

The interplay between selective attention and working memory: A behavioural, neural and computational perspective



Mengya Zhang
Robinson College

This dissertation is submitted for the degree of Doctor of Philosophy

February 2021

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the preface and specified in the text. It is not substantially the same as any work that has already been submitted before for any degree or other qualification except as declared in the preface and specified in the text. It does not exceed the prescribed word limit for the School of Clinical Medicine Degree Committee.

The interplay between selective attention and working memory: A behavioural, neural and computational perspective

Mengya Zhang

Abstract

Selective attention (SA), the process by which information is prioritized for processing according to its relevance to current goals, and working memory (WM), the temporary storage and/or manipulation of information in mind, are considered to be important building blocks in human cognition. Both are essential for coordinating thought and action, and both are foundational for the emergence of other more complex executive functions, like planning and problem solving. The complicated interplay between SA and WM has been investigated across a growing number of experimental studies, with attentional processes influencing various stages of WM, and vice versa. Behavioural evidence suggests that SA can bias processing of information as we anticipate, encode and maintain contents in memory, whilst WM can serve to maintain a template as we search. Neuroimaging studies have observed a highly similar frontoparietal network subserving both processes, indicating anatomical and functional overlap in their corresponding neural mechanisms. Nevertheless, despite substantial evidence for cognitive and neural overlap, almost everything we know about the relationship between SA and WM is derived using group-average performance. In reality, some individuals may rely on shared sub-processes to perform tasks more so than others. In this thesis we extended previous work by understanding this individual variability. The first experimental chapter describes the development of two behavioural paradigms tapping SA and WM. These paradigms are better suited to address this question, relative to previous experimental approaches, because they are matched on task-specific features while being independently scalable in terms of difficulty. The second experimental chapter used functional magnetic resonance imaging (fMRI) in combination with these tasks to identify the neural correlates of individual differences in the strength of SA-WM coupling across participants. The third experimental chapter builds upon the neuroimaging study and addresses whether computational models trained to perform the same set of tasks share any mechanistic properties observed in the human brain, providing a useful framework in which predictions about the relationship between cognitive processes can be readily tested. Lastly, in the final experiment we used cognitive training to test whether altering SA would lead to changes in the related WM system, and whether these gains are modulated by baseline individual differences in the strength of their coupling. Together, along with an opening General Introduction and concluding Discussion, these chapters explore heterogeneity in the relationship between SA and WM from multiple perspectives, integrating advances in human cognition, neuroimaging and computational modelling

Acknowledgement

I would first like to thank my supervisor, Dr Duncan Astle. His wisdom, patience and encouragement have shaped every bit of this thesis and kept me going during the most difficult times. What I appreciate the most is his open-mindedness, without which this work would not have become a showcase of multiple exploratory techniques.

I acknowledge the CSC and Cambridge Trust for jointly providing my scholarship, and the Medical Research Council for funding to this research. I would also like to thank my colleagues from the 4D group and CBU in general for your generous and bright advices on this study and PhD life, and for creating the incredible vibe at tea times and Christmas parties. In particular, my esteemed colleague Joseph Peter Rennie, with whom I worked closely together for the first half of my PhD and shared many interesting/strange conversations and laughter. I'm also grateful for Jonathan Jones and Edwin Dalmaijer for their comments and technical expertise on the thesis.

And to my father who passed away during the write-up of this thesis. There is nothing in the world I wouldn't give up to go back in time and see you again. You were not a man of science but it is because of your unconditional love and support I feel very safe to follow my interest in research. I hope you are proud of me for this little accomplishment. I love you deeply. To my mother who is the strongest woman I know of, for shielding me from harm and heartbreaks and teaching me the merit of being responsible and kind. And to my wider families who lent us a helping hand in the hard time.

Finally, to my fiancé Alex whom I met in the CBU, for your love and quirkiness which kept me sane during the thousands of hours we spent together in the house in the past year. Thank you for being my biggest support.

Table of Contents

Chapter 1: General Introduction	11
1.1 The role of selective attention in theories of working memory	11
1.2 Experimental evidence for the role of attention in working memory	14
1.3 Dynamic interplay between SA and WM	18
1.4 The role of working memory in theories of selective attention	19
1.5 Neural overlap between SA and WM	21
1.6 Summary and foreshadowing the rest of the thesis	24
Chapter 2: Task development and behavioural validation.....	26
2.1 Introduction	26
2.2 Methods	28
2.2.1 <i>Participants</i>	28
2.2.2 <i>Stimuli, tasks and procedures</i>	28
2.2.3 <i>Analyses</i>	32
2.3 Results	35
2.3.1 <i>Performance on the new SA and WM tasks</i>	35
2.3.2 <i>Test-retest reliability</i>	38
2.3.3 <i>Comparison tasks and between-task relationships</i>	38
2.3.4 <i>Subgroups</i>	39
2.4 Discussion	44
Chapter 3: Individual differences in the shared neural substrate of selective attention and working memory.....	49
3.1 Introduction	49
3.2 Methods	51
3.2.1 <i>Participants</i>	51
3.2.2 <i>Experimental design and procedure</i>	52
3.2.3 <i>Behavioural analysis</i>	53
3.2.4 <i>fMRI acquisition, preprocessing and analysis</i>	54
3.2.5 <i>Group-level univariate analysis and conjunction</i>	55
3.2.6 <i>Group-constrained subject-specific (GCSS) analyses</i>	55
3.3 Results	57
3.3.1 <i>Behavioural results of in-scanner tasks</i>	57
3.3.2 <i>Group-level univariate results</i>	58
3.3.3 <i>GCSS results</i>	61
3.4 Discussion	63

