 0.03$, $BF_{01} = 33.33$). A main effect of tested at WM emerged, however, ($F(1, 23) = 42.38$, $MSE = 309{,}338.61$, $p < .001$, $\eta_p^2 = .65$; $BF_{10} > 10{,}000$), whereby participants exhibited faster performance in the tested condition ($M = 2651.30$, $SE = 136.32$) relative to the not tested condition ($M = 3020.83$, $SE = 156.87$). There was a significant interaction between probe value and SP ($F(3, 69) = 3.09$, $MSE = 233{,}980.55$, $p = .033$, $\eta_p^2 = .12$), although the BF analysis provided evidence of no effect ($BF_{10} = 0.11$, $BF_{01} = 9.09$). The frequentist analysis also revealed a significant interaction between probe value and tested at WM ($F(1, 23) = 4.73$, $MSE = 276{,}021.56$, $p = .040$, $\eta_p^2 = .17$), although the BF was in favour of no effect ($BF_{10} = 0.21$, $BF_{01} = 4.76$). No other significant interactions emerged ($F \leq 1.96$, $p \geq .128$, $BF_{10} \leq 0.54$, $BF_{01} \geq 1.85$).

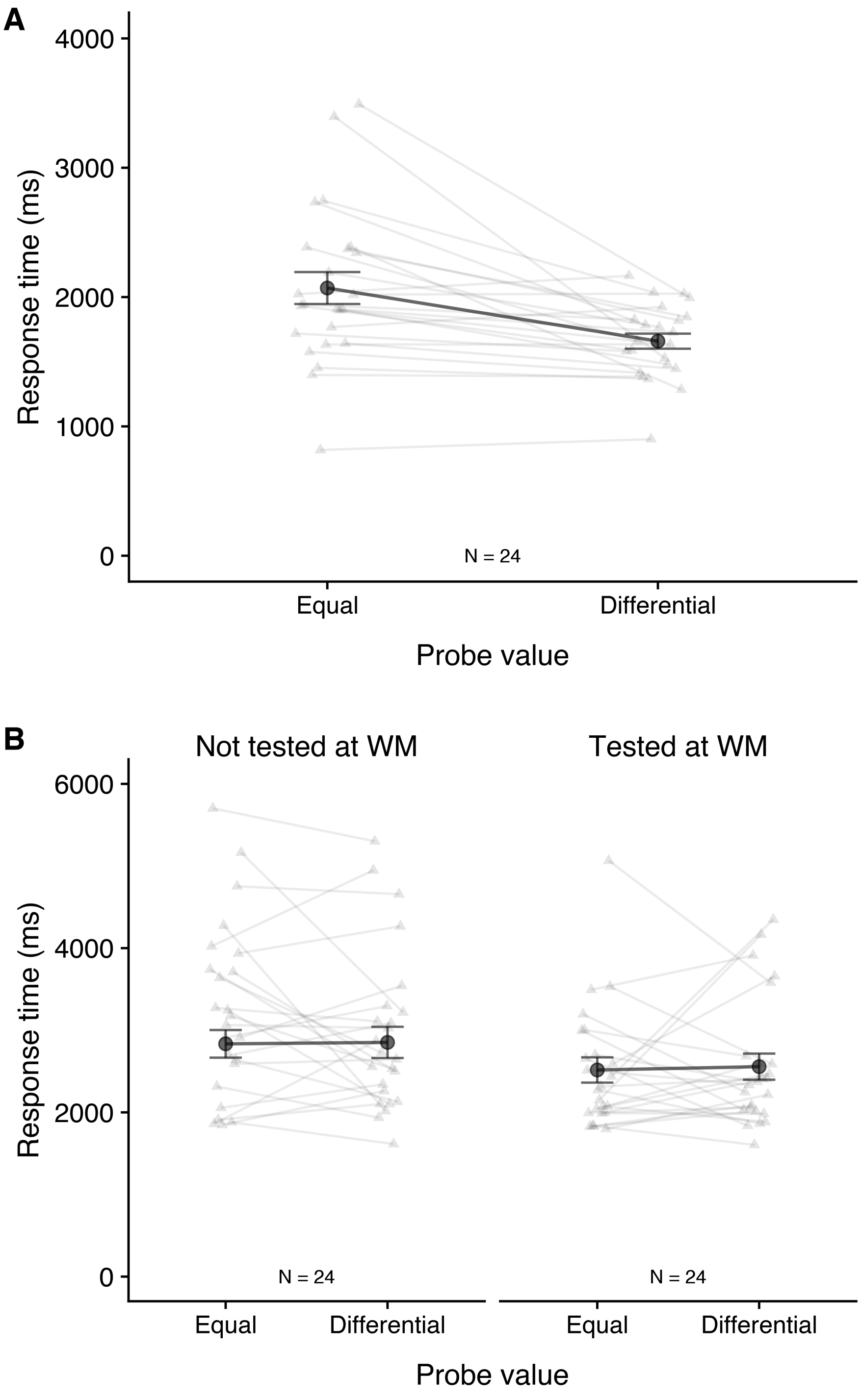


Figure 5.7. Mean RT at SP1 for the individual participants in the WM phase (A) and the LTM phase (B). Part A displays the means as a function of probe value, whilst part B presents the means as a function of probe value and tested at WM. The lighter lines with triangular points reflect means for individual participants. The darker line with circular points reflects mean RTs. Error bars display SE. Note that parts A and B are on different scales.

To investigate the interaction between probe value and SP, four paired samples t-tests (corrected using Bonferroni-Holm) were conducted to investigate whether an effect of probe value emerged at each SP, when the tested at WM variable was ignored. After correction, no significant effects emerged ($t \geq -1.08$ and ≤ 2.17, $p \geq .164$, $d \geq -0.22$ and ≤ 0.44, although the BF analysis was slightly in favour of an effect for the comparison at SP4 ($BF_{10} = 1.53$). At the other SPs, the BF analysis indicated evidence of no effect ($BF_{10} \leq 0.40$; $BF_{01} = 2.50$).

To investigate the interaction between probe value and tested at WM, two Bonferroni-Holm corrected paired sample t-tests were conducted to investigate whether an effect of probe value emerged in the tested at WM and not tested at WM conditions, whilst ignoring the effect of SP. There was no effect of probe value when the item was not tested at WM ($t(23) = -0.46$, $p = .654$, $d = -0.09$, $BF_{10} = 0.24$, $BF_{01} = 4.17$). There was also no significant effect of probe value when the item was tested at WM, although the difference approached significance ($t(23) = 2.21$, $p = .074$, $d = 0.45$, $BF_{10} = 1.65$).

5.2.2.3. Excluding participants who performed at ceiling at baseline

As shown in Figure 5.7A, several participants performed at ceiling in the WM phase. Although at-ceiling performance in the differential probe value condition at SP1 may limit the size of the probe value effect, 100% accuracy in the equal probe value condition completely prevents participants from obtaining a boost. This is likely to have reduced the effect size at WM, and potentially affected whether a boost was observed at WM. To investigate this, subsidiary analysis was conducted after removing the four participants who exhibited 100% accuracy in the equal probe value condition at WM. This analysis can be considered similar to that in Experiment 3 and 4, whereby participants were excluded from the analysis if they were likely to exhibit ceiling performance based on their FDR score. It is, however, important to note that this analysis within the current experiment was entirely post-hoc, whereas the decision to remove participants based on their FDR performance in Experiments 3 and 4 was made at the design phase. All of the statistical tests reported above were repeated with the reduced sample size. There were a few small differences that might affect the conclusions drawn. Firstly, the probe value accuracy effect at SP1 in the WM phase was considerably larger in this analysis ($d = 1.02$, relative to $d = 0.72$ when the data from all 24 participants were analysed). Secondly, a significant RT cost was observed at SP4 in the WM phase, with participants exhibiting faster responses in the equal probe value condition relative to the differential probe value condition ($p = 001$). When RTs were considered at LTM with the smaller sample size, no interaction between probe value and SP emerged, as well as no interaction between probe value and tested at WM.

5.2.2.4. Questionnaire

As in Chapter 3, ratings were transformed, with 0 reflecting no perceived effect of probe value, positive scores indicating perceived benefits, and negative scores reflecting perceived costs. The mean perceived effects to SP1 and the less valuable items are presented in Table 5.1 for WM and LTM. Four one-sample t-tests were conducted to investigate whether the ratings given significantly differed from zero (no perceived effect). The p-values were corrected using Bonferroni-Holm.

Participants believed that prioritising SP1 would result in a memory boost to this item at both WM ($t(23) = 4.38, p < .001, d = 0.89, BF_{10} = 133.64$) and LTM ($t(23) = 2.91, p = .010, d = 0.59, BF_{10} = 5.89$). They also believed that prioritisation would come at a cost to the less valuable items at both the WM ($t(23) = -3.12, p = .010,\ d = -0.64, BF_{10} = 8.94$) and LTM phases ($t(23) = -3.55, p = .006,\ d = -0.72, BF_{10} = 21.92$). Taken together, this demonstrates perceived boosts to the more valuable item and costs to the less valuable items in both the WM and LTM phases.

Table 5.1.

Mean (and SE) perceived effects of probe value obtained from the questionnaire. Values could range from -4 to 4, with negative values indicating perceived costs, whilst positive values reflect perceived benefits. Zero reflects no perceived effect.

	Effect to the more valuable item (SP1)	Effect to the less valuable items
Working memory	1.42 (0.32)	-1.17 (0.37)
Long-term memory	1.13 (0.39)	-1.25 (0.35)

Next, two paired sample t-tests (corrected using Bonferroni-Holm) were conducted to investigate whether the perceived boosts at SP1 and the perceived costs differed between the WM and LTM phases. The perceived boost to the more valuable item did not significantly differ between the WM ($M = 1.42, SE = 0.32$) and LTM phases ($M = 1.13, SE = 0.39$; $t(23) = 0.74, p = .940,\ d = 0.15, BF_{10} = 0.27$, $BF_{01} = 3.70$). A similar pattern emerged concerning the perceived costs to the less valuable items, with no significant difference emerging between the WM ($M = -1.17$,

$SE = 0.37$) and LTM phases ($M = -1.25$, $SE = 0.35$; $t(23) = 0.35$, $p = .940$, $d = 0.07$, $BF_{10} = 0.23$, $BF_{01} = 4.35$).

5.3. Discussion

The current study investigated whether prioritising a more valuable item for a WM task resulted in a durable boost that could be observed at LTM. Also of interest was whether increasing probe value resulted in faster RTs at both test phases. In the WM phase, participants responded faster and more accurately at SP1 in the differential probe value condition relative to the condition in which all items were equally valuable. Increasing probe value also enhanced accuracy at SP1 in the LTM phase, but only if the item had been tested at WM. However, further analysis suggested that this effect is likely to have been driven by a carry-over effect from WM, rather than a boost at LTM *per se*. RTs also did not significantly differ as a function of probe value at LTM.

Within the study, recognition was implemented to ensure that participants did not perform at floor in the surprise LTM task. This differs from most previous studies exploring the effects of probe value on WM, which have primarily used cued-recall (e.g. Chapters 2 and 3, Allen & Ueno, 2018; Berry et al., 2018; Hitch et al., 2018; Hu et al., 2014; 2016). Evidence that probe value effects emerged at WM therefore extends previous research, providing further evidence that the effect is robust. This is also in line with other paradigms, such as cueing, where boosts have been found in tasks employing a variety of retrieval methods, including cued recall, recognition, and delayed estimation (Souza & Oberauer, 2016). However, the size of the probe value effect was somewhat smaller at SP1 than in previous research (i.e. d

= 0.72 when all of the participants were included and *d* = 1.02 when participants who performed at ceiling in the equal probe value condition were excluded, relative to *d* = 1.33 in Chapter 2). There may be several reasons for this. Firstly, even after excluding participants who performed at ceiling in the equal probe value condition, several participants exhibited 100% accuracy in the differential probe value condition (see Figure 4A). Although at-ceiling performance in the differential probe value condition would not have prevented participants from obtaining a boost, it is likely to have constrained its size. This might, therefore, explain why a somewhat smaller effect size was observed in this study relative to previous research (e.g. Chapter 2). The smaller effect size observed may also partially reflect participants using different strategies as a result of the retrieval method. Thiede (1996) found that participants expect recall tests to be more difficult and to perform better on recognition tests. If participants believe they will retrieve most or all of the items, focusing on the more valuable object may not be considered beneficial or potentially even harmful to overall performance (Middlebrooks et al., 2017). Evidence of this was provided by Middlebrooks et al. (2017), who found that the effect of value on LTM performance was smaller in participants who anticipated a recognition test relative to participants who expected to recall the information. Such findings are also in line with Chapter 3, in which it was suggested that the retrieval method used may have affected the strategies participants used to retain the items that differed in value. This therefore further highlights the need for research to investigate how methodological factors, such as the retrieval method, affect the size of probe value effects at WM.

At LTM, a probe value effect only emerged when the more valuable item had been assessed at WM. However, further analysis revealed that this effect was abolished when performance at WM was taken into account. This suggests that the probe value manipulation did not yield any LTM boost *per se*, with boosts instead occurring due to a carry-over effect from participants being more likely to respond correctly at WM. Evidence that a probe value boost was not observed at LTM *per se* contrasts with findings from the cueing literature, in which retro-cues have been found to enhance memory after delays of approximately 25 minutes (Reaves et al., 2016; Strunk et al., 2018). These differences in findings might reflect differences between the paradigms, with retro-cues resulting in lasting memory enhancements, whilst prioritising a more valuable item for a WM test does not. There are, however, several methodological distinctions between these studies which might explain the differences in findings.

Firstly, one possibility is that the differences in findings relates to the frequency with which the targeted item was tested at WM. In both Reaves et al. (2016) and Strunk et al. (2018), the cued item was tested at WM with 100% validity. In contrast, the probe value manipulation employed within the current study did not predict which item would be tested, with each of the four SPs tested 25% of the time. This might have resulted in a form of retrieval induced forgetting in the current study, whereby testing one item at WM disrupted later memory for non-tested items within the same trial (Murayama, Miyatsu, Buchli, & Storm, 2014). This might therefore explain why no probe value effect was observed at LTM when SP1 was not tested. However, if so, one would have expected a clear probe value effect to emerge at LTM when SP1 was tested at WM. This was observed, although it appears to be driven by boosts at WM rather than an independent effect at LTM. As such, differences in the frequency with which the more valuable/cued items were tested at

WM is unlikely to account for the differing outcomes between the current study and previous research using cueing (Reaves et al., 2016; Strunk et al., 2018).

Alternatively, the differences in findings might relate to the surprise nature of the LTM test in the current study. In this experiment, the LTM test was a surprise, whereas participants in previous retro-cue studies were aware of the LTM test, and even able to practice this phase prior to encoding (Reaves et al., 2016; Strunk et al., 2018). Participants might therefore have either consciously or unconsciously taken distinct approaches in the different experiments. When one does not expect to be tested further on an item, it might be beneficial to suppress or even remove the representation from memory, such that resources can be focused on the next trial (Ecker, Lewandowsky, & Oberauer, 2014; Lewis-Peacock, Kessler, & Oberauer, 2018; Souza et al., 2014). This strategy would not be useful if one expects to be tested on the item in the future, however, as it is likely to result in poor memory at the latter test. In the current experiment, participants may have taken this approach as they were unaware of the delayed test. This might therefore explain why an independent probe value boost was not observed at LTM. This could be examined in future work by exploring whether participants' awareness of the LTM test affects the size and nature of the probe value effect observed.

These findings also contrast with previous research which has found that items associated with higher monetary rewards are remembered better at LTM (e.g. Adcock et al., 2006; Gruber & Otten, 2010; Murty and Adcock, 2014; Shigemune et al., 2010; Wittmann et al., 2005). There may be several reasons for this. Firstly, most studies investigating the effects of reward on memory have employed a longer retention interval than that used in the current study, such as 24 hours (e.g. Adcock et al., 2006). This may be important, as previous research has suggested that reward activates the mesolimbic reward system, which, in turn, releases dopamine into the hippocampal memory system (Murayama & Kitagami, 2014; Shohamy & Adcock, 2010). Hippocampus-dependent memory consolidation may take some time, with sleep considered to be heavily implicated (Marshall & Born, 2007). As such, the delay between the encoding phase and the LTM test may not have been long enough in the current study, with a period of sleep potentially needed. Supporting this, Murayama and Kuhbander (2011) found that items associated with a higher reward were remembered more accurately than low value items after a one-week delay, but not after a delay of 10 minutes. Similarly, Braun, Wimmer, & Shohamy (2018) found that neutral items that appeared closer to a reward were remembered better after a 24-hour delay, but not a 15 minutes delay. There is, however, some evidence that items associated with a reward are remembered better after shorter delays. For instance, Gruber and Otten (2010) found that items associated with higher monetary rewards were remembered better after delays of 15 minutes. As such, this is unlikely to fully explain why a probe value effect did not emerge in the current experiment. Nevertheless, it would be useful to investigate how the delay between the WM test and the LTM test affects retention of the more valuable information. The methodology may need to be adapted in order to explore this though, as it is likely that the current paradigm would result in at-floor performance after longer delays, such as 24 hours or a week.