Chapter 4: Task representations in the brain and artificial neural networks	67
4.1 Introduction	67
4.2 Methods	70
4.2.1 <i>Recurrent neural network model</i>	70
4.2.2 <i>Coding the behavioural tasks</i>	71
4.2.3 <i>Dissecting the RNN model</i>	73
4.2.4 <i>Representational similarity analysis</i>	75
4.2.5 <i>fMRI data pre-processing and comparisons to RNN using RSA</i>	75
4.3 Results	76
4.3.1 <i>Training RNNs on the attention and short-term memory tasks</i>	76
4.3.2 <i>Relationships between representations of tasks by RNN</i>	78
4.3.3 <i>Clustering produced two functionally separable subsets of network neurons</i>	79
4.3.4 <i>Representational similarity analysis of the clustered neurons</i>	81
4.3.5 <i>Relating the RDMs of the network model to brain activity</i>	82
4.4 Discussion	85
4.4.1 <i>The emerging network structure suggests overlapping neural circuits supporting SA and WM</i>	86
4.4.2 <i>Brain regions showing similarity to the overlapping RNN neurons coincides with a frontoparietal network</i>	87
4.4.3 <i>Functional roles of the identified brain networks</i>	88
4.4.4 <i>Limitations and future directions</i>	89
Chapter 5: Does training selective attention improve working memory?	91
5.1 Introduction	91
5.2 Methods	94
5.2.1 <i>Participants and procedures</i>	94
5.2.2 <i>Training tasks</i>	95
5.2.3 <i>Transfer tasks</i>	95
5.2.4 <i>Data analysis</i>	97
5.3 Results	97
5.3.1 <i>Pre-training group differences</i>	97
5.3.2 <i>On-task training effects</i>	99
5.3.3 <i>Transfer effects</i>	100
5.4 Discussion	105
Chapter 6: General discussion	109
6.1 <i>Relationship between SA and WM as assessed by our new tasks and individual differences</i>	110
6.2 <i>Common neural correlates for SA and WM and their functional roles</i>	114

6.3 <i>Limitations</i>	116
6.4 <i>Future directions</i>	118
6.5 <i>Conclusions</i>	120
References:	121

List of Figures

Figure 1. Schematics for SA task.....	30
Figure 2. Schematic for WM task.....	30
Figure 3. Distribution of search slopes of log-transformed RTs	36
Figure 4. Distribution of KL divergence for WM task.	36
Figure 5. Mixture model fit to all participants' data.....	37
Figure 6. Model fit to an example subject, whose performance was poor.	37
Figure 7. Profiles of subgroups.....	40
Figure 8. CFA baseline model.	44
Figure 9. Schematics for SA and WM tasks in fMRI study	53
Figure 10. GCSS pipeline	57
Figure 11. SA task. RT on target present trials by difficulty level	58
Figure 12. WM task. Orientation error by memory load	58
Figure 13. SA task vs. baseline contrast	60
Figure 14. WM task vs. baseline contrast	60
Figure 15. SA difficulty-dependent activity.	60
Figure 16. WM load-dependent activity.	60
Figure 17. Conjunction of the two task vs. baseline contrasts.....	60
Figure 18. Percentage of participants showing conjunctive activation at each parcel.....	61
Figure 19. Group difference in correlation strength was found in right superior parietal lobule (SPL).....	62
Figure 20. Correlation coefficients between SA difficulty-related and WM load-related activity by group in right SPL.....	62
Figure 21. RNN structure from (Ehrlich et al., 2021).....	71
Figure 22. Network input and output on the SA task.	73
Figure 23. Network input and output on WM task.	73
Figure 24. Behavioural accuracy in each task from fMRI study (top), and RNN (bottom). ...	78
Figure 25. Selectivity of RNN neurons to SA and WM tasks.	79
Figure 26. Mean silhouette score with each K-number solution	79
Figure 27. Accuracy difference between lesioning a cluster and same number of random units.....	81
Figure 28. Results of sequential lesioning.....	81

Figure 29. Representational dissimilarity matrix of cluster 1 neurons between task conditions.	82
Figure 30. Representational dissimilarity matrix of cluster 2 neurons between task conditions.	82
Figure 31. Neuron selectivity by clusters	82
Figure 32. Strength of averaged group correlation between brain and RNN cluster RDMs ...	85
Figure 33. ROIs that showed significant difference in brain-RNN similarity (cluster 1 > cluster 2).....	85
Figure 34. Training trajectories of the highWM and lowWM groups on the SA training task.	100
Figure 35. Training trajectories of the control group (highWM) on the control training task.	100
Figure 36. Distribution of level reached in SA training at the end of the first and last sessions	100
Figure 37. Difficulty level reached in control training at the end of the first and last sessions	100
Figure 38. WM task performance before and after training by group.	102
Figure 39. Forward spatial span.....	102
Figure 40. Backward spatial span	102
Figure 41. Forward digit span.....	102
Figure 42. Backward digit span	102
Figure 43. Stroop task (RT difference).....	102
Figure 44. Abstract reasoning task (accuracy).....	103

List of Tables

Table 1. Descriptive statistics of all tasks.....	37
Table 2. Pearson’s correlations between representative measures of all assessment tasks	39
Table 3. Pearson’s correlation between tasks within each subgroup.....	42
Table 4. Model fit indices and model comparison statistics.....	43
Table 5. Modification index for metric invariance model in highSA+highWM group.....	44
Table 6. Conjunction results from group-level task vs. baseline contrasts.....	61
Table 7. Conjunction results from GCSS analysis.....	62
Table 8. Regions of interest more similar to neurons in cluster 1 compared to cluster 2.....	84
Table 9. Results of one-way ANOVAs and pairwise comparisons for pre-training group differences on transfer tasks.....	98
Table 10. Results of ANCOVA tests for post-training differences with baseline performance as covariate.....	104

Chapter 1: General Introduction

Higher-order cognition enables complex adaptive behaviours. These complex behaviours recruit multiple underlying processes, and this thesis focuses on two that are necessary for multiple different types of higher-order cognition. Selective attention (SA) is the process of selectively allocating processing resources to a specific aspect of environmental input, while ignoring others, according to both current goals and the relative differences in the saliency of those inputs (Broadbent, 2013; Duncan, 1984; Treisman, 1964). Working memory (WM) is the temporary storage and/or manipulation of information in mind to support other goal-directed tasks (Baddeley & Hitch, 1974). Both are essential for coordinating thoughts and actions. They work in concert, and play an important role in other so-called ‘executive functions’, such as planning, inhibition control and learning (Miyake & Shah, 1999). In recent years, SA and WM have been increasingly viewed as overlapping constructs, with a growing number of experiments investigating the influence of attentional control process on various stages of WM, and vice versa (Awh & Jonides, 2001; Chun, 2011; Eimer, 2014; Gazzaley & Nobre, 2012). This trend is deeply rooted in the various theoretical frameworks of WM. When placed within its historical context, the interest in the interplay between SA and WM is not new. Even from its first formalisation, WM has incorporated some attentional process that enables it to go beyond pure passive storage (Baddeley & Hitch, 1974; Cowan, 1999; Engle et al., 1999). This is our starting point in understanding the interplay between WM and SA.

1.1 The role of selective attention in theories of working memory

The now famous multi-component model proposed by Baddeley and Hitch (1974) incorporated the existence of a hypothetical component, the ‘central executive’. This putative controlling resource prioritises incoming information and the manipulation of stored representations within the two domain-specific components: the phonological loop, responsible for storing verbal material, and the visuospatial sketchpad, responsible for integrating visual (e.g., colour and shape) and spatial information (e.g., locations). The closest the central executive ever came to a fully mechanistic explanation was when it was likened to an already described model, the Supervisory Attentional System (SAS, Norman & Shallice, 1986). This model formalised the trade-off between automatic and controlled processing, particularly in relation to the selection of action schemas. This ability to select amongst multiple possible representations made this an attractive model to borrow from. However, despite this attempt to provide some formalisation to the central executive, most were unconvinced that this crucial element of the multi-component model had been sufficiently specified. Donald (1993) famously said “*the ‘central executive’ sits atop the mountain of working memory like a gigantic Buddha, an inscrutable, immaterial, omnipresent homunculus, at whose busy desk the buck stops every time memory theorists run out of alternatives*”.

In later attempts to nail down the nature and function of this admittedly vague attentional resource, Baddeley and colleagues used the now well-deployed dual-task methodology. To provide one example: in a series of experiments, Baddeley and colleagues (1996) recruited two different age groups to perform a task wherein they had to press a key as soon as a target stimulus was presented, whilst counting the total number of stimuli. This running count would have to be reported when cued, thus introducing a concurrent memory task. The attentional demand was varied in the form of irrelevant stimuli appearing, in either the same or different sensory modality as the target, as well as an occasional instruction to switch target to a new stimulus, again in the same or different modality as the old target. Older participants (mean age = 72) were slower than their younger counterparts (mean age = 42) in all conditions. But having controlled for the general decline in cognitive ability, the age-related effect only remains in the conditions where same-modality distractors are present. Therefore, aging may have a specific detrimental effect on the focus of attention beyond the impact of general intelligence and processing speed, especially when irrelevant information is in the same modality. This is a clear example of the mechanistic role for the central executive within working memory – filtering out distractions. Furthermore, there are clear age associations with this mechanism. Baddeley also theorised other functions for the central executive, such as task switching and divided attention, which were tested in the context of Alzheimer’s patients who showed selective deficits in dual-task performance (Spinnler et al., 1988).