Alternatively, the differences in findings might relate to the reward offered. Previous studies which have reported that more valuable items are remembered better at LTM have used monetary rewards (e.g. Adcock et al., 2006; Gruber & Otten, 2010; Murty and Adcock, 2014; Shigemune et al., 2010; Wittmann et al., 2005). This contrasts with the current study, which used notional points. Although a notional points system yields robust probe value effects at WM in adults (Allen & Ueno, 2018; Hitch et al., 2018; Hu et al., 2014; 2016; Sandry et al., 2014; Chapter 2-3), it may not be motivating enough to generate independent effects at LTM. This therefore highlights a need for further research to explore whether the effects of value on memory differ depending on the reward offered, and whether the boost observed is larger when monetary rewards are offered as opposed to notional points. Some evidence for this has been provided in children, with the emergence of a probe value boost at WM contingent on the reward offered (Chapter 4).

Another possibility is that a probe value effect did not emerge at LTM *per se* because of the retrieval method used. As discussed previously, the probe value effect observed at WM was smaller than in previous work. This may have resulted from participants performing near ceiling or from participants not trying particularly hard to prioritise the high value item because the WM task was relatively easy (Middlebrooks et al., 2017; Thiede, 1996). If individuals were not trying particularly hard to prioritise the more valuable item for the WM test, the size of the effect at LTM may have been attenuated. Secondly, recognition tests present additional items at retrieval that may interfere with the memory traces of the to-be-remembered items (Criss, Malmberg, & Shiffrin, 2011). At WM, there is evidence that probe value effects are vulnerable to interference during the retention interval (Hitch et al., 2018; Hu et al., 2014). As such, it is possible that a probe value effect did not emerge at LTM because of the interference at the WM and/or the LTM test. There was, however, a significant probe value boost at WM, indicating that interference did not eliminate the effect at this stage. Nevertheless, the effects of interference may be larger at LTM, where the representations are likely to be weaker and not as readily accessible. The level of interference is also likely to have been larger during this phase as participants would have been exposed to more lure items by this point (i.e. lures presented at both the WM and LTM tests). Given these limitations, it would be useful for further research to investigate whether prioritising an item for a WM test affects LTM when other retrieval methods are implemented. Cued-recall has commonly been used to investigate the effects of probe value at WM (Allen & Ueno, 2018; Berry et al., 2018; Hitch et al., 2018; Hu et al., 2014; 2016; Chapters 2 and 4), although it was not implemented in the current study due to concerns of possible floor effects at LTM. However, further research could adapt this methodology, employing a cued-recall test at WM and a recognition test at LTM. This could be examined by overlaying the images on coloured square, with the WM test assessing the relationship between these features. At LTM, participants could then be asked to recognise the object in a 4-AFC test (as in the current experiment). This would have two benefits: (1) the WM phase would be similar to previous research (e.g. Chapter 2 & 4, Allen & Ueno, 2018; Berry et al., 2018; Hu et al., 2014; 2016; Hitch et al., 2018); and (2) the LTM phase would be identical to that used in the current experiment, meaning that sensitive performance levels are likely to be observed. Moreover, such a design might maximise the likelihood of finding an effect at LTM if one exists, as previous research has revealed that value effects are larger in participants who expect a recall test, relative to participants who expect to be tested using recognition (Middlebrooks et al., 2017). As participants would be encoding the items in anticipation of a cued-recall test at WM, this might result in the emergence of independent value effects at LTM. It is, however, important to note that some effects of interference may persist at LTM, as lures will be presented during this test (Criss et al., 2011). Nevertheless, the interference is likely to be smaller than in the current study, as lures would not be presented during the WM phase.

A final possibility is that increasing the value of an item does not yield performance boosts at LTM *per se*, but that it does alter the nature of the representations. For instance, items associated with higher rewards might be recollected better than items associated with lower rewards. Such outcomes would be in line with findings from Chan and McDermott (2007), who found that asking participants to recall information in an intermediate test did not affect later recognition memory. However, it did affect the way in which the decisions were executed at the final test, with participants reporting they were more likely to recollect the information that was tested earlier. In contrast, the familiarity judgements were unchanged. It would therefore be useful to investigate whether increasing the value of an item affects the extent to which individuals draw from recollection and familiarity in a recognition test, even if it does not affect overall performance levels. To do so, further research could ask participants to indicate whether to provide remember, know, and guess judgements on a trial-by-trial basis.

A secondary aim of the study was to investigate whether prioritising a more valuable item affects RTs as well as accuracy. A RT boost was observed at WM, with faster responses at SP1 in the differential probe value condition relative to the equal probe value condition. This difference was large (d = -0.80; Cohen, 1988), with participants responding, on average, 400ms faster in the differential probe value condition than the equal probe value condition. This indicates that the performance boost observed does not result from a speed-accuracy trade off, where participants respond more accurately in the differential probe value condition because they take longer to consider their answer. A SP effect also emerged, which was partially driven by faster RTs at SP4 relative to SP2 and SP3. Taken together, this provides further evidence to support claims that the more valuable item and the final item are held in a privileged state within WM (such as the FoA) that renders them easier to access (Hitch et al., 2018; Hu et al., 2014; 2016). Such findings are also in line with Sandry et al. (2014), who found that the most valuable item and the final item are responded to faster at WM, relative to the other items. In contrast to this, RTs did not differ as a function of probe value or SP at LTM, suggesting that the boosts observed are temporary and not observable beyond WM.

However, these findings should be interpreted with caution as the RT measure was not as 'pure' as others commonly used within cognitive psychology, such as a key press. Nevertheless, several measures were put in place to ensure that the RT measure was valid. Firstly, at the start of each trial, the mouse appeared at the centre of the screen, ensuring that each item was equidistant from the cursor. The images were also large, ensuring they were easy to select. Furthermore, the retrieval method was the same for all conditions, allowing comparisons between them. It would, however, be beneficial for further research to investigate the effects using a more fine-grained measure, such as reaction time obtained using a computer keyboard. Alternatively, it might be useful to collect RTs using a computer mouse in a further experiment, but to also track the trajectory of the movement. This would allow us to gain a wealth of information, including a measure of competition between the different options (i.e. if participants began to move to a particular item but switched to another before clicking; Maldonado, Dunbar, & Chemla, 2019; Zgonnikov, Aleni, Piiroinen, O'Hora, & di Bernardo, 2017).

The questionnaire revealed that participants perceived boosts to the more valuable item and costs to the other items at both the WM and the LTM phases. The perceived effects did not significantly differ between the phases. Evidence that participants perceived a probe value boost to the more valuable item at WM is in line with the behavioural outcomes observed. In contrast, although participants perceived a cost to the less valuable items, no significant effects emerged at these SPs. At LTM, participants also perceived a probe value boost. Whilst this effect did not emerge in the condition where items were not tested at WM, there was some evidence of a probe value effect when the items were tested at WM (although this was driven by a boost at WM and not an effect at LTM *per se*). Furthermore, although participants perceived costs to the less valuable items at LTM, no significant costs to the less valuable items emerged in either the tested at WM or the not tested at WM conditions. Replicating the findings from Chapter 3, these findings therefore indicate that participants have a fair, but not complete, awareness of the effects of probe value.

There are several other effects found within this chapter which warrant some discussion. In the analysis involving all of the participants, there was no significant cost to accuracy or RT at the less valuable SPs. This pattern was observed during both the WM and the LTM phases. When participants who performed at ceiling in the equal probe value condition were excluded, RTs in the WM task were significantly slower at SP4 in the differential probe value condition relative to the equal probe value condition. This is somewhat in line with the previous experiments, where inconsistent costs have been found using cued recall (Chapter 2) and serial recall (Chapter 3). Furthermore, although not of primary interest, a testing effect was observed, whereby items that were tested at WM were recognised more accurately and faster at LTM than items that were not assessed previously. This is in line with previous research, which has found that testing information enhances later memory (e.g. Kang, McDermott, & Roediger, 2007; Nungester & Duchastel, 1982; Rowland, 2014). However, it is important to note that the items that had been tested at WM had been viewed twice by participants (i.e. during encoding and during the WM test phase). In contrast, the items that had not been tested at WM had only been presented once (i.e. during encoding). The ‘testing effect’ observed might therefore result from a combination of the additional presentation, as well as a testing benefit.

As well as providing theoretical insights, the outcomes of this study might also have important practical implications. Firstly, although the effect appeared to be driven by a carry-over effect from WM, a probe value effect was observed at LTM when the item was tested at WM. This is promising, as it suggests that such a strategy may be beneficial in educational environments or other situations where an individual needs to remember information that differs in value over a prolonged period. However, at least in the current study, it appears that a better approach might be to assess memory for the item at WM, and in doing so, present it a second time. This resulted in a large performance enhancement and may therefore be a more useful way of boosting performance for particularly important information. Such findings are in line with previous work, which has suggested that testing information enhances retention more than other strategic approaches (Rickard & Pan, 2018). However, before these manipulations are implemented in a practical setting, further research should first be conducted to explore how prioritisation and testing at WM affects longer-term memory on tasks more analogous to everyday life.

5.4. Conclusions

This study reflects a first attempt at examining whether prioritising a more valuable item for a WM test enhances memory on a surprise LTM test. A significant effect of probe value was observed at LTM, but only if the item had been tested at WM.

Moreover, when performance at WM was controlled for, there was no longer a significant probe value effect. This suggests that prioritising an item for a WM test does not enhance LTM *per se*, but that the effect might emerge due to a carry-over effect from WM. These outcomes contrast with studies which have reported that cueing an item at WM enhances LTM for the item, as well as research which has examined the effects of monetary values on LTM (without an intervening WM test). As such, further research is needed to examine why these differences in findings might have emerged.

A secondary aim of the present study was to investigate whether probe value affects response times at both WM and LTM. A boost was observed at WM, with participants responding faster at SP1 when this item was associated with a higher value. Responses to the final item were also faster than the other less valuable items. This provides further support for previous claims that the more valuable item and the final item are held within a privileged state within WM (i.e. the FoA), which renders them more accessible. In contrast, neither the more valuable item nor the final item was recognised faster at LTM, suggesting that these items remain more accessible for a limited amount of time.

Taken together, the current study suggests that prioritising a more valuable item enhances both accuracy and RT at WM. This then, in turn, improves performance at LTM for these items. As such, it can be concluded that prioritising an item for a WM test does yields a durable boost if the item is tested at WM, but that this reflects a carry-over effect from WM rather than an effect at LTM *per se*.

CHAPTER 6

GENERAL DISCUSSION

6.1. Thesis overview

WM is vital to many everyday activities, including planning (Cowan, 2010), reasoning (Süß et al., 2002), learning (Alloway, 2006), and language comprehension (Daneman & Merikle, 1996). Recently, research has focused on examining how individuals can optimise the efficiency of WM to ensure that they are focusing their existing capacity on retaining particularly important or valuable information. This work has primarily employed cueing paradigms, whereby a visual stimulus points towards one of the items immediately before or after item presentation (e.g. Astle et al., 2012; Duarte et al., 2013; Gressman & Janczyk, 2016; Loaiza & Souza, 2018; Myers et al., 2018; Mok et al., 2016, Rerko et al., 2014; Shimi et al., 2014; Shimi & Scerif, 2017; Souza & Oberauer, 2016). This generally informs participants that this item will, or is likely to, be tested at retrieval (Souza & Oberauer, 2016). Other paradigms have been used less commonly, such as probe frequency, whereby participants are told that one feature (e.g. a colour or shape) is more likely to be tested at retrieval (Cowan et al., 2010; Gorgoraptis et al., 2011; Klyszejko et al., 2014), and probe value, whereby participants are informed that one particular item (e.g. a SP) is relatively more valuable than the other items (Allen & Ueno, 2018; Hitch et al., 2018; Hu et al. 2014; 2016; Sandry et al., 2014).

Although cueing paradigms have been focused on to date, this approach has little practical validity, as it would require the re-design of the environment in order to implement the visual cues. Moreover, both visual cues and probe frequency generally require one item to be tested more than the others, which would also place constraints on how it might be implemented in practical context. In contrast to this, the probe value paradigm simply assigns more points to one of the items, which can be done quickly and easily. The manipulation might also be considered most analogous to everyday life, as information encountered as part of memory tasks (e.g. WM, LTM) might naturally differs in its importance or value. Despite this, research examining the effects of probe value on WM has been limited to date.

The current thesis therefore aimed to extend this literature, by examining; (A) whether the probe value manipulation encourages individuals to direct their attention in a different way to probe frequency; (B) whether probe value effects emerge in a serial-recall auditory-verbal WM task, and whether these effects are reliant on rehearsal and executive control mechanisms; (C) whether children can prioritise particularly valuable information in WM across memory loads when they are sufficiently motivated to do so; and (D) whether prioritising a more valuable item for a WM test creates durable boosts that can be observed on a surprise LTM test. These research questions were explored across eight experiments and four empirical chapters. The outcomes and implications of each chapter will now be discussed in turn.

6.2. Summary of the key findings

6.2.1. Chapter 2

Terms such as “cueing” and “prioritisation” have been used interchangeably to refer to visual cueing, probe frequency, and probe value manipulations (Gorgoraptis et al., 2011; Klyszejko et al., 2014; Myers et al., 2018). However, within the literature, there is some existing evidence that probe value and cueing manipulations encourage individuals to direct their attention in different ways. For instance, probe value effects in cued-recall visual WM tasks are considered to be entirely strategic in nature (Hu et al., 2016), whilst cueing benefits may be obtained through a combination of strategic and automatic processes (e.g. Berryhill et al., 2012; Gunseli et al., 2015; Hommel et al., 2001; Shimi et al., 2014). However, little research has explored how probe frequency effects emerge and whether the manipulation involves the same cognitive mechanisms as the probe value manipulation. Chapter 2 therefore aimed to examine this.

In Experiment 1, significant probe value effects were observed, replicating previous research using cued-recall visual WM tasks (Allen & Ueno, 2018; Hitch et al., 2018; Hu et al., 2014; 2016). Significant probe frequency effects were also observed. This extends previous findings which has demonstrated that individuals can direct their attention to particular colours (Gorgoraptis et al., 2011; Klyszejko et al., 2014) or shapes (Cowan et al., 2010) that are more likely to be tested. These effects were additive, suggesting that the size of the probe value effect does not differ depending on the frequency with which the item is tested. It also suggests that probe value and probe frequency might involve distinct cognitive mechanisms. Experiment 2 provided further evidence for the latter claim, by demonstrating that probe frequency effects are not reliant on executive resources in a cued-recall visual WM task, as probe value effects appear to be (Hu et al., 2016). Alternatively, it may be that probe value and probe frequency involve the same mechanism, but that the way in which they access it may differ.

It is, however, important to note these experiments do not entirely rule out a strategic component to probe frequency effects. One possibility is that individuals strategically attended to the item that was most likely to be tested, but that doing so did not improve performance beyond the automatic effects observed (Hasher & Zacks, 1979). Alternatively, the complex concurrent task may have left some residual executive control resources, which could be used to strategically focus on the item that was more likely to be tested. Further research could examine this by investigating whether probe frequency effects are observed when a more cognitively demanding concurrent task is used, or by identifying whether probe frequency effects possess the characteristics expected of automatic processes (Hasher & Zacks, 1979). However, regardless of the outcomes of this, the current experiments provide clear evidence that probe frequency effects in WM are relatively more automatic than probe value effects.

These findings have important implications for the relationship between WM and attention, by demonstrating that not all forms of attentional direction are functionally equivalent. The probe value and probe frequency effects should therefore be explored further, as doing so is likely to provide novel insights. These findings also warn researchers against drawing generalisations across the paradigms, as this may result in erroneous conclusions. The findings might also have some practical implications, by demonstrating that individuals can experience probe value and probe frequency boosts simultaneously, with both independently enhancing performance.