Later a new component, the ‘episodic buffer’ was introduced to bridge the gap between long-term memory (LTM) and WM, and provide a multidimensional platform that allows features from different domain-specific components to be bound into chunks or episodes (Baddeley, 2000). It was assumed that binding was attentionally demanding and thus relied heavily on the central executive. This would predict that an attentionally demanding concurrent task should have a very substantial effect on the capacity to bind information, in contrast to minor effects from disrupting the other subsystems (Baddeley, 2012). The results were mixed. Using the dual-task procedure, presentation of the stimulus array was accompanied by a demanding task such as counting backward by threes in conditions of binding (e.g. remembering both colour and shape) and single feature (e.g. remembering colour or shape). Indeed, there was an overall drop in performance with the concurrent task, but this was equivalent across conditions (Baddeley et al., 2011). In a different experiment, however, they found a differential effect between the binding and single feature conditions. Coloured shapes were presented sequentially, followed by a memory probe probing objects at any given positions. Performance of the earlier items in the sequence, but not the final one showed poorer retention of the bound stimuli. This was explored further using simultaneous presentations, with a single additional distracter object appearing between the presentation array and the final probe. Binding was differentially impaired by the presence of the distracter, despite it being task irrelevant and to-be-ignored. The authors took this as evidence that

although visual binding per se is not attention demanding, maintaining bindings against distraction is (Allen et al., 2006).

Another class of WM models that has recently gained prominence features a critical role for the allocation of attention to internal representations, as a form of WM maintenance. Unlike the multi-component theory (Baddeley & Hitch, 1974) which emphasizes modular distinctions between separate components, this family of models is perhaps better thought of as a set of *process accounts* of WM maintenance. Arguably the most representative of them is the embedded-processes model proposed and formalized by Cowan (1988, 1999, 2001). Information held in memory is thought to exist in different states, distinguished by attentional selection. Information comes in from the environment through a very brief sensory store, activating features in LTM corresponding to the sensory properties of the incoming information, such as letters, words, digits, etc. These mental representations can exist in two distinct states: a capacity-limited state referred to as the focus of attention, which is embedded in a more expansive state referred to as the activated LTM. It is purported that the focus of attention can hold four (± 1) chunks of information at any given moment, decided by voluntary and involuntary (bottom-up) orienting of attention (Cowan, 2001). The items previously in the focus of attention but which have no importance to future behavioural goals transition into activated LTM, which is not constrained by capacity limit per se but is susceptible to temporal decay and interference effects. In a variation of this model, Oberauer (2001, 2002, 2005) modified the Sternberg recognition task, in which response time for a recognition judgement of a memory probe increases with number of items maintained in mind (Sternberg, 1969). In his task, two lists of items were presented and a retro-cue was used to indicate that one of them as relevant for recognition. When the probe occurred 1 s after the cue, response time was only affected by the size of the relevant list, whereas the irrelevant list was not completely forgotten as probes from this list were more slowly rejected compared to new probes. The latter effect disappeared 5 s or longer after the retro-cue. This is taken as a demonstration of the distinct representational states whereby objects exist: the relevant list is hypothesized to remain within the focus of attention whilst the irrelevant list is removed from the “central spotlight” but persists in activated LTM for a brief period (Oberauer, 2001). Moreover, by varying the cue-to-probe interval, Oberauer (2005) estimated that it takes approximately 1s to remove irrelevant items from the focus of attention, and the cost of switching attention to other items within the activated LTM depends on the number of eligible candidates that WM has direct access to (Oberauer, 2002).

As will become clearer as this chapter unfolds, this embedded-processes account of WM can to some extent explain a whole range of phenomena, like the nature of individual differences in WM capacity, the strong relationships between WM and other cognitive abilities (e.g. fluid reasoning). This is because the scope of focused attention plays a key role in WM maintenance itself. In contrast to passive storage tasks, a task with a very similar trial procedure, but which prevents covert rehearsal, correlates well with other cognitive aptitudes. For example, in a version of running memory span, digits are presented

fast and the sequence stops at an unpredictable point, after which the participants recall as many items as possible from the end of the list. Since rehearsal is prevented, information must be retrieved from activated LTM and brought into the focus of attention. In theory at least, this should make this task better at capturing variance in the capacity limits of the attentional focus or, put more simply, the ability to maintain representations in an activated state for processing (Cowan et al., 2005, 2006). A further illustration for this perspective is shown in children too young to use covert verbal rehearsal (unlike older children and adults). For these youngsters even a simple span task (no manipulating information was involved) is correlated with more general cognitive abilities. In other words, various WM tasks correlate with other cognitive abilities to the extent that attention is required for storage and/or processing (Cowan, 2008).

This very brief overview of WM theory contrasts two different ways of conceptualising the role of attentional control. The first is as a component of WM, with something akin to SA stepping in at particular moments to interact with the ongoing processes of other components. The second class of theory is process-based, with SA underpinning the very maintenance process itself. Whilst this kind of theoretical work provides the backdrop to thinking about the role of SA within WM, it will soon become clear that most experimental evidence does not really distinguish between these contrasting conceptual frameworks.

1.2 Experimental evidence for the role of attention in working memory

Following the multicomponent model, many adopted a similar procedure of concurrent task performance to investigate the relationship between SA and WM. The premise of the dual-task paradigm is that if two tasks compete for the same capacity-limited resource, performing both concurrently should result in interference, relative to single-task conditions. Within this dual-task context, attentional capture is one particular phenomenon that has been used to demonstrate the recruitment of attentional resources during WM maintenance. In attentional capture, searching for a target defined by a unique feature (e.g. a red circle among green circles) is relatively easy compared to occasional appearances of an irrelevant but unique distractor (e.g. a red diamond) which captures attention and slows down the search. Lavie and de Fockert (2005) presented stimuli like these, but each stimulus also contained either a horizontal or a tilted line. Participants were required to find the red circle and report the direction of the line. This was done either in isolation or within the retention period of a verbal WM task, with varying memory loads. The effect of the unique distractor was enhanced under high WM load, compared to low or no load. This suggested that WM indeed involved some attentional process that is also responsible for the shifting of spatial attention and/or the shielding of processing from potent distractors. Within the literature there are many examples of studies like this, such as the Flanker task (Eriksen & Eriksen, 1974), in which participants need to indicate the direction

of target central arrows flanked by congruent or incongruent arrows on both sides. Attentional control and response inhibition are thought necessary for successful performance on this task. Again, dual-task interference is present in the flanker task (Pratt et al., 2011). Event-related potential (ERP) components associated with early top-down attentional modulation for both type of distractors and late attentional shift for incongruent distractors diminish in the dual-task condition. This indicates that when WM is in active use, the ability to bias goal-directed visual processing, as well as the ability to suppress interference from flankers, is reduced.

However, whilst there is substantial evidence that SA in some form seems to be recruited when WM is in use, it has remained somewhat unclear what precisely it is doing. Moreover, it is clear that some aspects of attention are not recruited during WM. For example, again using a dual task logic, inefficient visual search (such as determining whether a “T” existed among randomly rotated “L”s; search time increases linearly with set size) and spatial WM tasks interfere with one another, whereas inefficient search and non-spatial WM tasks using colours and shapes do not (Oh & Kim, 2004; Woodman et al., 2001). One mechanistic account for the role of SA within WM is that the former allows for spatial rehearsal, inspired largely by the demonstration that visual search during spatial WM can interfere with subsequent recall (Awh & Jonides, 2001), but not so for maintaining other types of information. That said, others have demonstrated that non-spatial and verbal WM tasks produce comparable interference effects to the spatial variants (Anderson et al., 2008; Han & Kim, 2004), challenging the notion that SA and WM only share domain-specific resources. Instead, these authors argue that a domain-general “executive” process subserves successful performance on both tasks, hence the broader dual-task interference. The bottom line is that whether the interference effects are domain-specific or domain-general is unknown – there are good examples for each finding. Anderson et al (2008) reflected on the impact of task choice on the results, noting that whether the non-spatial WM task contains additional requirements (e.g. manipulation of information, retaining temporal order of objects) could account for the inconsistencies.

Another approach often used to explore the relationship between SA and WM is to test whether individual differences in one ability correlate with that of the other. The typical procedure involves testing a large sample of participants with WM span tasks and then selecting two subgroups: those with high and those with low WM capacity (WMC). These participants then perform other tasks assessing SA (e.g. Astle et al., 2014; Kane et al., 2001; Lustig et al., 2001; Turner & Engle, 1989). If the low WMC group’s performance is worse on the secondary SA task, in comparison to that of their high WMC counterparts, then the cognitive process taxed by the second task is interpreted as interacting with, or relying upon, WM. There are numerous studies that show WMC grouping predicts performance on tasks requiring attentional control, where participants must orient their attention to targets, dimensions, locations, or events presented among distractors, such as in flanker and dichotic-listening tasks (e.g. Bleckley et al., 2003; Colflesh & Conway, 2007; Conway & Kane, 2001; Heitz & Engle,