6.2.2. Chapter 3

Research to date has explored probe value effects using visual stimuli (Chapter 2; Allen & Ueno, 2018; Berry et al., 2018; Hitch et al., 2018; Hu et al., 2014; 2016) or verbal stimuli presented visually (Sandry et al., 2014). Chapter 3 reflected a first attempt at examining whether participants can also direct their attention to more valuable information in an auditory-verbal WM task. Serial recall was used as the retrieval method, as this is commonly used in verbal WM tasks. Experiment 3 revealed that participants could prioritise a more valuable item appearing near the beginning, middle or end of a verbal sequence. This came at a cost to the less valuable items in some, but not all, conditions. This suggests that individuals can effectively prioritise more valuable information in verbal WM. It also provides evidence that the FoA in WM is likely to be modality general, as is suggested in several prominent WM models (Cowan, 2005; Hitch et al., 2018; Hu et al., 2014; Oberauer, 2013).

Two further experiments then examined whether value effects would emerge in this paradigm when rehearsal (Experiment 4), or rehearsal and executive control resources (Experiment 5) were disrupted. Based on previous research (e.g. Hu et al., 2016), it was predicted that value effects would be observed when rehearsal abilities were reduced, but that the boost would be smaller in size or completely abolished when executive control resources were disrupted. When rehearsal was disrupted, the value effect was significantly larger in size, indicating that the emergence of effects is not reliant on rehearsal (Experiment 4). Contrary to the predictions, the effect also remained significant when both rehearsal and executive control resources were reduced (Experiment 5).

This initially appears to suggest that value effects in a serial-recall auditory-verbal WM task are not reliant on rehearsal or executive control. However, observation of the serial position curves revealed that accuracy was at, or near floor, for the majority of the less valuable items when participants prioritised a more valuable item under complex concurrent task conditions. Furthermore, under complex concurrent task conditions, participants omitted significantly more of the low value items in the prioritisation condition than the control condition. Taken together, this suggests that participants may have abandoned these items in order to retain the more valuable item when executive control resources was reduced. This contrasts with Hu et al. (2016), who found that probe value effects were significantly reduced or abolished under concurrent task conditions in a cued-recall visual WM task. However, it is not clear whether these distinct outcomes reflect differences in the modality (i.e. verbal vs visual information) or the retrieval method used (i.e. serial vs cued recall). Nevertheless, it is clear that the effects of value and the approaches participants take may differ depending on task factors. This highlights a need for further research to examine how other factors might affect probe value boosts in WM. Such research would allow one to delineate those characteristics of value effects that vary depending on task factors, and those that are present regardless of task specifics.

As subjective awareness of the effects of probe value have not been explored to date, a further aim of Chapter 3 was to examine whether participants would be aware of the benefits and costs associated with prioritising a more valuable item. Across all three experiments, it was found that participants were generally aware that focusing on the more valuable item might enhance memory for this item, whilst negatively affecting memory for the less valuable items. This is in line with Berryhill et al. (2012), who found that participants were aware that retro-cues improved memory for the cued item. Nevertheless, within the current experiments, participants sometimes overestimated the costs and underestimated the benefits of prioritisation, showing a lack of complete understanding about the effects of value. Further analysis was also completed to explore whether performance on an FDR task would correlate with the size of the boost to the more valuable item and costs to the less valuable items. No significant correlations were observed with BF analysis also in favour of no effect, suggesting that simple storage capacity is not related to the size of value effects.

6.2.3. Chapter 4

Chapter 4 focused on examining whether children can also exhibit probe value effects in visual WM. Berry et al. (2018) recently reported that children aged 7-10 cannot prioritise more valuable information in WM, a finding that was attributed to underdeveloped executive control. Chapter 4 examined whether children aged 7-8 and 9-10 years can direct their attention based on value when they are particularly motivated to do so. Also of interest was whether these effects differed depending on memory load. The cued-recall visual WM task implemented in Chapter 2 was used, as this was similar to that used in Berry et al. (2018). To increase motivation, the main memory task was set within a child-friendly alien story and game. Participants were also told they would receive a physical prize if they collected enough ‘energy points’. Experiment 6 investigated effects using sequential presentation, whilst Experiment 7 presented items simultaneously. Significant probe value effects were found regardless of whether three or four items were presented in Experiment 6. This therefore provides the first evidence that children can prioritise more valuable information in WM, provided they are sufficiently motivated to do so. The probe value effects did not differ as a function of age group, although the effect sizes observed in both groups were considerably smaller than previous studies with adults (e.g. Chapter 2). This therefore suggests that further developmental changes are likely to occur between 10 years of age and adulthood. Such findings are in line with the cueing and probe frequency literature, whereby children can direct their attention in WM, although not always as effectively as adults (Cowan et al., 2010; Shimi et al., 2014; Shimi & Scerif, 2017).

Using simultaneous presentation (Experiment 7), significant probe value effects were observed when four items were presented, extending the findings from Experiment 6 and providing further evidence that children can prioritise more valuable information in WM when they are motivated. However, no significant probe value effects were observed when three items were presented. As performance in this condition was already relatively high, this was taken as evidence that children as young as 7-8 years of age selectively apply strategies when they believe doing so will enhance performance (Shimi & Scerif, 2017).

These findings have important implications, as they suggest that children can direct their attention to more valuable information in WM earlier than was previously assumed (i.e. in Berry et al., 2018). A further novel finding is that children can direct their attention in WM when items are presented sequentially (Experiment 6), as all previous studies reporting effects of cueing (Shimi et al., 2014; Shimi & Scerif, 2017) and probe frequency (Cowan et al., 2010) have presented items simultaneously. Furthermore, evidence that children successfully prioritised the more valuable item in this series of experiments, but not in Berry et al.

(2018) where motivation was reduced, has broad implications for researchers interested in children’s development. In line with previous research, it suggests that children may show cognitive abilities earlier in development if the task is meaningful and age appropriate (Borke, 1975; McGarrigle & Donaldson, 1974; Rose & Blank, 1974; Light et al., 1979; Watanabe, 2017). It also provides further evidence that children may not be as motivated to perform well in experiments as adults (Brewer et al., 2013). Based on this, it is recommended that researchers avoid simply taking a paradigm that yields significant effects in adults and exploring these effects developmentally. Instead, tasks should be adapted to ensure they are interesting and appealing to children, as this is likely to enhance motivation and encourage children to perform to their best of their abilities.

6.2.4. Chapter 5

The previous chapters have demonstrated that adults can direct their attention to more valuable information during visual and verbal WM tasks, and that children can also prioritise a valuable item when they are sufficiently motivated to do so. For the first time, Chapter 5 examined whether prioritising an item for a WM test yields durable effects that can be observed on a surprise LTM test. At WM, significant probe value effects were observed. This extends previous findings by demonstrating that individuals can prioritise more valuable information when real-world objects are used (as opposed to geometric shapes) and when recognition is used as the retrieval method (Chapters 2-4; Allen & Ueno, 2018; Hitch et al., 2018; Hu et al., 2014; 2016; Sandry et al., 2014). At LTM, a significant probe value effect was observed, but only when the more valuable item had been tested at WM. Moreover, further subsidiary analysis revealed that this effect was abolished when the analysis was limited to trials where participants responded correctly at WM. This suggests that prioritisation of a more valuable item for a WM test yields significant probe value effects, which then enhances performance on a LTM test.

Evidence that no independent effect emerged at LTM contrasts with findings from the cueing literature, in which retro-cues have been found to enhance memory after a delay (Reaves et al., 2016; Strunk et al., 2018). It also contrasts with findings from the value-based LTM literature, whereby increasing the value of an item enhances later memory (e.g. Adcock et al., 2006; Gruber & Otten, 2010). There are several reasons why independent LTM effects might not have emerged within this study. For instance, it may be that participants must be aware of the LTM test in order to obtain an independent LTM effect. Alternatively, it may be that the recognition test employed may have resulted in near ceiling performance or increased interference, which then prevented an independent effect at LTM emerging. Other factors may also have resulted in this finding, including the reward offered and the length of the delay before the LTM test. It would be useful for further research to examine these possibilities.

A second aim of Chapter 5 was to examine whether manipulating probe value affects RTs at both WM and LTM. To date, only one study had assessed effects to RTs at WM (Sandry et al., 2014), with other research focusing solely on accuracy as the outcome measure (Chapters 2-4; Allen & Ueno, 2018; Berry et al., 2018; Hitch et al., 2018; Hu et al., 2014; 2016). However, if probe value effects result from the more valuable item being held in a specialist sub-region that increases accessibility to the item (such as the FoA; Hitch et al., 2018; Hu et al., 2016), prioritising a more valuable item should also result in faster RTs. This was observed, with participants responding faster at SP1 when this item was more valuable, relative to the condition where all items were equally valuable. In contrast, no significant RT effects were observed at LTM. This suggests that probe value increases accessibility at WM, although this effect is short-lived and not durable. As such, this provides further evidence that prioritising a more valuable item results in that item being held in the FoA within WM (Hitch et al., 2018; Hu et al., 2014). This renders it more easily accessible, but these effects dissipate when the item leaves the sub-region.

A final aim was to further explore awareness of probe value effects, as examined for the first time in Chapter 3. Further supporting these previous outcomes, it was found that participants have a reasonable, but not complete, understanding about the effects of probe value.

6.3. Theoretical implications

This thesis primarily aimed to explore whether adults and children can prioritise more valuable information in working memory in a variety of task contexts. However, the outcomes of the work may also provide some insights into the cognitive mechanisms that may underlie these effects. In the majority of the experiments (Experiments 1-2 and 4-8), participants were asked to engage in articulatory suppression in order to disrupt articulation of the to-be-remembered information (Baddeley, 1986). Probe value effects were observed in all experiments, suggesting that individuals could still direct their attention towards the high value item when articulation of the memoranda was disrupted. Further supporting this claim, the probe value effect was actually significantly larger in size in Experiment 4 when participants concurrently engaged in articulatory suppression relative to a condition in which no suppression was required. Taken together, these findings suggest that verbal rehearsal is unlikely to be the main cognitive mechanism driving probe value effects in WM.

What might then be the cognitive mechanisms driving value effects in WM? Hu et al. (2016) found that probe value effects in a cued-recall visual WM task were either abolished or considerably reduced if individuals engaged in a cognitively demanding concurrent task. This pattern of findings was not observed in Experiment 5, with participants able to prioritise more valuable verbal information whilst completing a cognitively demanding concurrent task. They did, however, appear to do so by abandoning the less valuable items. This was therefore taken as evidence that individuals directed residual attentional resources towards retaining the more valuable digit. Taken together, this suggests that the cognitive mechanism underlying probe value effects in WM is likely to be reliant on attentional resources. The effects might therefore emerge from a process such as WM consolidation or attentional refreshing. For instance, individuals might either consolidate the more valuable item into WM better than the other items or refresh it more than the other items. Previous research can provide some insights into these claims. For instance, several studies have found that probe value effects in WM are disrupted by visual interference (i.e. a to-be-ignored suffix) presented after the set of items (Allen & Ueno, 2018; Hitch et al., 2018; Hu et al., 2014). Such outcomes are not consistent with a biased consolidation account, as consolidation is thought to protect items from perceptual interference (De Schrijver & Barrouillet, 2017). An attentional refreshing account might, however, predict this pattern of results, as the representation of the more valuable item could become distorted if it is being actively refreshed when the perceptual interference (e.g. the post-stimulus suffix) is encountered. Moreover, given that individuals are told the value information in advance of encoding, it is likely that effects at least partially reflect participants directing their attention to the high value item during encoding (e.g. by diverting their gaze towards the spatial location of the more valuable item during the encoding phase or the ISI). These suggestions are speculative, however, warranting further research to investigate the cognitive mechanisms underlying probe value effects in WM.

The findings presented in this body of work can inform WM models in several important ways. Multicomponent models of WM suggest that information enters the system from the external world and is separated into modality-specific stores (Baddeley, 2000; Logie; 2011). Control processes are carried out by the central executive (Baddeley, 1986; 2000), or through interactions between multiple executive functions (Logie, 1995; 2011; 2016). Within Baddeley's model (Baddeley, 2000), the final item and particularly goal relevant information can be held in a modality-general FoA within the episodic buffer, which temporarily increases accessibility (Hu et al., 2014). The way in which these types of information enters the FoA is thought to differ, however, with the final item entering the sub-system automatically, whilst goal-relevant information is held strategically (Hu et al., 2016). The current findings are consistent with this claim, by demonstrating that information that is more likely to be tested at retrieval, and hence more task-relevant, might enter the FoA relatively automatically (Chapter 2). This body of work is also consistent with claims that the FoA is modality-general (Chapter 3), and that holding items within the sub-system yields only temporary boosts (Chapter 5). Finally, Chapter 4 is consistent with suggestions that a FoA exists in children (Berry et al., 2019), although there is some evidence that this group might be less willing or less able to use this sub-system to focus their attention on particularly important information relative to adults.

The current findings also support several key features of Cowan's embedded processes model (Cowan, 1999). Cowan's model suggests that WM comprises an activated form of LTM, as well as a FoA that can store integrated information (Cowan, 1999; 2016). The FoA is thought to zoom in and out on particular representations (Cowan, 2005), with information able to enter the sub-system voluntarily or automatically (Cowan, 1988). The evidence described above is therefore also consistent with Cowan's interpretation of the FoA. For instance, evidence that the prioritisation of valuable information yields only transient boosts (Chapter 5) would be predicted from Cowan's model which states that the FoA is involved in temporary processes. Similarly, evidence that the FoA can hold both verbal and visuo-spatial representations (Chapter 3) is a key feature of Cowan's model. Evidence that the effects of value under concurrent task demands may differ depending on task factors such as the modality or the retrieval method (Chapter 3) might also be well explained by Cowan's FoA, which suggests that the sub-system may zoom in and out on particular representations depending on task demands (Cowan, 2005).

Thus, although the current findings are consistent with elements of both multicomponent and domain-general models, they cannot distinguish between them. This highlights important similarities between the models in terms of the FoA and how particular information in WM can be stored in a highly accessible manner. Such similarities have also been highlighted in the literature, where it has been suggested that the models may not be as distinct as they were once considered (Cowan, Saults, & Blume, 2014; Gray et al., 2017; Hornung, Brunner, Reuter, & Martin, 2011; Hu et al., 2016). This may be particularly true since the addition of the episodic buffer component (Baddeley, 2000; Logie, 2011), which might be considered similar in nature to Cowan's conceptualisation of the FoA (Gray et al., 2017).

6.4. Limitations and further directions

6.4.1. Limitations

One limitation of the studies conducted within this thesis is that they sometimes differ from previous research on more than one task factor, thus limiting the conclusions that can be drawn. For instance, Chapter 3 aimed to examine whether adults could prioritise more valuable information in verbal WM. As serial recall is commonly used to assess memory for verbal material, this retrieval method was implemented. This has some benefits, as the emergence of probe value boosts across these differing features demonstrates the robustness of the effect. However, where differences emerge, it is unclear which factor is driving the difference. For instance, in Chapter 3, it was found that the probe value boost was observed even when participants completed an attentionally-demanding concurrent task. This contrasts with Hu et al. (2016), where the probe value effects were reduced or abolished when participants completed a cognitively demanding concurrent task. However, these studies differ in the type of material presented and the retrieval method, thus making it impossible to conclude which factor is driving the differences in findings.

Another limitation is that conclusions are sometimes drawn by comparing the outcomes of the studies reported in this thesis to previous research. Within Chapter

4, where the main conclusions were drawn from comparing the findings to Berry et al. (2018), cross-experimental analysis were completed as the data is publicly available online. However, this was not possible in Chapter 2 and 3, which compared the findings to Hu et al. (2016). Within these experiments, it would have been optimal to include a condition that sought to replicate Hu et al. (2016). This would have provided further support for these findings, and also allowed us to complete cross-experimental comparisons, analysing all of the data within a single set of analyses.

Finally, the questionnaires used in Chapters 3 and 5 were retrospective in nature, asking participants to retrieve thoughts about all of the conditions at the end of experiment. This may have created some level of interference or inaccuracy when participants were asked to think about the earlier conditions. Furthermore, directly asking participants what they believe the effects would be to the more valuable item and the less valuable items may have led participants to give answers they thought were socially desirable or logical. In order to overcome this, further research using cued-recall or recognition could ask participants to give confidence judgements on every trial or on randomly selected trials (Berryhill et al., 2012).

6.4.2. Future directions

Chapter 2 demonstrated that probe value effects differ from probe frequency effects. There is also some evidence in the literature that probe value effects differ from visual cueing (Allen & Ueno, 2018; Berryhill et al., 2012; Gunseli et al., 2015; Hommel et al., 2001; Shimi et al., 2014; Hitch et al., 2018; Hu et al., 2014; Makovski & Jiang, 2007; Makovski et al., 2008; Matsukura et al., 2007, van

Moorselaar et al., 2014; Souza & Oberauer, 2016). It would, however, be useful to further examine this, as such research would reveal characteristics that are similar and different between the paradigms. For instance, research could explore whether probe frequency and visual cueing effects are vulnerable to a post-stimulus suffix, as probe value effects appear to be (Allen & Ueno, 2018; Hitch et al., 2018; Hu et al., 2014). It would also be useful to examine whether individuals can prioritise more valuable information retrospectively, and whether any effects are similar, or smaller in size, to when participants are provided with the value information prior to encoding. This would further bridge the literature between the probe value and cueing literature, and identify further similarities or differences between the two paradigms. Furthermore, evidence that probe value effects are reduced in size when participants are only told the value information after item presentation would suggest that probe value effects at least partially reflect participants directing their attention to the more valuable item at encoding.

The remaining chapters within this thesis (i.e. Chapters 3-5) extended findings in the probe value literature by revealing some conditions in which individuals can prioritise more valuable information in WM. It would, however, be useful to explore probe value effects in adults and children in a range of other task contexts. This would provide further information on the conditions in which adults and children do, and do not, experience probe value boosts. Within children, research could examine whether probe value effects emerge when information is presented verbally, or when a different retrieval method is used (e.g. serial recall or recognition). It would also be useful to explore whether both age groups exhibit effects when precision-based measures are used (e.g. retrieval of a colour using a colour wheel). Mixture modelling could then be performed to explore the mechanisms underlying such effects. This would reveal whether effects reflect an increased probability of recalling the target item, increased precision, decreased probability of recalling a non-target item, or decreased guessing (Bays et al., 2009; Oberauer, Stoneking, & Wabersich, & Lin, 2017). Additional research is also warranted to further explore how the probe value manipulation affects RT in different populations (e.g. in children) and paradigms (e.g. a serial-recall visual WM task akin to that used in Chapter 3).

Individual differences in the size of the probe value effects emerged across all experiments, with some individuals experiencing a large boost whilst others exhibited no effect or even a cost to the more valuable item. It would therefore useful for further research to explore how probe value effects differ across individuals, and factors that might predict the size of probe value effects. Chapter 3 revealed that performance on a FDR task is not related to the size of value effects in a serial-recall verbal WM task. However, performance on WM tasks that require both storage and processing (such as BDR) might predict the size of effects. It may also be that executive functions more broadly predict the size of probe value effects. Levels of attentional control may be particularly important, as this would be likely to affect how individuals distribute their attention between the more valuable item and the less valuable items (Anderson, Laurent, & Yantis, 2011). Alternatively, or additionally, probe value effects might vary depending on the value one places on cognitive effort. This is thought to vary between individuals and may substantially affect performance on cognitively demanding tasks (Chevalier, 2018). As probe value effects are thought to rely on executive resources in at least some contexts (Hu et al., 2016), individuals that exhibit probe value boosts may be those more willing to exert cognitive effort (Chevalier, 2018). It would also be beneficial for further research to explore how probe value effects differ within individuals (e.g. across trials or task contexts).

One of the primary reasons for focusing on probe value effects within this thesis is that this paradigm might more readily translate into real-life settings relative to other manipulations such as probe frequency or visual cueing. Furthermore, in real-life settings, information might naturally differ by importance or value. As probe value effects have been observed in a range of task contexts in both adults and children, it would now be useful to establish whether the effects also emerge in more realistic settings or using more realistic material. One task that may lend itself nicely to such explorations is following instructions. This ability relies heavily on WM (Gathercole et al., 2008), is commonly required in classroom settings (Gathercole et al., 2008; Waterman et al., 2017), and is identified as an area of difficulty in children with poor WM (Alloway, Gathercole, Kirkwood, & Elliott, 2009; Atkinson, Allen, & Waterman, In Preparation). However, within classroom settings, all instructions given to children might not be equally important. For instance, in the following instructional sequence “finish your sentence, put your worksheet in the middle of the table, put your chair under the table, and come sit on the carpet”, sitting on the carpet may be the most important instruction, as it would ensure that children are in the correct location, settled, and ready for the next learning activity. Teachers could therefore add a reward to this final instruction, which might increase the likelihood of children remembering it. Other tasks that also rely on WM could be explored, including problem-solving and mental arithmetic. The reward offered may need to be appealing, however, as this appears to be an important factor in determining whether children prioritise more valuable information in WM (Chapter 4).

Finally, it would be useful to explore whether probe value effects are observed in other populations, such as older adults. Research examining visual cueing effects in older adults have yielded mixed findings, with some reporting that older adults are as able to direct their attention in WM as younger adults (Gilchrist, Duarte, & Verhaeghen, 2016; Loaiza & Souza, 2018; Mok et al., 2016; Strunk et al., 2018; Souza, 2016), whilst others have reported that they are less able to do so (Duarte et al., 2013; Newsome et al., 2015; Yi & Friedman, 2014). At least under some conditions, directing attention to more valuable information in WM appears to be cognitively demanding (Hu et al., 2016). One might therefore expect that older adults may be less able to prioritise particularly valuable information relative to younger adults. However, Atkinson et al. (2017) found that older adults benefited from focusing on a subset of items (relative to all items) to a similar extent as young adults in a visual WM task. As such, it is possible that older adults might be able to use their executive control resources to successfully direct their attention in WM. Such findings would provide insights into how the ability to prioritise more valuable items changes across the lifespan, whilst also establishing whether older adults are able to flexibly focus their attention on certain information in WM.

6.5. Conclusions

This thesis has provided evidence that the probe value manipulation may involve distinct cognitive mechanisms to probe frequency, whereby one of the items is more likely to be tested. Taken alongside existing literature suggesting that probe value and visual cueing encourage individuals to direct attention in different ways, this may be taken as evidence that probe value effects in WM arise from distinct cognitive mechanisms to either probe frequency or visual cueing. The remaining empirical chapters then focused on exploring probe value effects in more detail. Chapter 3 examined whether young adults can prioritise more valuable information in an auditory-verbal WM task requiring serial recall. Clear probe value effects were observed, which were not diminished in size when rehearsal and executive control were disrupted. However, when both rehearsal abilities and executive control resources were reduced, participants often abandoned the less valuable items in order to retain the more valuable item. Taken with previous work, this suggests that young adults can prioritise more valuable information in WM regardless of modality. Chapter 4 then focused on exploring whether children aged 7-10 years can prioritise more valuable information when sufficiently motivated to do so. There was clear evidence of this, with children exhibiting probe value effects across memory loads and presentation types. There was, however, some evidence that the children selectively applied the probe value strategy when they believed it would enhance performance. Finally, Chapter 5 examined whether prioritising a more valuable item for a WM test enhanced LTM accuracy. A boost was observed at LTM when the more valuable item was tested at WM, although this appeared to be driven by an enhancement to WM performance rather than an independent LTM effect. Taken together, these studies provide evidence that probe value effects in WM are distinct, robust, and observable in a range of different task settings and populations. If the more valuable item is tested at WM, this can also yield LTM boosts, although this appear to be driven by enhanced performance at WM rather than an independent effect at LTM *per se*.

REFERENCES

Adcock, R. A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., & Gabrieli, J. D. E. (2006). Reward-motivated learning: mesolimbic activation precedes memory formation. *Neuron*, *50*(3), 507–517. https://doi.org/10.1016/j.neuron.2006.03.03

Allen, R. J., Baddeley, A. D., & Hitch, G. J. (2006). Is the binding of visual features in working memory resource-demanding? *Journal of Experimental Psychology: General*, *135*(2), 298–313. https://doi.org/10.1037/0096-3445.135.2.298

Allen, R. J., Baddeley, A. D., & Hitch, G. J. (2014). Evidence for two attentional components in visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(6), 1499–1509. https://doi.org/10.1037/xlm000000

Allen, R. J., & Ueno, T. (2018). Multiple high-reward items can be prioritized in working memory but with greater vulnerability to interference. *Attention, Perception & Psychophysics*, *80*(7), 1731–1743. https://doi.org/10.3758/s13414-018-1543-6

Alloway, T. P. (2006). How does working memory work in the classroom? *Educational Research and Reviews*, *1*(4), 134–139. Retrieved from http://dspace.stir.ac.uk/handle/1893/786

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

Alloway, T. P., Alloway, R. G., & Wootan, S. (2014). Home sweet home: does where you live matter to working memory and other cognitive skills? *Journal of Experimental Child Psychology*, *124*, 124–131. https://doi.org/10.1016/j.jecp.2013.11.012

Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child Development*, *80*(2), 606–621. https://doi.org/10.1111/j.1467-8624.2009.01282.x

Altmann, E. M., & Schunn, C. D. (2012). Decay versus interference: a new look at an old interaction. *Psychological Science*, *23*(11), 1435–1437. https://doi.org/10.1177/0956797612446027

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

Anderson, P. J., Lee, K. J., Roberts, G., Spencer-Smith, M. M., Thompson, D. K., Seal, M. L., … Pascoe, L. (2018). Long-Term Academic Functioning Following Cogmed Working Memory Training for Children Born Extremely Preterm: A Randomized Controlled Trial. *The Journal of Pediatrics*, *202*, 92–97.e4. https://doi.org/10.1016/j.jpeds.2018.07.003

Arán-Filippetti, V. (2013). Structure and Invariance of Executive Functioning Tasks across Socioeconomic Status: Evidence from Spanish-Speaking Children. *The Spanish Journal of Psychology*, *16*. https://doi.org/10.1017/sjp.2013.102

Astle, D. E., Nobre, A. C., & Scerif, G. (2012). Attentional control constrains visual short-term memory: insights from developmental and individual differences.

Quarterly Journal of Experimental Psychology , *65*(2), 277–294. https://doi.org/10.1080/17470218.2010.492622

Astle, D. E., Summerfield, J., Griffin, I., & Nobre, A. C. (2012). Orienting attention to locations in mental representations. *Attention, Perception & Psychophysics*, *74*(1), 146–162. https://doi.org/10.3758/s13414-011-0218-3

Atkinson, A. L., Baddeley, A. D., & Allen, R. J. (2018). Remember some or remember all? Ageing and strategy effects in visual working memory. *Quarterly Journal of Experimental Psychology* , *71*(7), 1561–1573. https://doi.org/10.1080/17470218.2017.1341537

Atkinson, A. L., Berry, E. D. J., Waterman, A. H., Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2018). Are there multiple ways to direct attention in working memory? *Annals of the New York Academy of Sciences*, *1424*(1), 115–126. https://doi.org/10.1111/nyas.13634

Atkinson, A. L., Waterman, A. H., & Allen, R. J. (2019). Can children prioritize more valuable information in working memory? An exploration into the effects of motivation and memory load. *Developmental Psychology*, *55*(5), 967–980. https://doi.org/10.1037/dev0000692

Atkinson, A. L, Allen, R. J, Baddeley, A. D, Hitch, G. J, Waterman, A. H. (Under Revision). Can valuable information be prioritised in verbal working memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Atkinson, A. L, Allen, R. J., Waterman, A. H. (In Preparation). *To what extent do teachers and other educational professionals understand working memory? Analysis of 1,425 responses*.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In *Psychology of learning and motivation* (Vol. 2, pp. 89–195).

Attig, M., & Hasher, L. (1980). The processing of frequency of occurrence information by adults. *Journal of Gerontology*, *35*(1), 66–69. https://doi.org/10.1093/geronj/35.1.66

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences*, *5*(3), 119–126. https://doi.org/10.1016/s1364-6613(00)01593-x

Backer, K. C., & Alain, C. (2012). Orienting attention to sound object representations attenuates change deafness. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(6), 1554–1566. https://doi.org/10.1037/a0027858

Backer, K. C., Binns, M. A., & Alain, C. (2015). Neural dynamics underlying attentional orienting to auditory representations in short-term memory. *Journal of Neuroscience*, *35*(3), 1307–1318. https://doi.org/10.1523/JNEUROSCI.1487-14.2015

Baddeley, A. (1986). *Working memory*. Oxford: Oxford University Press, Clarendon Press

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: looking back and looking forward. *Nature Reviews Neuroscience*, *4*(10), 829–839. https://doi.org/10.1038/nrn1201

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Baddeley, A. D. (1997). *Human memory: Theory and practice* (2nd Ed.). Psychology Press.

Baddeley, A. D., & Broadbent, D. E. (1983). Working memory. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *302*(1110), 311–324. https://doi.org/10.1098/rstb.1983.0057

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

Baddeley, A. D., & Hitch, G. (1993). The recency effect: implicit learning with explicit retrieval? *Memory & Cognition*, *21*(2), 146–155. https://doi.org/10.3758/bf03202726

Baddeley, A. D., & Hitch, G. J. (1994). Developments in the concept of working memory. *Neuropsychology*, *8*(4), 485-493.

Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2009). Working memory and binding in sentence recall. *Journal of Memory and Language*, *61*(3), 438–456. https://doi.org/10.1016/j.jml.2009.05.004

Baddeley, A., Lewis, V., & Vallar, G. (1984). Exploring the Articulatory Loop. *The Quarterly Journal of Experimental Psychology Section A*, *36*(2), 233–252. https://doi.org/10.1080/14640748408402157

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental*

Psychology: General, *133*(1), 83–100. https://doi.org/10.1037/0096-3445.133.1.83

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 570–585. https://doi.org/10.1037/0278-7393.33.3.570

Barrouillet, P., & Camos, V. (2012). As Time Goes By: Temporal Constraints in Working Memory. *Current Directions in Psychological Science*, *21*(6), 413–419. https://doi.org/10.1177/0963721412459513

Bayliss, D. M., Bogdanovs, J., & Jarrold, C. (2015). Consolidating working memory: Distinguishing the effects of consolidation, rehearsal and attentional refreshing in a working memory span task. *Journal of Memory and Language*, *81*, 34–50. https://doi.org/10.1016/j.jml.2014.12.004

Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, *9*(10), 7.1–11. https://doi.org/10.1167/9.10.7

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

Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E.-J., Berk, R., … Johnson, V. E. (2018). Redefine statistical significance. *Nature Human Behaviour*, *2*(1), 6–10. https://doi.org/10.1038/s41562-017-0189-z

Berry, E. D. J., Waterman, A. H., Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2018). The limits of visual working memory in children: Exploring prioritization and recency effects with sequential presentation.

Developmental Psychology, *54*(2), 240–253. https://doi.org/10.1037/dev0000427

Berryhill, M. E., Richmond, L. L., Shay, C. S., & Olson, I. R. (2012). Shifting attention among working memory representations: testing cue type, awareness, and strategic control. *Quarterly Journal of Experimental Psychology* , *65*(3), 426–438. https://doi.org/10.1080/17470218.2011.604786

Bireta, T. J., & Mazzei, C. M. (2016). Does the isolation effect require attention? *Memory & Cognition*, *44*(1), 1–14. https://doi.org/10.3758/s13421-015-0538-y

Blackwell, K. A., & Munakata, Y. (2014). Costs and benefits linked to developments in cognitive control. *Developmental Science*, *17*(2), 203–211. https://doi.org/10.1111/desc.12113

Borke, H. (1975). Piaget's mountains revisited: Changes in the egocentric landscape. *Developmental Psychology*, *11*(2), 240–243. https://doi.org/10.1037/h0076459

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences*, *105*(38), 14325–14329. https://doi.org/10.1073/pnas.0803390105

Braun, E. K., Wimmer, G. E., & Shohamy, D. (2018). Retroactive and graded prioritization of memory by reward. *Nature Communications*, *9*(1), 4886. https://doi.org/10.1038/s41467-018-07280-0

Brewer, R., Anthony, L., Brown, Q., Irwin, G., Nias, J., & Tate, B. (2013). Using Gamification to Motivate Children to Complete Empirical Studies in Lab Environments. *Proceedings of the 12th International Conference on*

Interaction Design and Children, 388–391. https://doi.org/10.1145/2485760.2485816

Broadbent, D. E. (1958). *Perception and communication. Elmsford, NY, US.* Pergamon Press. http://dx. doi. org/10.1037/10037-000.