2007; Redick & Engle, 2006). Although, the exact boundary conditions that determine the interaction between SA and WM remains contentious. In a series of experiments, Kane and colleagues (2006) selected high and low WMC participants, who then performed visual search tasks with varying configurations, including efficient and inefficient letter search (F among Os or Es), classic conjunction search (colour and direction of lines), spatial configuration search (F among Es and 90° tilted Ts) and constrained search that requires endogenous control of spatial attention (start from the top position and search clockwise). There were no group differences in search efficiency or error rate. This led the investigators to conclude that WM differences do not generalize to SA tasks. Implying that WM is not needed to actively maintain goals, to restrain prepotent responses or constrain attentional focus to particular stimuli or locations in space amid distractors". Nonetheless, several factors could potentially explain the discrepancies between their findings and the dual-task approach, especially considering both employed quite similar tasks. First, the correlational and between-subject nature of the high vs. low WMC method may be less sensitive than the dual-task manipulation (Vandierendonck, 2014). Second, modifications to the task design might vary the results in important ways. For example, adding additional attentional constraints to the search process can change the outcome, with search performance now distinguishing high and low-capacity individuals (Poole & Kane, 2009). In this case participants had to limit their search to several cued target positions (the target will be shown at one of these locations) interspersed among more distractor positions. The authors compared WM effects when stimuli only appear at cued positions versus at all positions, and with long versus short latencies between the cue and search array. WMC predicts performance on trials with stimuli are shown at all locations and only when the attentional focus needs to be maintained over a brief time. The former suggests that WM is related to the ability to ignore lures appearing outside of focus of attention, while the latter implies whilst both groups are equally capable of initiating attentional focus, over time high WMC individuals maintain it better. Others have also shown that the *type* of visual search is the crucial for determining the impact of individual differences in WMC. Sobel et al. (2007) found that although WMC did not predict performance on trials in which saliency-based pop-out search was effective in identifying the target, it did so when the automatic bottom-up search must be overridden by a more difficult top-down search.

To summarise briefly so far: as with the dual-task methodology, the individual differences approach has yielded plenty of positive evidence that SA plays some performance-limiting role within WM. Both of these approaches have so far shown that the scope and nature of the SA-WM interaction depends upon the specific task implementations. The dual-task approach has produced more compelling evidence for the overlap between visual search and WM, although there remains some debate about which kinds of WM maintenance are interfered with by concurrent visual search. In contrast, the individual differences literature largely suggests that WM does not interact with the search process per se, but rather with the executive processes that guide search.

Another source of experimental evidence relevant to the role of SA within WM is an electrophysiological marker sensitive to the amount of information stored in WM (Vogel & Machizawa, 2004). The contralateral delayed activity (CDA) is a slow negative wave first observed during retention phase of change detection task at posterior electrodes and is thought to reflect the number of items maintained in visual WM. Many independent studies have replicated this finding and critically, established that the CDA is not related to other factors such as general task difficulty, but rather the amplitude change of which is correlated with individual's WMC (Diamantopoulou et al., 2011; Jost et al., 2011; Kang & Woodman, 2014; Kuo et al., 2012). Seminal work by Vogel and colleagues (2005) examined the moment-to-moment contents of WM through CDA amplitude and observed that selection efficiency differed substantially across individuals. These differences were in turn strongly predicted by their WMC. Specifically, participants were presented with a bilateral array of coloured rectangles of varying orientations and were asked to remember the orientations of only the red items in either the left or right hemifield, as indicated by an arrow. Memory for these remembered items was tested with an array that was either identical to the original memory array or differed by one orientation. Some trials contained two red items with two blue items in the same array. These blue items were distractors. Whereas other trials contained just two or four red target rectangles in each hemifield. The critical finding was that in high capacity individuals the CDA during retention reflected the number of *memory targets* in the trial, whereas in low capacity participants it reflected the overall number of items regardless of their status. In other words, the ERP magnitude associated with two-target-two-distractor condition was equivalent to that associated with two-target-no-distractor condition for high WMC group. Whereas in the low WMC group the former ERP magnitude was higher than latter, and was closer to that associated with the four-target-no-distractor condition. The finding suggests that high WMC participants were able to more efficiently filter out distractors from entering WM, implying that the allocation of memory capacity to irrelevant objects may be a primary source of differences in overall storage capacity. This effect was also shown in an aging study that investigated the relationship between filtering efficiency and age-related reductions in WMC (Jost et al., 2011). In comparison with young adults, older participants (64 year-old and above) had lower performance in the change detection task, accompanied by reduced filtering efficiency during early retention stage. This aligns with other findings that demonstrate that old adults allocated more attention to irrelevant information in visual processing/WM encoding, linking the impairment in WM to a selective deficit in suppressing distractions (Gazzaley et al., 2008; Gazzaley, et al., 2005). Whilst many have questioned whether the CDA really does reflect the number of items in WM (e.g. Astle et al. 2014), the clear differences in lateralized components for high and low WMC participants, specifically for distractors, does speak to the idea that SA may allow for the prioritization of what is stored within WM.