Brockmole, J. R., & Logie, R. H. (2013). Age-Related Change in Visual Working Memory: A Study of 55,753 Participants Aged 8–75. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00012

Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The Bank of Standardized Stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PloS One*, *5*(5), e10773. https://doi.org/10.1371/journal.pone.0010773

Brodeur, M. B., Guérard, K., & Bouras, M. (2014). Bank of Standardized Stimuli (BOSS) phase II: 930 new normative photos. *PloS One*, *9*(9), e106953. https://doi.org/10.1371/journal.pone.0106953

Calia, C., Darling, S., Havelka, J., & Allen, R. J. (2019). Visuospatial bootstrapping: Binding useful visuospatial information during verbal working memory encoding does not require set-shifting executive resources. *Quarterly Journal of Experimental Psychology*, *72*(4), 913-921.

Camos, V., & Barrouillet, P. (2014). Attentional and non-attentional systems in the maintenance of verbal information in working memory: the executive and phonological loops. *Frontiers in Human Neuroscience*, *8*, 900. https://doi.org/10.3389/fnhum.2014.00900

Camos, V., Johnson, M., Loaiza, V., Portrat, S., Souza, A., & Vergauwe, E. (2018). What is attentional refreshing in working memory? *Annals of the New York Academy of Sciences*, *1424*(1), 19–32. https://doi.org/10.1111/nyas.13616

Camos, V., Lagner, P., & Barrouillet, P. (2009). Two maintenance mechanisms of verbal information in working memory. *Journal of Memory and Language*, *61*(3), 457–469. https://doi.org/10.1016/j.jml.2009.06.002

Camos, V., Mora, G., & Oberauer, K. (2011). Adaptive choice between articulatory rehearsal and attentional refreshing in verbal working memory. *Memory & Cognition*, *39*(2), 231–244. https://doi.org/10.3758/s13421-010-0011-x

Carretti, B., Borella, E., & De Beni, R. (2007). Does strategic memory training improve the working memory performance of younger and older adults? *Experimental Psychology*, *54*(4), 311–320. https://doi.org/10.1027/1618-3169.54.4.311

Castel, A. D., Benjamin, A. S., Craik, F. I. M., & Watkins, M. J. (2002). The effects of aging on selectivity and control in short-term recall. *Memory & Cognition*, *30*(7), 1078–1085. https://doi.org/10.3758/bf03194325

Castel, A. D., Humphreys, K. L., Lee, S. S., Galván, A., Balota, D. A., & McCabe, D. P. (2011). The development of memory efficiency and value-directed remembering across the life span: a cross-sectional study of memory and selectivity. *Developmental Psychology*, *47*(6), 1553–1564. https://doi.org/10.1037/a0025623

Castel, A. D., Lee, S. S., Humphreys, K. L., & Moore, A. N. (2011). Memory capacity, selective control, and value-directed remembering in children with and without attention-deficit/hyperactivity disorder (ADHD). *Neuropsychology*, *25*(1), 15–24. https://doi.org/10.1037/a0020298

Chan, J. C. K., & McDermott, K. B. (2007). The testing effect in recognition memory: a dual process account. *Journal of Experimental Psychology:*

Learning, Memory, and Cognition, *33*(2), 431–437. https://doi.org/10.1037/0278-7393.33.2.431

Chevalier, N. (2018). Willing to Think Hard? The Subjective Value of Cognitive Effort in Children. *Child Development*, *89*(4), 1283–1295. https://doi.org/10.1111/cdev.12805

Chevalier, N., James, T. D., Wiebe, S. A., Nelson, J. M., & Espy, K. A. (2014). Contribution of reactive and proactive control to children's working memory performance: Insight from item recall durations in response sequence planning. *Developmental Psychology*, *50*(7), 1999–2008. https://doi.org/10.1037/a0036644

Chun, M. M. (2011). Visual working memory as visual attention sustained internally over time. *Neuropsychologia*, *49*(6), 1407–1409. https://doi.org/10.1016/j.neuropsychologia.2011.01.029

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd Ed.). New York, NY: Routledge.

Comblain, A. (1994). Working memory in Down's syndrome: Training the rehearsal strategy. *Down's Syndrome, Research and Practice*, *2*(3), 123–126. Retrieved from https://orbi.uliege.be/handle/2268/13380

Connor, C. E., Egeth, H. E., & Yantis, S. (2004). Visual attention: bottom-up versus top-down. *Current Biology: CB*, *14*(19), R850–R852. https://doi.org/10.1016/j.cub.2004.09.041

Conrad, R. (1971). The chronology of the development of covert speech in children. *Developmental Psychology*, *5*(3), 398–405. https://doi.org/10.1037/h0031595

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*(5), 769–786. https://doi.org/10.3758/bf03196772

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 Process Model of Working Memory. In A. Miyake. and P, Shah. (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 62-101). Cambridge University Press.

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. (2005). Working memory capacity. Hove, England: Psychology Press.

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. (2014). Working Memory Underpins Cognitive Development, Learning, and Education. *Educational Psychology Review*, *26*(2), 197–223. https://doi.org/10.1007/s10648-013-9246-y

Cowan, N. (2016). Working Memory Maturation: Can We Get at the Essence of Cognitive Growth? *Perspectives on Psychological Science*, *11*(2), 239–264. https://doi.org/10.1177/1745691615621279

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., Morey, C. C., AuBuchon, A. M., Zwilling, C. E., & Gilchrist, A. L. (2010). Seven-year-olds allocate attention like adults unless working memory is overloaded. *Developmental Science*, *13*(1), 120–133. https://doi.org/10.1111/j.1467-7687.2009.00864.x

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

Criss, A. H., Malmberg, K. J., & Shiffrin, R. M. (2011). Output interference in recognition memory. *Journal of Memory and Language*, *64*(4), 316–326. https://doi.org/10.1016/j.jml.2011.02.003

Cumming, G. (2014). The new statistics: why and how. *Psychological Science*, *25*(1), 7–29. https://doi.org/10.1177/0956797613504966

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3*(4), 422–433. https://doi.org/10.3758/BF03214546

De Schrijver, S., & Barrouillet, P. (2017). Consolidation and restoration of memory traces in working memory. *Psychonomic Bulletin & Review*, *24*(5), 1651–1657. https://doi.org/10.3758/s13423-017-1226-7

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

Dhaliwal, S., & Campbell, M. J. (2010). Misinterpreting p-values in research. *The Australasian Medical Journal*, *1*(1), 1–2. https://doi.org/10.4066/AMJ.2009.191

Dobbs, A. R., & Rule, B. G. (1989). Adult age differences in working memory. *Psychology and Aging*, *4*(4), 500–503. https://doi.org/10.1037//0882-7974.4.4.500

Donaldson, M., & McGarrigle, J. (1974). Some clues to the nature of semantic development. *Journal of Child Language*, *1*(2), 185–194. https://doi.org/10.1017/S0305000900000635

Dovis, S., Van der Oord, S., Wiers, R. W., & Prins, P. J. M. (2012). Can Motivation Normalize Working Memory and Task Persistence in Children with Attention-Deficit/Hyperactivity Disorder? The Effects of Money and Computer-Gaming. *Journal of Abnormal Child Psychology*, *40*(5), 669–681. https://doi.org/10.1007/s10802-011-9601-8

Duarte, A., Hearons, P., Jiang, Y., Delvin, M. C., Newsome, R. N., & Verhaeghen, P. (2013). Retrospective attention enhances visual working memory in the young but not the old: an ERP study. *Psychophysiology*, *50*(5), 465–476. https://doi.org/10.1111/psyp.12034

Ecker, U. K. H., Lewandowsky, S., & Oberauer, K. (2014). Removal of information from working memory: A specific updating process. *Journal of Memory and Language*, *74*, 77–90. https://doi.org/10.1016/j.jml.2013.09.003

Elke, S., & Wiebe, S. A. (2017). Proactive control in early and middle childhood: An ERP study. *Developmental Cognitive Neuroscience*, *26*, 28–38. https://doi.org/10.1016/j.dcn.2017.04.005

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.

Faulkenberry, T. J. (2018). Computing Bayes factors to measure evidence from experiments: An extension of the BIC approximation. *Biometrical Letters*, *55*(1), 31–43.

Festini, S. B., & Reuter-Lorenz, P. A. (2013). The short- and long-term consequences of directed forgetting in a working memory task. *Memory*, *21*(7), 763–777. https://doi.org/10.1080/09658211.2012.754900

Flavell, J. H., Beach, D. R., & Chinsky, J. M. (1966). Spontaneous verbal rehearsal in a memory task as a function of age. *Child Development*, *37*(2), 283–299.

Fliessbach, K., Hoppe, C., Schlegel, U., Elger, C. E., & Helmstaedter, C. (2006). NeuroCogFX--a computer-based neuropsychological assessment battery for the follow-up examination of neurological patients. *Fortschritte Der Neurologie-Psychiatrie*, *74*(11), 643–650.

Forbes, N. F., Carrick, L. A., McIntosh, A. M., & Lawrie, S. M. (2009). Working memory in schizophrenia: a meta-analysis. *Psychological Medicine*, *39*(6), 889–905. https://doi.org/10.1017/S0033291708004558

Fougnie, D. (2009). The relationship between attention and working memory. In N. B. Johansen (Ed.), *New research on short-term memory* (pp. 1–45). New York: Nova Science.

Galton, F. (1883). *Inquiries into human faculty and its development* (Everyman edition). London: Dent.

Gathercole, S. E., Durling, E., Evans, M., Jeffcock, S., & Stone, S. (2008). Working memory abilities and children's performance in laboratory analogues of classroom activities. *Applied Cognitive Psychology*, *22*(8), 1019–1037.

Gathercole, S. E., Lamont, E., & Alloway, T. P. (2006). Working memory in the classroom. In S. Pickering (Ed.), *Working memory and education* (pp. 219–240). Oxford, UK: Elsevier.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, *40*(2), 177-190.

Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. (2004). Working memory skills and educational attainment: evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology*, *18*(1), 1–16. https://doi.org/10.1002/acp.934

Gilchrist, A. L., Duarte, A., & Verhaeghen, P. (2016). Retrospective cues based on object features improve visual working memory performance in older adults. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *23*(2), 184–195. https://doi.org/10.1080/13825585.2015.1069253

Gold, J. M., Hahn, B., Zhang, W. W., Robinson, B. M., Kappenman, E. S., Beck, V. M., & Luck, S. J. (2010). Reduced capacity but spared precision and maintenance of working memory representations in schizophrenia. *Archives of General Psychiatry*, *67*(6), 570–577. https://doi.org/10.1001/archgenpsychiatry.2010.65

Goodman, S. N. (2001). Of P-values and Bayes: a modest proposal. *Epidemiology*, *12*(3), 295–297. https://doi.org/10.1097/00001648-200105000-00006

Gorgoraptis, N., Catalao, R. F. G., Bays, P. M., & Husain, M. (2011). Dynamic updating of working memory resources for visual objects. *Journal of Neuroscience*, *31*(23), 8502–8511. https://doi.org/10.1523/JNEUROSCI.0208-11.2011

Gray, S., Green, S., Alt, M., Hogan, T. P., Kuo, T., Brinkley, S., & Cowan, N. (2017). The Structure of Working Memory in Young Children and Its Relation to Intelligence. *Journal of Memory and Language*, *92*, 183–201. https://doi.org/10.1016/j.jml.2016.06.004

Gressmann, M., & Janczyk, M. (2016). The (Un)Clear Effects of Invalid Retro-Cues. *Frontiers in Psychology*, *7*, 244. https://doi.org/10.3389/fpsyg.2016.00244

Gruber, M. J., & Otten, L. J. (2010). Voluntary control over prestimulus activity related to encoding. *Journal of Neuroscience*, *30*(29), 9793–9800. https://doi.org/10.1523/JNEUROSCI.0915-10.2010

Gunseli, E., van Moorselaar, D., Meeter, M., & Olivers, C. N. L. (2015). The reliability of retro-cues determines the fate of noncued visual working memory representations. *Psychonomic Bulletin & Review*, *22*(5), 1334–1341. https://doi.org/10.3758/s13423-014-0796-x

Hackman, D. A., Gallop, R., Evans, G. W., & Farah, M. J. (2015). Socioeconomic status and executive function: developmental trajectories and mediation. *Developmental Science*, *18*(5), 686–702. https://doi.org/10.1111/desc.12246

Halsey, L. G. (2019). The reign of the p-value is over: what alternative analyses could we employ to fill the power vacuum? *Biology Letters*, *15*(5), 20190174. https://doi.org/10.1098/rsbl.2019.0174

Hammer, R., Tennekoon, M., Cooke, G. E., Gayda, J., Stein, M. A., & Booth, J. R. (2015). Feedback associated with expectation for larger-reward improves visuospatial working memory performances in children with ADHD. *Developmental Cognitive Neuroscience*, *14*, 38–49. https://doi.org/10.1016/j.dcn.2015.06.002

Hartshorne, J. K. (2008). Visual working memory capacity and proactive interference. *PloS One*, *3*(7), e2716. https://doi.org/10.1371/journal.pone.0002716

Hasher, L., & Chromiak, W. (1977). The processing of frequency information: An automatic mechanism? *Journal of Verbal Learning and Verbal Behavior*, *16*(2), 173–184. https://doi.org/10.1016/S0022-5371(77)80045-5

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, *108*(3), 356–388. https://doi.org/10.1037/0096-3445.108.3.356

Hasher, L., & Zacks, R. T. (1984). Automatic processing of fundamental information: the case of frequency of occurrence. *The American Psychologist*, *39*(12), 1372–1388. https://doi.org/10.1037//0003-066x.39.12.1372

Hasher, L., Zacks, R. T., Rose, K. C., & Sanft, H. (1987). Truly incidental encoding of frequency information. *The American Journal of Psychology*, *100*(1), 69–91.

Hitch, G. J., & Halliday, M. S. (1983). Working memory in children. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *302*(1110), 325-340. https://doi.org/10.1098/rstb.1983.0058

Hitch, G. J., Halliday, M. S., Dodd, A., & Littler, J. E. (1989). Development of rehearsal in short-term memory: Differences between pictorial and spoken stimuli. *British Journal of Developmental Psychology*, *7*(4), 347–362.

Hitch, G. J., Hu, Y., Allen, R. J., & Baddeley, A. D. (2018). Competition for the focus of attention in visual working memory: perceptual recency versus executive control. *Annals of the New York Academy of Sciences*, *1424*(1), 64–75. https://doi.org/10.1111/nyas.13631

Hebb, D. O. (1961). Distinctive features of learning in the higher animal. In J. F. Delafresnaye (Ed.), Brain mechanisms and learning (pp. 37-46). Oxford: Blackwell.

Hollingworth, A., & Maxcey-Richard, A. M. (2013). Selective maintenance in visual working memory does not require sustained visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(4), 1047–1058. https://doi.org/10.1037/a0030238

Holmes, J., & Adams, J. W. (2006). Working memory and children’s mathematical skills: Implications for mathematical development and mathematics curricula. *Educational Psychology*, *26*(3), 339–366.

Hommel, B., Pratt, J., Colzato, L., & Godijn, R. (2001). Symbolic control of visual attention. *Psychological Science*, *12*(5), 360–365. https://doi.org/10.1111/1467-9280.00367

Hornung, C., Brunner, M., Reuter, R. A. P., & Martin, R. (2011). Children’s working memory: Its structure and relationship to fluid intelligence. *Intelligence*, *39*(4), 210–221. https://doi.org/10.1016/j.intell.2011.03.002

Hu, Y., Allen, R. J., Baddeley, A. D., & Hitch, G. J. (2016). Executive control of stimulus-driven and goal-directed attention in visual working memory. *Attention, Perception & Psychophysics*, *78*(7), 2164–2175. https://doi.org/10.3758/s13414-016-1106-7

Hu, Y., Hitch, G. J., Baddeley, A. D., Zhang, M., & Allen, R. J. (2014). Executive and perceptual attention play different roles in visual working memory: evidence from suffix and strategy effects. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(4), 1665–1678. https://doi.org/10.1037/a0037163

Hunt, R. R. (1995). The subtlety of distinctiveness: What von Restorff really did. *Psychonomic Bulletin & Review*, *2*(1), 105–112. https://doi.org/10.3758/BF03214414

Hunt, R. R., & Lamb, C. A. (2001). What causes the isolation effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(6), 1359–1366.

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

James, W. (1890). *Principles of Psychology, Vol. 1.(1905 ed.)*. London, UK: Methuen & Co.

Janczyk, M., & Berryhill, M. E. (2014). Orienting attention in visual working memory requires central capacity: decreased retro-cue effects under dual-task conditions. *Attention, Perception & Psychophysics*, *76*(3), 715–724. https://doi.org/10.3758/s13414-013-0615-x

Jaroslawska, A. J., Gathercole, S. E., & Holmes, J. (2018). Following instructions in a dual-task paradigm: Evidence for a temporary motor store in working memory. *Quarterly Journal of Experimental Psychology* , *71*(11), 2439–2449. https://doi.org/10.1177/1747021817743492

Jaroslawska, A. J., Gathercole, S. E., Logie, M. R., & Holmes, J. (2016). Following instructions in a virtual school: Does working memory play a role? *Memory & Cognition*, *44*(4), 580–589. https://doi.org/10.3758/s13421-015-0579-2

Jeffreys, H. (1961). *Theory of probability (3rd edition)*. New York: Oxford University Press.

Johnston, J. C., McCann, R. S., & Remington, R. W. (1995). Chronometric evidence for two types of attention. *Psychological Science*, *6*(6), 365–369.

Jolicoeur, P., & Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology*, *36*(2), 138–202. https://doi.org/10.1006/cogp.1998.0684

Jonides, J., Naveh-Benjamin, M., & Palmer, J. (1985). Assessing automaticity. *Acta Psychologica*, *60*(2), 157–171. https://doi.org/10.1016/0001-6918(85)90053-8

Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: a review of our current understanding. *Neuropsychology Review*, *17*(3), 213–233. https://doi.org/10.1007/s11065-007-9040-z

Kang, S. H. K., McDermott, K. B., & Roediger, H. L., III. (2007). Test format and corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology*, *19*(4-5), 528–558.

Katsuki, F., & Constantinidis, C. (2014). Bottom-up and top-down attention: different processes and overlapping neural systems. *The Neuroscientist*, *20*(5), 509–521. https://doi.org/10.1177/1073858413514136

Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences*, *14*(7), 317–324. https://doi.org/10.1016/j.tics.2010.05.002

Klyszejko, Z., Rahmati, M., & Curtis, C. E. (2014). Attentional priority determines working memory precision. *Vision Research*, *105*, 70–76. https://doi.org/10.1016/j.visres.2014.09.002

Komori, M. (2016). Effects of Working Memory Capacity on Metacognitive Monitoring: A Study of Group Differences Using a Listening Span Test. *Frontiers in Psychology*, *7*, 285. https://doi.org/10.3389/fpsyg.2016.00285

Krefeld-Schwalb, A. (2018). The Retro-Cue Benefit for Verbal Material and Its Influence on the Probability of Intrusions Under Dual-Task Conditions. *Experimental Psychology*, *65*(3), 128–138. https://doi.org/10.1027/1618-3169/a000400

Ku, Y. (2018). Selective attention on representations in working memory: cognitive and neural mechanisms. *PeerJ*, *6*, e4585. https://doi.org/10.7717/peerj.4585

Kumar, S., Joseph, S., Pearson, B., Teki, S., Fox, Z. V., Griffiths, T. D., & Husain, M. (2013). Resource allocation and prioritization in auditory working memory. *Cognitive Neuroscience*, *4*(1), 12–20. https://doi.org/10.1080/17588928.2012.716416

Kuo, B.-C., Stokes, M. G., & Nobre, A. C. (2012). Attention modulates maintenance of representations in visual short-term memory. *Journal of Cognitive Neuroscience*, *24*(1), 51–60. https://doi.org/10.1162/jocn_a_0008

Lakens, D., McLatchie, N., Isager, P. M., Scheel, A. M., & Dienes, Z. (2018). Improving Inferences about Null Effects with Bayes Factors and Equivalence Tests. *The Journals of Gerontology: Series B,* gby065. https://doi.org/10.1093/geronb/gby065

Landman, R., Spekreijse, H., & Lamme, V. A. F. (2003). Large capacity storage of integrated objects before change blindness. *Vision Research*, *43*(2), 149–164. https://doi.org/10.1016/s0042-6989(02)00402-9

LaRocque, J. J., Eichenbaum, A. S., Starrett, M. J., Rose, N. S., Emrich, S. M., & Postle, B. R. (2015). The short- and long-term fates of memory items retained outside the focus of attention. *Memory & Cognition*, *43*(3), 453–468. https://doi.org/10.3758/s13421-014-0486-y

Lee, E., Cowan, N., Vogel, E. K., Rolan, T., Valle-Inclán, F., & Hackley, S. A. (2010). Visual working memory deficits in patients with Parkinson’s disease are due to both reduced storage capacity and impaired ability to filter out irrelevant information. *Brain*, *133*(9), 2677–2689. https://doi.org/10.1093/brain/awq197

Lee, J., & Park, S. (2005). Working memory impairments in schizophrenia: a meta-analysis. *Journal of Abnormal Psychology*, *114*(4), 599–611. https://doi.org/10.1037/0021-843X.114.4.599

Lee, M. D., and Wagenmakers, E. J. (2014). *Bayesian Cognitive Modeling: A Practical Course*. Cambridge: Cambridge University Press.

Lemaire, B., & Portrat, S. (2018). A Computational Model of Working Memory Integrating Time-Based Decay and Interference. *Frontiers in Psychology*, *9*, 416. https://doi.org/10.3389/fpsyg.2018.00416

Leppink, J., O’Sullivan, P., & Winston, K. (2017). Evidence against vs. in favour of a null hypothesis. *Perspectives on Medical Education*, *6*(2), 115–118. https://doi.org/10.1007/s40037-017-0332-6

Lewandowsky, S., Oberauer, K., & Brown, G. D. A. (2009). No temporal decay in verbal short-term memory. *Trends in Cognitive Sciences*, *13*(3), 120–126. https://doi.org/10.1016/j.tics.2008.12.003

Lewis-Peacock, J. A., Kessler, Y., & Oberauer, K. (2018). The removal of information from working memory. *Annals of the New York Academy of Sciences*, *1424*(1), 33–44. https://doi.org/10.1111/nyas.13714

Li, Q., & Saiki, J. (2014). The effects of sequential attention shifts within visual working memory. *Frontiers in Psychology*, *5*, 965. https://doi.org/10.3389/fpsyg.2014.00965

Light, P. H., Buckingham, N., & Robbins, A. H. (1979). The conservation task as an interactional setting. *British Journal of Educational Psychology*, *49*(3), 304–310.

Loaiza, V. M., & Souza, A. S. (2018). Is refreshing in working memory impaired in older age? Evidence from the retro-cue paradigm. *Annals of the New York Academy of Sciences*, *1424*(1), 175–189. https://doi.org/10.1111/nyas.13623

Locke, J. (1841). *An Essay Concerning Human Understanding*.

Logie, R. H. (1995). Visuo-spatial working memory Hove. *UK: Lawrence Erlbaum Associates Ltd*.

Logie, R. H. (1996). The seven ages of working memory. *Working Memory and Human Cognition*, 31–65.

Logie, R. H. (2011). The Functional Organization and Capacity Limits of Working Memory. *Current Directions in Psychological Science*, *20*(4), 240–245. https://doi.org/10.1177/0963721411415340

Logie, R. H. (2016). Retiring the central executive. *Quarterly Journal of Experimental Psychology*, *69*(10), 2093–2109. https://doi.org/10.1080/17470218.2015.1136657

Logie, R. H., & Cowan, N. (2015). Perspectives on working memory: introduction to the special issue. *Memory & Cognition*, *43*(3), 315–324. https://doi.org/10.3758/s13421-015-0510-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

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

Macken, B., Taylor, J., & Jones, D. (2015). Limitless capacity: a dynamic object-oriented approach to short-term memory. *Frontiers in Psychology*, *6*, 293. https://doi.org/10.3389/fpsyg.2015.00293

Macken, W. J., Phelps, F. G., & Jones, D. M. (2009). What causes auditory distraction? *Psychonomic Bulletin & Review*, *16*(1), 139–144. https://doi.org/10.3758/PBR.16.1.139

Makovsik, T., & Jiang, Y. V. (2007). Distributing versus focusing attention in visual short-term memory. *Psychonomic Bulletin & Review*, *14*(6), 1072–1078. https://doi.org/10.3758/bf03193093

Makovski, T. (2012). Are multiple visual short-term memory storages necessary to explain the retro-cue effect? *Psychonomic Bulletin & Review*, *19*(3), 470–476. https://doi.org/10.3758/s13423-012-0235-9

Makovski, T., Sussman, R., & Jiang, Y. V. (2008). Orienting attention in visual working memory reduces interference from memory probes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(2), 369–380. https://doi.org/10.1037/0278-7393.34.2.369

Maldonado, M., Dunbar, E., & Chemla, E. (2019). Mouse tracking as a window into decision making. *Behavior Research Methods*, *51*(3), 1085–1101. https://doi.org/10.3758/s13428-018-01194-x

Marshall, L., & Born, J. (2007). The contribution of sleep to hippocampus-dependent memory consolidation. *Trends in Cognitive Sciences*, *11*(10), 442–450. https://doi.org/10.1016/j.tics.2007.09.001

Martinussen, R., Hayden, J., Hogg-Johnson, S., & Tannock, R. (2005). A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *44*(4), 377–384. https://doi.org/10.1097/01.chi.0000153228.72591.73

Matsukura, M., Luck, S. J., & Vecera, S. P. (2007). Attention effects during visual short-term memory maintenance: protection or prioritization? *Perception & Psychophysics*, *69*(8), 1422–1434. https://doi.org/10.3758/bf03192957

McElree, B. (2001). Working memory and focal attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(3), 817–835. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11394682

McElree, B. (2006). Accessing Recent Events. In *Psychology of Learning and Motivation* (Vol. 46, pp. 155–200). https://doi.org/10.1016/S0079-7421(06)46005-9

McElree, B., & Dosher, B. A. (1989). Serial position and set size in short-term memory: The time course of recognition. *Journal of Experimental Psychology: General*, *118*(4), 346–373. https://doi.org/10.1037/0096-3445.118.4.346

McKelvie, S. J. (1987). Learning and awareness in the Hebb digits task. *The Journal of General Psychology, 114*(1), 75-88. https://doi.org/10.1080/00221309.1987.9711057

McNamara, D. S., & Scott, J. L. (2001). Working memory capacity and strategy use. *Memory & Cognition*, *29*(1), 10–17. https://doi.org/10.3758/bf03195736

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

Middlebrooks, C. D., Kerr, T., & Castel, A. D. (2017). Selectively Distracted: Divided Attention and Memory for Important Information. *Psychological Science*, *28*(8), 1103–1115. https://doi.org/10.1177/0956797617702502

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.

Ministry of Housing, Communities & Local Governments. (2015). English indices of deprivation 2015. Retrieved from https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015

Mok, R. M., Myers, N. E., Wallis, G., & Nobre, A. C. (2016). Behavioral and Neural Markers of Flexible Attention over Working Memory in Aging. *Cerebral Cortex* , *26*(4), 1831–1842. https://doi.org/10.1093/cercor/bhw011

Morales, J., Calvo, A., & Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *Journal of Experimental Child Psychology*, *114*(2), 187–202. https://doi.org/10.1016/j.jecp.2012.09.002

Morey, C. C. (2018). The case against specialized visual-spatial short-term memory. *Psychological Bulletin*, *144*(8), 849–883. https://doi.org/10.1037/bul0000155

Morey, C. C., Morey, R. D., van der Reijden, M., & Holweg, M. (2013). Asymmetric cross-domain interference between two working memory tasks: Implications for models of working memory. *Journal of Memory and Language*, *69*(3), 324–348. https://doi.org/10.1016/j.jml.2013.04.004

Morey, R. D., & Rouder, J. N. (2011). Bayes factor approaches for testing interval null hypotheses. *Psychological Methods*, *16*(4), 406–419. https://doi.org/10.1037/a0024377

Morrison, A. B., & Chein, J. M. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychonomic Bulletin & Review*, *18*(1), 46–60. https://doi.org/10.3758/s13423-010-0034-0

Morrison, A. B., Rosenbaum, G. M., Fair, D., & Chein, J. M. (2016). Variation in strategy use across measures of verbal working memory. *Memory & Cognition*, *44*(6), 922–936. https://doi.org/10.3758/s13421-016-0608-9

Murayama, K., & Kitagami, S. (2014). Consolidation power of extrinsic rewards: reward cues enhance long-term memory for irrelevant past events. *Journal of*

Experimental Psychology: General, *143*(1), 15–20. https://doi.org/10.1037/a0031992

Murayama, K., & Kuhbandner, C. (2011). Money enhances memory consolidation--But only for boring material. *Cognition*, *119*(1), 120–124.

Murayama, K., Miyatsu, T., Buchli, D., & Storm, B. C. (2014). Forgetting as a consequence of retrieval: a meta-analytic review of retrieval-induced forgetting. *Psychological Bulletin*, *140*(5), 1383–1409. https://doi.org/10.1037/a0037505

Murty, V. P., & Adcock, R. A. (2014). Enriched encoding: reward motivation organizes cortical networks for hippocampal detection of unexpected events. *Cerebral Cortex* , *24*(8), 2160–2168. https://doi.org/10.1093/cercor/bht063

Myers, N. E., Chekroud, S. R., Stokes, M. G., & Nobre, A. C. (2018). Benefits of flexible prioritization in working memory can arise without costs. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(3), 398–411. https://doi.org/10.1037/xhp0000449

Myers, N. E., Stokes, M. G., & Nobre, A. C. (2017). Prioritizing Information during Working Memory: Beyond Sustained Internal Attention. *Trends in Cognitive Sciences*, *21*(6), 449–461. https://doi.org/10.1016/j.tics.2017.03.010

Naveh-Benjamin, M., & Jonides, J. (1985). The effects of rehearsal on frequency coding. *Bulletin of the Psychonomic Society*, *23*(4), 387–390. https://doi.org/10.3758/BF03330192

Newsome, R. N., Duarte, A., Pun, C., Smith, V. M., Ferber, S., & Barense, M. D. (2015). A retroactive spatial cue improved VSTM capacity in mild cognitive impairment and medial temporal lobe amnesia but not in healthy older adults.

Neuropsychologia, *77*, 148–157. https://doi.org/10.1016/j.neuropsychologia.2015.08.017

Nguyen, L. T., Marini, F., Zacharczuk, L., Llano, D. A., & Mudar, R. A. (2019). Theta and alpha band oscillations during value-directed strategic processing. *Behavioural Brain Research*, *367*, 210–214. https://doi.org/10.1016/j.bbr.2019.03.052

Nieuwenstein, M., & Wyble, B. (2014). Beyond a mask and against the bottleneck: retroactive dual-task interference during working memory consolidation of a masked visual target. *Journal of Experimental Psychology: General*, *143*(3), 1409–1427. https://doi.org/10.1037/a0035257

Niklaus, M., Singmann, H., & Oberauer, K. (2019). Two distinct mechanisms of selection in working memory: Additive last-item and retro-cue benefits. *Cognition*, *183*, 282–302. https://doi.org/10.1016/j.cognition.2018.11.01

Nissen, M. J., Willingham, D., & Hartman, M. (1989). Explicit and implicit remembering: when is learning preserved in amnesia? *Neuropsychologia*, *27*(3), 341–352. https://doi.org/10.1016/0028-3932(89)90023-7

Noble, K. G., McCandliss, B. D., & Farah, M. J. (2007). Socioeconomic gradients predict individual differences in neurocognitive abilities. *Developmental Science*, *10*(4), 464–480. https://doi.org/10.1111/j.1467-7687.2007.00600.x

Noble, K. G., Norman, M. F., & Farah, M. J. (2005). Neurocognitive correlates of socioeconomic status in kindergarten children. *Developmental Science*, *8*(1), 74–87. https://doi.org/10.1111/j.1467-7687.2005.00394.x

Nobre, A. C., Griffin, I. C., & Rao, A. (2007). Spatial attention can bias search in visual short-term memory. *Frontiers in Human Neuroscience*, *1*, 4. https://doi.org/10.3389/neuro.09.004.2007

Nungester, R. J., & Duchastel, P. C. (1982). Testing versus review: Effects on retention. *Journal of Educational Psychology*, *74*(1), 18. Retrieved from https://psycnet.apa.org/journals/edu/74/1/18/

Oberauer, K. (2002). Access to information in working memory: exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(3), 411–421.

Oberauer, K., & Hein, L. (2012). Attention to Information in Working Memory. *Current Directions in Psychological Science*, *21*(3), 164–169. https://doi.org/10.1177/0963721412444727

Oberauer, K., Jones, T., & Lewandowsky, S. (2015). The Hebb repetition effect in simple and complex memory span. *Memory & Cognition*, *43*(6), 852-865. 10.3758/s13421-015-0512-8.

Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D. A., Conway, A., Cowan, N., … Ward, G. (2018). Benchmarks for models of short-term and working memory. *Psychological Bulletin*, *144*(9), 885–958. https://doi.org/10.1037/bul0000153

Oberauer, K., Stoneking, C., Wabersich, D., & Lin, H.-Y. (2017). Hierarchical Bayesian measurement models for continuous reproduction of visual features from working memory. *Journal of Vision*, *17*(5), 11. https://doi.org/10.1167/17.5.11

Owen, A. M., Iddon, J. L., Hodges, J. R., Summers, B. A., & Robbins, T. W. (1997). Spatial and non-spatial working memory at different stages of Parkinson's disease. *Neuropsychologia*, *35*(4), 519–532. https://doi.org/10.1016/s0028-3932(96)00101-7

Page, M., Cumming, N., Norris, D., Hitch, G. J., & McNeil, A. M. (2006). Repetition learning in the immediate serial recall of visual and auditory materials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(4), 716-733. https://doi.org/10.1037/0278-7393.32.4.716

Pashler, H., Johnston, J. C., & Ruthruff, E. (2001). Attention and performance. *Annual Review of Psychology*, *52*, 629–651. https://doi.org/10.1146/annurev.psych.52.1.629

Pertzov, Y., Bays, P. M., Joseph, S., & Husain, M. (2013). Rapid forgetting prevented by retrospective attention cues. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(5), 1224–1231. https://doi.org/10.1037/a0030947

Piaget, J., and Inhelder, B. (1967). The Child's Conception of Space. New York, NY: Norton & Co.

Pinto, Y., van der Leij, A. R., Sligte, I. G., Lamme, V. A. F., & Scholte, H. S. (2013). Bottom-up and top-down attention are independent. *Journal of Vision*, *13*(3), 16. https://doi.org/10.1167/13.3.16

Portrat, S., Camos, V., & Barrouillet, P. (2009). Working memory in children: a time-constrained functioning similar to adults. *Journal of Experimental Child Psychology*, *102*(3), 368–374. https://doi.org/10.1016/j.jecp.2008.05.005

Quatto, P., Ripamonti, E., & Marasini, D. (2019). Best uses of p-values and complementary measures in medical research: Recent developments in the frequentist and Bayesian frameworks. *Journal of Biopharmaceutical Statistics*, 1–22. https://doi.org/10.1080/10543406.2019.1632874

R Core Team. (2016). *R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2016. R Foundation for Statistical Computing*.

Morey, R. D., Rouder, J. N., Jamil, T., Urbanek, S., Forner, K., & Ly, A. (2018). Package ‘BayesFactor’ (Version 0.9.12-4.2. Retrieved from https://cran.r-project.org/web/packages/BayesFactor/BayesFactor.pdf

Reaves, S., Strunk, J., Phillips, S., Verhaeghen, P., & Duarte, A. (2016). The lasting memory enhancements of retrospective attention. *Brain Research*, *1642*, 226–237. https://doi.org/10.1016/j.brainres.2016.03.048

Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., … Engle, R. W. (2013). No evidence of intelligence improvement after working memory training: a randomized, placebo-controlled study. *Journal of Experimental Psychology: General*, *142*(2), 359–379. https://doi.org/10.1037/a0029082

Rerko, L., & Oberauer, K. (2013). Focused, unfocused, and defocused information in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(4), 1075–1096. https://doi.org/10.1037/a0031172

Rerko, L., Souza, A. S., & Oberauer, K. (2014). Retro-cue benefits in working memory without sustained focal attention. *Memory & Cognition*, *42*(5), 712–728. https://doi.org/10.3758/s13421-013-0392-8

Richardson, J. T. E., & Baddeley, A. D. (1975). The effect of articulatory suppression in free recall. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 623–629. https://doi.org/10.1016/S0022-5371(75)80049-1

Rickard, T. C., & Pan, S. C. (2018). A dual memory theory of the testing effect. *Psychonomic Bulletin & Review*, *25*(3), 847–869. https://doi.org/10.3758/s13423-017-1298-4

Ricker, T. J., & Cowan, N. (2014). Differences between presentation methods in working memory procedures: a matter of working memory consolidation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(2), 417–428. https://doi.org/10.1037/a0034301

Ricker, T. J., Nieuwenstein, M. R., Bayliss, D. M., & Barrouillet, P. (2018). Working memory consolidation: insights from studies on attention and working memory. *Annals of the New York Academy of Sciences*, *1424*(1), 8–18.

Ricker, T. J., & Sandry, J. (2018). The relationship between masking and short-term consolidation during recall from visual working memory. *Annals of the New York Academy of Sciences*, *1424*(1), 91–101.

Ristic, J., & Kingstone, A. (2006). Attention to arrows: pointing to a new direction. *Quarterly Journal of Experimental Psychology*, *59*(11), 1921–1930. https://doi.org/10.1080/17470210500416367

Robison, M. K., & Unsworth, N. (2017). Working memory capacity, strategic allocation of study time, and value-directed remembering. *Journal of Memory and Language*, *93*, 231–244. https://doi.org/10.1016/j.jml.2016.10.007

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374. https://doi.org/10.1016/j.jmp.2012.08.001

Rouder, J. N., Morey, R. D., Verhagen, J., Swagman, A. R., & Wagenmakers, E. J. (2017). Bayesian analysis of factorial designs. *Psychological Methods*, *22*(2), 304.

Rowland, C. A. (2014). The effect of testing versus restudy on retention: a meta-analytic review of the testing effect. *Psychological Bulletin*, *140*(6), 1432–1463. https://doi.org/10.1037/a0037559

Ruiz-Ruano García, A. M., & López Puga, J. (2018). Deciding on Null Hypotheses using P-values or Bayesian alternatives: A simulation study. *Psicothema*, *30*(1), 110–115. https://doi.org/10.7334/psicothema2017.308

Sala, S. D., Darling, S., & Logie, R. H. (2010). Items on the left are better remembered. *The Quarterly Journal of Experimental Psychology*, *63*(5), 848–855. https://doi.org/10.1080/17470211003690672

Sandry, J., Schwark, J. D., & MacDonald, J. (2014). Flexibility within working memory and the focus of attention for sequential verbal information does not depend on active maintenance. *Memory & Cognition*, *42*(7), 1130–1142. https://doi.org/10.3758/s13421-014-0422-1

Schendel, Z. A., & Palmer, C. (2007). Suppression effects on musical and verbal memory. *Memory & Cognition*, *35*(4), 640–650. https://doi.org/10.3758/bf03193302

Schmidt, B. K., Vogel, E. K., Woodman, G. F., & Luck, S. J. (2002). Voluntary and automatic attentional control of visual working memory. *Perception & Psychophysics*, *64*(5), 754–763. https://doi.org/10.3758/BF03194742

Seigneuric, A., & Ehrlich, M. F. (2005). Contribution of working memory capacity to children’s reading comprehension: A longitudinal investigation. *Reading and Writing*, *18*(7-9), 617-656.

Seigneuric, A., Ehrlich, M.-F., Oakhill, J. V., & Yuill, N. M. (2000). Working memory resources and children's reading comprehension. *Reading and Writing*, *13*(1), 81–103. https://doi.org/10.1023/A:1008088230941

Shepherdson, P., Oberauer, K., & Souza, A. S. (2018). Working memory load and the retro-cue effect: A diffusion model account. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(2), 286–310. https://doi.org/10.1037/xhp0000448

Shigemune, Y., Abe, N., Suzuki, M., Ueno, A., Mori, E., Tashiro, M., … Fujii, T. (2010). Effects of emotion and reward motivation on neural correlates of episodic memory encoding: a PET study. *Neuroscience Research*, *67*(1), 72–79. https://doi.org/10.1016/j.neures.2010.01.003

Shimi, A., Nobre, A. C., Astle, D., & Scerif, G. (2014). Orienting attention within visual short-term memory: Development and mechanisms. *Child Development*, *85*(2), 578–592.

Shimi, A., & Scerif, G. (2015). The interplay of spatial attentional biases and mental codes in VSTM: Developmentally informed hypotheses. *Developmental Psychology*, *51*(6), 731–743. https://doi.org/10.1037/a0039057

Shimi, A., & Scerif, G. (2017). Towards an integrative model of visual short-term memory maintenance: Evidence from the effects of attentional control, load, decay, and their interactions in childhood. *Cognition*, *169*, 61–83. https://doi.org/10.1016/j.cognition.2017.08.005

Shin, H., Bjorklund, D. F., & Beck, E. F. (2007). The adaptive nature of children's overestimation in a strategic memory task. *Cognitive Development*, *22*(2), 197–212. https://doi.org/10.1016/j.cogdev.2006.10.001

Shipstead, Z., Redick, T. S., & Engle, R. W. (2010). Does working memory training generalize? *Psychologica Belgica*, *50*(3), 245–276.

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

Shohamy, D., & Adcock, R. A. (2010). Dopamine and adaptive memory. *Trends in Cognitive Sciences*, *14*(10), 464–472. https://doi.org/10.1016/j.tics.2010.08.002

Siegert, R. J., Weatherall, M., Taylor, K. D., & Abernethy, D. A. (2008). A meta-analysis of performance on simple span and more complex working memory tasks in Parkinson's disease. *Neuropsychology*, *22*(4), 450–461. https://doi.org/10.1037/0894-4105.22.4.450

Souza, A. S., & Oberauer, K. (2015). Time-based forgetting in visual working memory reflects temporal distinctiveness, not decay. *Psychonomic Bulletin & Review*, *22*(1), 156–162. https://doi.org/10.3758/s13423-014-0652-z

Souza, A. S., & Oberauer, K. (2016). In search of the focus of attention in working memory: 13 years of the retro-cue effect. *Attention, Perception & Psychophysics*, *78*(7), 1839–1860. https://doi.org/10.3758/s13414-016-1108-5

Souza, A. S., Rerko, L., & Oberauer, K. (2014). Unloading and reloading working memory: attending to one item frees capacity. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 1237–1256. https://doi.org/10.1037/a0036331

Souza, A. S., Rerko, L., & Oberauer, K. (2016). Getting more from visual working memory: Retro-cues enhance retrieval and protect from visual interference.

Journal of Experimental Psychology: Human Perception and Performance, *42*(6), 890–910. https://doi.org/10.1037/xhp0000192

St Clair-Thompson, H. L., & Allen, R. J. (2013). Are forward and backward recall the same? A dual-task study of digit recall. *Memory & Cognition*, *41*(4), 519–532. https://doi.org/10.3758/s13421-012-0277-2

St Clair-Thompson, H., Stevens, R., Hunt, A., & Bolder, E. (2010). Improving children's working memory and classroom performance. *Educational Psychology Review*, *30*(2), 203–219. https://doi.org/10.1080/01443410903509259

Strunk, J., Morgan, L., Reaves, S., Verhaeghen, P., & Duarte, A. (2018). Retrospective attention in short-term memory has a lasting effect on long-term memory across age. *The Journals of Gerontology. Series B, 74*(8), 1317-1325.

Süß, H.-M., Oberauer, K., Wittmann, W. W., Wilhelm, O., & Schulze, R. (2002). Working-memory capacity explains reasoning ability—and a little bit more. *Intelligence*, *30*(3), 261–288. https://doi.org/10.1016/S0160-2896(01)00100-3

Theeuwes, J. (2018). Visual Selection: Usually Fast and Automatic; Seldom Slow and Volitional. *Journal of Cognition*, *1*(1), 29. https://doi.org/10.5334/joc.13

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

Torralva, T., Roca, M., Gleichgerrcht, E., Bekinschtein, T., & Manes, F. (2009). A neuropsychological battery to detect specific executive and social cognitive impairments in early frontotemporal dementia. *Brain*, *132*(5), 1299–1309. https://doi.org/10.1093/brain/awp041

Turley-Ames, K. J., & Whitfield, M. M. (2003). Strategy training and working memory task performance. *Journal of Memory and Language*, *49*(4), 446–468. https://doi.org/10.1016/S0749-596X(03)00095-0

van Moorselaar, D., Gunseli, E., Theeuwes, J., & N L Olivers, C. (2014). The time course of protecting a visual memory representation from perceptual interference. *Frontiers in Human Neuroscience*, *8*, 1053. https://doi.org/10.3389/fnhum.2014.01053

van Moorselaar, D., Olivers, C. N. L., Theeuwes, J., Lamme, V. A. F., & Sligte, I. G. (2015). Forgotten but not gone: Retro-cue costs and benefits in a double-cueing paradigm suggest multiple states in visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(6), 1755–1763. https://doi.org/10.1037/xlm0000124

Vergauwe, E., & Cowan, N. (2015). Attending to items in working memory: evidence that refreshing and memory search are closely related. *Psychonomic Bulletin & Review*, *22*(4), 1001–1006. https://doi.org/10.3758/s13423-014-0755-6

Vergauwe, E., & Langerock, N. (2017). Attentional refreshing of information in working memory: Increased immediate accessibility of just-refreshed representations. *Journal of Memory and Language*, *96*, 23–35. https://doi.org/10.1016/j.jml.2017.05.001

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., Woodman, G. F., & Luck, S. J. (2005). Pushing around the locus of selection: evidence for the flexible-selection hypothesis. *Journal of Cognitive Neuroscience*, *17*(12), 1907–1922. https://doi.org/10.1162/089892905775008599

von Restorff, H. (1933). Über die Wirkung von Bereichsbildungen im Spurenfeld. *Psychologische Forschung*, *18*(1), 299–342. https://doi.org/10.1007/BF02409636

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., … Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, *25*(1), 35–57. https://doi.org/10.3758/s13423-017-1343-3

Wang, B., Yan, C., Wang, Z., Olivers, C. N. L., & Theeuwes, J. (2017). Adverse orienting effects on visual working memory encoding and maintenance. *Psychonomic Bulletin & Review*, *24*(4), 1261–1267. https://doi.org/10.3758/s13423-016-1205-4

Wasserstein, R. L., Lazar, N. A., & Others. (2016). The ASA's statement on p-values: context, process, and purpose. *The American Statistician*, *70*(2), 129–133.

Waszak, F., Li, S.-C., & Hommel, B. (2010). The development of attentional networks: cross-sectional findings from a life span sample. *Developmental Psychology*, *46*(2), 337–349. https://doi.org/10.1037/a0018541

Watanabe, N. (2017). Acquiring Piaget's conservation concept of numbers, lengths and liquids as ordinary play. *Journal of Educational and Developmental Psychology*, *7*(1), 210–217.

Waterman, A. H., Atkinson, A. L., Aslam, S. S., Holmes, J., Jaroslawska, A., & Allen, R. J. (2017). Do actions speak louder than words? Examining children's ability to follow instructions. *Memory & Cognition*, *45*(6), 877–890. https://doi.org/10.3758/s13421-017-0702-7

Watkins, M. J., & Bloom, L. C. (1999). *Selectivity in memory: An exploration of willful control over the remembering process.* Unpublished Manuscript.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale*. Psychological Corporation.

Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E.-J. (2011). Statistical Evidence in Experimental Psychology: An Empirical Comparison Using 855 t Tests. *Perspectives on Psychological Science*, *6*(3), 291–298. https://doi.org/10.1177/1745691611406923

Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, *131*(1), 48–64. https://doi.org/10.1037//0096-3445.131.1.48

Wittmann, B. C., Schott, B. H., Guderian, S., Frey, J. U., Heinze, H.-J., & Düzel, E. (2005). Reward-related FMRI activation of dopaminergic midbrain is associated with enhanced hippocampus-dependent long-term memory formation. *Neuron*, *45*(3), 459–467. https://doi.org/10.1016/j.neuron.2005.01.010

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

Wundt, W. (1948). Principles of physiological psychology. In W. Dennis (Ed.), *Readings in the history of psychology*. New York, NY: Appleton-Century-Crofts. Original work published 1873.

Yantis, S. (2000). Goal-directed and stimulus-driven determinants of attentional control. *Attention and Performance*, *18*, 73–103.

Yi, Y., & Friedman, D. (2014). Age-related differences in working memory: ERPs reveal age-related delays in selection- and inhibition-related processes. *Aging, Neuropsychology and Cognition*, *21*(4), 483–513. https://doi.org/10.1080/13825585.2013.833581

Yussen, S. R., & Levy, V. M. (1975). Developmental changes in predicting one's own span of short-term memory. *Journal of Experimental Child Psychology*, *19*(3), 502–508. https://doi.org/10.1016/0022-0965(75)90079-X

Zacks, R. T., & Hasher, L. (2002). Frequency processing: A twenty-five year perspective. *Frequency Processing and Cognition*, 21–36.