1.3 Dynamic interplay between SA and WM

In the previous section we came across one specific functional role that SA may play within WM – the ‘gatekeeper’. A number of researchers have argued that the primary role for attention is the biasing of perceptual representations, such that only some gain access to the limited capacity of WM (Awh et al., 2006). However, one particular subfield of WM research has shown that there can be a far more dynamic interplay between these processes, with SA biases operating at *multiple stages* of WM encoding, maintenance and retrieval. This interplay is revealed by cueing participants to orient attention at multiple different points of information processing. The basic premise is to present participants with memoranda to maintain in WM, and then to cue their attention to different locations, features or items within the memory array. This cueing is done with a so-called ‘retro-cue’ procedure, to manipulate the allocation of attention during WM maintenance by retrospectively managing information about which remembered items would be tested (Griffin & Nobre, 2003; Landman et al., 2003). Retro-cueing can be distinguished from a post-cue, which acts as a direct reminder of the probed item. Instead, a retro-cue provides information about the relevance of given items (e.g. spatial locations or object category) to guide internal shifts of attention towards memory representations (Gazzaley & Nobre, 2012). Just like when attention is oriented within perceptual representations, retro-cues confer large behavioural benefit in WM performance (Lepsien & Nobre, 2007; Sligte et al., 2008), both through removal of uncued items from WM storage and strengthening of cued items (Q. Li & Saiki, 2014; Rerko & Oberauer, 2013). Astle et al (2012) showed that the role of a retro-cue could be somewhat flexible: when WM loads are within capacity limits (four in the relevant experiment) valid cues act to prioritize memory search so that the differential effect between valid and invalid cues are manifested in the reaction times but not in the accuracy of responses; whereas when size of the memory array exceeds capacity limit, retro-cues are used to discard uncued items, therefore invalid cues induce a catastrophic effect to accuracy. Finally, retro-cues have even been shown to be able to reinstate items to WM after behavioural measures indicate that they have been lost (Murray et al., 2013). This result would surely please those who have advocated WM as the spotlight of attention within LTM, as it implies that reorienting attention can rescue items just outside that spotlight (e.g. Oberauer, 2013). Taken together, pre-, retro- and post-cue paradigms have shown that SA is highly flexible and can operate at multiple points within a standard WM trial, biasing processing as information is encoded, maintained and retrieved, and even reinstating items hitherto thought lost.

Of course, simply because SA can operate at multiple time points does not mean that it acts in the same way at each. These cueing paradigms have been adapted and used alongside multiple neuroimaging modalities showing that the attentional benefit afforded by retro-cues is achieved by top-down signals biasing persistent delay activity in the posterior cortex reflecting the amount of information maintained (Kuo et al., 2012). Using face and scene stimuli, Lepsien and Nobre (2007) manipulated the expectation of a probed category during WM delay and reported the activity in the face- and scene-processing areas

was increased if the object representation currently being elevated by SA matched the preferred category of the area, and dropped if attention was switched to the other category. Moreover, prefrontal cortex (PFC) and posterior parietal regions have been associated with orientating attention within WM to both locations and categories (Lepsien et al., 2005; Roth et al., 2005), and are shown to exhibit strong functional connectivity with visual areas when attentional modulation occurs (Kuo et al., 2011). This has led some to suggest that the function of SA during WM maintenance is highly analogous to its role in perceptual processing, and that it is subserved by a common subset of functional networks (Tamber-Rosenau et al., 2011). In short, under this view top-down biases emanating from fronto-parietal cortex can act either to bias sensory representations according to relevant goals, in which case we refer to it as SA, or to refresh or enhance those representations in the absence of sensory input, in which case we refer to it as WM.

1.4 The role of working memory in theories of selective attention

So far, we have primarily focussed on the potential role of SA within WM. Next, we consider the reverse relationship. In visual environments where multiple objects compete for processing resources, the challenge is to find relevant information and to ignore information unrelated to current task goals. This is a fairly standard definition of SA (Desimone & Duncan, 1995; Duncan & Humphreys, 1992; Treisman & Gelade, 1980). To mimic this everyday task, visual search is arguably one of the most influential paradigms designed to uncover strategies, mechanisms and neural correlates mediating SA (Estes & Taylor, 1964; Neisser, 1964). The main setup is to vary the number of distractors among which the target is embedded and measure either the reaction time to find the target, or the accuracy in detecting the target, when the display is brief. The change in performance with increasing number of distractors in the display (i.e., set-size effect) is used to infer how attention operates and selects visual information (Eckstein, 2011).

In an early model of visual search, SA was conceptualized as a temporally serial mechanism that chooses and processes one item at a time. This was supported by the general finding that reaction times in search increase linearly as a function of set size (e.g. Shiffrin & Schneider, 1977). It was further popularized by feature integration theory (Treisman & Gelade, 1980) which proposed a two-stage process, during which individual features of objects were first processed in parallel and subsequently visual attention, which operated serially from item to item, was needed to assign or bind different features to an item. This model was successful in explaining the large difference between the shallow set-size functions (i.e. reaction time increases little with more distractors in the display) of feature search and the steep set-size functions for conjunction search. In feature search, a target is differentiated by a single attribute (e.g. colour or shape) and in conjunction search, it is defined only by the joint presence of two features.