Zacks, R. T., Hasher, L., & Sanft, H. (1982). Automatic encoding of event frequency: Further findings. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*(2), 106-116. https://doi.org/10.1037//0278-7393.8.2.106

Zgonnikov, A., Aleni, A., Piiroinen, P. T., O'Hora, D., & di Bernardo, M. (2017). Decision landscapes: visualizing mouse-tracking data. *Royal Society Open Science*, *4*(11), 170482. https://doi.org/10.1098/rsos.170482

APPENDICES

Appendix A: Questionnaire used in Experiment 3 (Chapter 3)

Part 1

The following question refers to the condition where you were told: "Try to remember as many digits as you can. For each digit you correctly recall, you will be given 1 point":

Did you use a strategy to try remember the digits in this condition? If so, what was it?

Part 2

The following questions refer to the condition where you were told: "Try to remember as many digits as you can, but try extra hard to remember the 3rd digit you hear. You will receive 4 points if you correctly recall this digit. For all other digits you recall, you will receive 1 point."

Did you use a strategy to try remember the digits in this condition? If so, what was it?

How do you think trying to Prioritize the 3rd position affected your memory for this digit compared to the condition where every digit was worth 1 point? (1 = large negative effect, 5 = no effect, 9 = large positive effect)

1	2	3	4	5	6	7	8	9

How do you think trying to Prioritize the 3rd position affected your memory for the other digits compared to the condition where every digit was worth 1 point? (1 = large negative effect, 5 = no effect, 9 = large positive effect)

1	2	3	4	5	6	7	8	9

Part 3

The following questions refer to the condition where you were told: Try to remember as many digits as you can, but try extra hard to remember the 5th digit you hear. You will receive 4 points if you correctly recall this digit. For all other digits you recall, you will receive 1 point."

Did you use a strategy to try remember the digits in this condition? If so, what was it?

How do you think trying to Prioritize the 5th position affected your memory for this digit compared to the condition where every digit was worth 1 point? (1 = large negative effect, 5 = no effect, 9 = large positive effect)

1	2	3	4	5	6	7	8	9

How do you think trying to Prioritize the 5th position affected your memory for the other digits compared to the condition where every digit was worth 1 point? (1 = large negative effect, 5 = no effect, 9 = large positive effect)

1	2	3	4	5	6	7	8	9

Part 4

The following questions refer to the condition where you were told: Try to remember as many digits as you can, but try extra hard to remember the 7th digit you hear. You will receive 4 points if you correctly recall this digit. For all other digits you recall, you will receive 1 point."

Did you use a strategy to try remember the digits in this condition? If so, what was it?

How do you think trying to Prioritize the 7th position affected your memory for this digit compared to the condition where every digit was worth 1 point? (1 = large negative effect, 5 = no effect, 9 = large positive effect)

1	2	3	4	5	6	7	8	9

How do you think trying to Prioritize the 7th position affected your memory for the other digits compared to the condition where every digit was worth 1 point? (1 = large negative effect, 5 = no effect, 9 = large positive effect)

1	2	3	4	5	6	7	8	9

Do you have any further comments about the experiment/any of the conditions?

Appendix B: Questionnaire used in Experiment 4 (Chapter 3)

Part 1
The following question refers to the condition you were told**: "Try to remember as many digits as you can. For each digit you correctly recall, you will be given 1 point"** and **you weren't asked to repeat anything during the presentation phase**:

Did you use a strategy to try remember the digits in this condition? If so, what was it?

Part 2
The following questions refer to the condition where you were told: **Try to remember as many digits as you can, but try extra hard to remember the 5th digit you hear. You will receive 4 points if you correctly recall this digit. For all other digits you recall, you will receive 1 point"** and **you weren't asked to repeat anything during the presentation phase**:

Did you use a strategy to try remember the digits in this condition? If so, what was it?

How do you think trying to Prioritize the 5th position affected your memory for this digit compared to the condition where every digit was worth 1 point and you weren't asked to repeat anything during the presentation phase? (1 = large negative effect, 5 = no effect, 9 = large positive effect)

1	2	3	4	5	6	7	8	9

How do you think trying to Prioritize the 5th position affected your memory for the other digits compared to the condition where every digit was worth 1 point and you weren't asked to repeat anything during the presentation phase? (1 = large negative effect, 5 = no effect, 9 = large positive effect)

1	2	3	4	5	6	7	8	9

Part 3
The following question refers to the condition you were told**: "Try to remember as many digits as you can. For each digit you correctly recall, you will be given 1 point"** and **you were asked to repeat a month of the year and a day of the week during the presentation phase:**

Did you use a strategy to try remember the digits in this condition? If so, what was it?

Part 4
The following questions refer to the condition where you were told: **Try to remember as many digits as you can, but try extra hard to remember the 5th**

digit you hear. You will receive 4 points if you correctly recall this digit. For all other digits you recall, you will receive 1 point” and **you were asked to repeat a day of the week and a month of the year during the presentation phase:**

Did you use a strategy to try remember the digits in this condition? If so, what was it?

How do you think trying to Prioritize the 5th position affected your memory for this digit compared to the condition where every digit was worth 1 point and you had to repeat a day of the week and a month of the year during the presentation phase? (1 = large negative effect, 5 = no effect, 9 = large positive effect)

1	2	3	4	5	6	7	8	9

How do you think trying to Prioritize the 5th position affected your memory for the other digits compared to the condition where every digit was worth 1 point and you had to repeat a day of the week and a month of the year during the presentation phase? (1 = large negative effect, 5 = no effect, 9 = large positive effect)

1	2	3	4	5	6	7	8	9

Do you have any further comments about the experiment/any of the conditions?

Appendix C: Questionnaire used in Experiment 5 (Chapter 3)

Part 1

The following question refers to the condition you were told**: "Try to remember as many digits as you can. For each digit you correctly recall, you will be given 1 point"** and **you were asked to repeat a day of the week and a month of the year:**

Did you use a strategy to try remember the digits in this condition? If so, what was it?

Part 2

The following questions refer to the condition where you were told: **Try to remember as many digits as you can, but try extra hard to remember the 4th digit you hear. You will receive 4 points if you correctly recall this digit. For all other digits you recall, you will receive 1 point"** and **you were asked to repeat a day of the week and a month of the year:**

Did you use a strategy to try remember the digits in this condition? If so, what was it?

How do you think trying to prioritise the 4th position affected your memory for this digit compared to the condition where every digit was worth 1 point and you had to repeat a day of the week and a month of the year? (1 = large negative effect, 5 = no effect, 9 = large positive effect)

1	2	3	4	5	6	7	8	9

How do you think trying to prioritise the 4th position affected your memory for the other digits compared to the condition where every digit was worth 1 point and you had to repeat a day of the week and a month of the year? (1 = large negative effect, 5 = no effect, 9 = large positive effect)

1	2	3	4	5	6	7	8	9

Do you have any further comments about the experiment/any of the conditions?

Part 3

The following question refers to the condition you were told**: "Try to remember as many digits as you can. For each digit you correctly recall, you will be given 1 point"** and **you were asked to alternate between the days of the week and the month of the year:**

Did you use a strategy to try remember the digits in this condition? If so, what was it?

Part 4

The following questions refer to the condition where you were told: **Try to remember as many digits as you can, but try extra hard to remember the 4th digit you hear. You will receive 4 points if you correctly recall this digit. For all other digits you recall, you will receive 1 point"** and **you were asked to alternate between the days of the week and the month of the year:**

Did you use a strategy to try remember the digits in this condition? If so, what was it?

How do you think trying to prioritise the 4th position affected your memory for this digit compared to the condition where every digit was worth 1 point and you had to alternate between the day of the week and a month of the year? (1 = large negative effect, 5 = no effect, 9 = large positive effect)

1	2	3	4	5	6	7	8	9

How do you think trying to prioritise the 4th position affected your memory for the other digits compared to the condition where every digit was worth 1 point and you had to alternate between the day of the week and a month of the year? (1 = large negative effect, 5 = no effect, 9 = large positive effect)

1	2	3	4	5	6	7	8	9

Do you have any further comments about the experiment/any of the conditions?

Appendix D: The alien story used in Experiments 6 and 7 (Chapter 4)

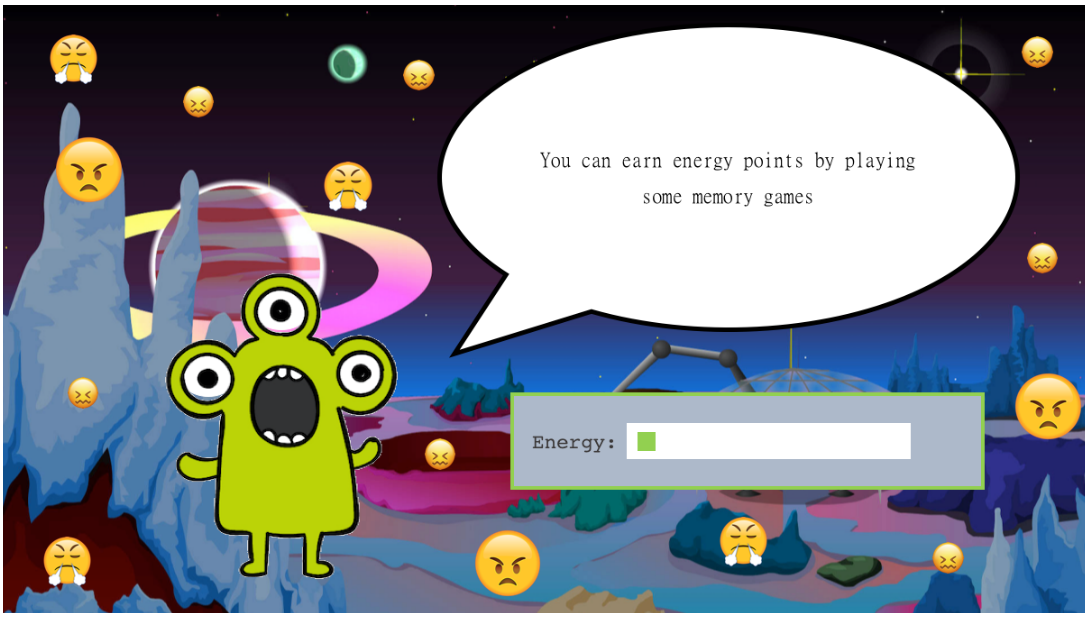
Story before the main memory task:


They have been causing lots of trouble.
Will you help me zap them?


To zap the aliens, you need energy.
How much energy do you have?

Oh no!

You are running low!!



You can earn energy points by playing some memory games


The more energy points you collect during these memory games, the more aliens you will be able to zap


If you get enough energy points, you will also be given a prize to take home!

Make sure you listen to find out how to earn energy points!

Screens used to inform children how many points they had collected:

3 item sequences:

Great job! You got lots of energy points!

Let's get some more!

Remember, the more energy you collect, the more aliens you will be able to zap

Great job! You got lots of energy points!

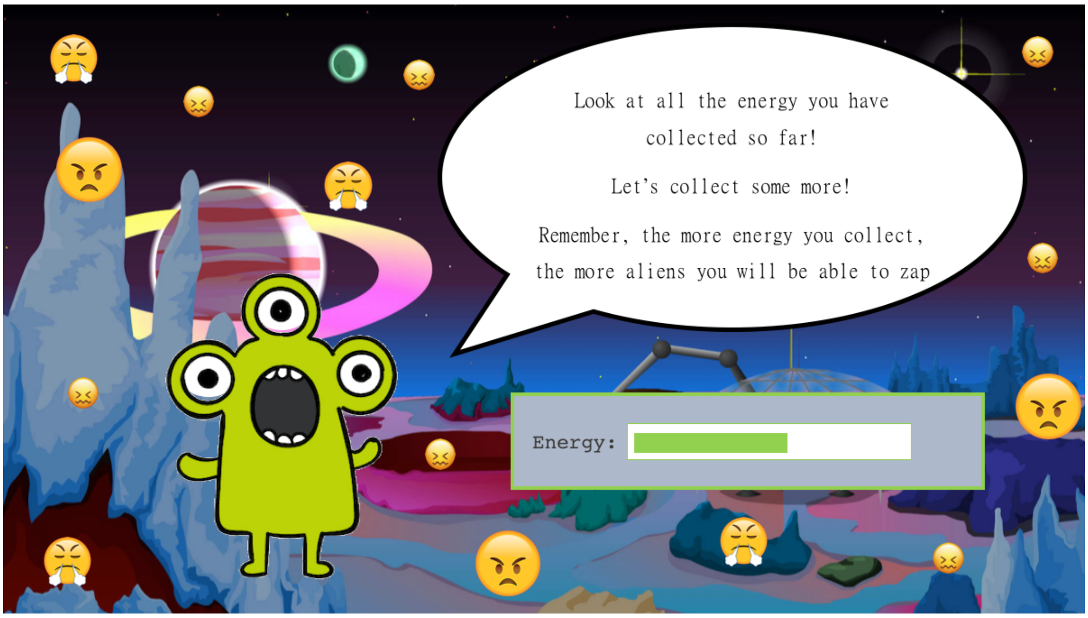
Let's get some more!

Remember, the more energy you collect, the more aliens you will be able to zap

Look at all the energy you have collected so far!

Let's collect some more!

Remember, the more energy you collect, the more aliens you will be able to zap



Look at all the energy you have collected so far!

Let's collect some more!

Remember, the more energy you collect, the more aliens you will be able to zap

Wow! That energy bar is getting full!

Can you collect a little more?

Remember, the more energy you collect, the more aliens you will be able to zap

4 item sequences:

Great job! You got lots of energy points!

Let’s get some more!

Remember, the more energy you collect, the more aliens you will be able to zap

Great job! You got lots of energy points!

Let’s get some more!

Remember, the more energy you collect, the more aliens you will be able to zap

Great job! You got lots of energy points!

Let's get some more!

Remember, the more energy you collect, the more aliens you will be able to zap

Look at all the energy you have collected so far!

Let's collect some more!

Remember, the more energy you collect, the more aliens you will be able to zap

Look at all the energy you have collected so far!

Let's collect some more!

Remember, the more energy you collect, the more aliens you will be able to zap

Wow! That energy bar is getting full!

Can you collect a little more?

Remember, the more energy you collect, the more aliens you will be able to zap

Wow! That energy bar is getting full!

Can you collect a little more?

Remember, the more energy you collect, the more aliens you will be able to zap

Appendix E: Questionnaire used in Experiment 8 (Chapter 5)

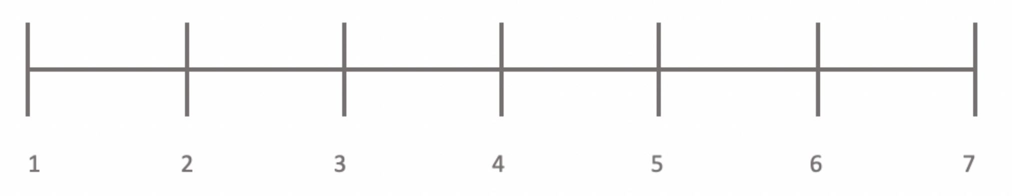
Please answer this questionnaire honestly. Your responses to this questionnaire will not affect the reward you receive for participating in the study.

1. Did you expect to be asked about the objects again after the working memory task (when you were shown the objects/tested on them the first time)?

Yes No

2. If yes, why?

3. To what extent did you think about the objects between encoding (when you were shown the objects/tested on them the first time) and the final test? (1 = not at all; 7 = all of the time).



4. Were you thinking about any objects in particular during the interval between encoding (when you were shown the objects/tested on them the first time) and the final test? If so, which ones?

Initial task:

The following questions refer to the test immediately after encoding:

How do you think trying to prioritise the 1st position affected your memory **for this item** compared to the condition in which every object was worth 1 point? (1 = large negative effect, 5 = no effect, 9 = large positive effect)

1	2	3	4	5	6	7	8	9

How do you think trying to prioritise the 1st position affected your memory for **the less valuable items** compared to the condition in which every object was worth 1 point? (1 = large negative effect, 5 = no effect, 9 = large positive effect)

1	2	3	4	5	6	7	8	9

Final task:

The following questions refer to the final test of the items you just completed:

How you think prioritising the 1st position during the initial task affected your memory for **these items** during the final test relative to the condition in which all items were worth 1 point? (1 = large negative effect, 5 = no effect, 9 = large positive effect)

1	2	3	4	5	6	7	8	9

How you think prioritising the 1st position during the initial task affected your memory for **the less valuable items** during the final test relative to the condition in which all items were worth 1 point? (1 = large negative effect, 5 = no effect, 9 = large positive effect)

1	2	3	4	5	6	7	8	9

Do you have any further questions about the experiment?