However, many have argued against the strict dichotomy between serial and parallel search, for instance, Duncan and Humphrey (1989) observed a continuum of increasing search difficulty (steeper slopes of set-size functions) as a result of decreased target-distractor similarity and increased distractor-distractor similarity (for preview, we have largely based our task design on this finding). This has given rise to a more nuanced model that yield a continuous variety of search slopes. The guided search model (Wolfe, 1994; Wolfe et al., 1989) states that search begins with a massively parallel stage that process information about basic visual features across large portion of visual field. An item is further processed in a later capacity-limited stage if they exceed an activation threshold. Attention then serially processes one item at a time guided by the item with the highest activation and before switching to the next one in line. Increasing the target–distractor heterogeneity will increase the probability that the target will be ranked highly in terms of activation, and thus, attention will reach the target faster, reducing predicted reaction times. This line of research (alongside a lot of single cell neurophysiology) fed directly into one of the most influence models of SA – biased competition theory (Desimone & Duncan, 1995). Within this theory, items are represented across multiple hierarchical receptive fields. Where competition occurs, say between targets and distracters, lower-level representations within the hierarchy must be biased. This biasing – either by the selective enhancement of relevant representations or suppression of irrelevant representations – then feeds forward through the processing hierarchy such that the competition at the higher order level of representation is biased in favour of the relevant item. The more competition there is, the greater the degree of biasing required and where targets and distracters share many features, the feature map that captures the unique difference between them will be proportionally smaller, and thus that bias more difficult to apply. Likewise, where distracters are highly heterogeneous, the map that uniquely captures the relevant feature will be proportionally smaller, and thus the bias more difficult to apply. This is why search slopes can vary systematically depending upon the graded degree of target-distractor similarity and distracter-distracter similarity. Researchers have also long been interested the role of WM within the visual search processes, from holding a representation of the search template during preparation (e.g. Desimone & Duncan, 1995; Van Der Velde & De Kamps, 2001) to representations of searched objects and making comparisons with the attentional template during object identification (Houtkamp & Roelfsema, 2009). Much like studies of role of SA within WM, neurophysiological studies using the CDA marker have investigated extensively the reverse and showed that CDA amplitude continues to increase as search progresses and is correlated to WMC estimated in a separate task (Emrich et al., 2009; Luria et al., 2016). Interestingly, contrary to the behavioural results from the abovementioned literature, higher WMC individuals (measured by change detection tasks) search more efficiently and accurately, albeit sometimes only in difficult search scenarios, possibly because they are better at rejecting distractors from occupying WM and keeping track of previously searched items. However, although the evidence for WM involvement in visual search is more straightforward than the reverse relationship, the mechanisms underlying this involvement may not be unitary. For example, Emrich et al (2009) found that CDA activity during

search is positively predictive of response time as well as participants' memory capacity, leading them to propose that high-capacity individuals search faster as they can prevent searched distractors from being revisited (M. S. Peterson et al., 2001). Conversely, other studies found that higher WMC is associated with smaller CDA amplitude, suggesting that they rely less on WM to perform the search task. Moreover, RT improvement as a result of informative cues restricting the number of locations to attend to (cued – uncued conditions) has a negative relationship with reductions in the CDA. In other words, those who benefit more from the spatial cues are more effective in reducing items stored in memory (Luria & Vogel, 2011). This is in line with studies examining the relationship between WM and filtering efficiency, which attribute the variance in memory capacity to difference in the ability to resisting irrelevant information from being encoded (e.g. Astle et al., 2014; Jost et al., 2011; Lee et al., 2010; Qi et al., 2014; Vogel et al., 2005). Last but not least, it is also possible that the CDA during search tasks represents the attentional templates necessary to guide search (Desimone & Duncan, 1995; Duncan & Humphreys, 1989), as when a new search target is used repetitively throughout subsequent trials, CDA amplitude decreases and search efficiency improves, signalling the transition of target representation from WM to long-term storage (Carlisle et al., 2011; Woodman & Arita, 2011). There is also some evidence that search difficulty and WMC independently influence the involvement of WM in visual search (Luria & Vogel, 2011), suggesting the interplay between WM and SA is dynamic and subject to individual differences.

1.5 Neural overlap between SA and WM

The heavy overlap of neural correlates for spatial SA and WM has been widely reported in the field (Awh et al., 1995; Corbetta et al., 1993; Smith et al., 1995). Together with the various behavioural findings that show that certain types of SA requirements can compete with WM maintenance, individual differences in SA and WM covary, and that attentional biases can operate within WM, this has led to some researchers to postulate SA as key component of spatial WM maintenance. The clearest examples of this is in the spatial rehearsal effect (Awh & Jonides, 2001). Corbetta et al (2002) compared brain activity when attention is directed to a peripheral location, relative to that when maintaining the attention (during which participants were waiting to see and respond to a stimulus) over a delay period within the same participants. A frontoparietal network consisting of the intraparietal sulcus (IPS) and frontal eye field (FEF) showed activity modulated by both type of trials, consistent with their roles in controlling visuospatial attention and programming eye movements (Corbetta, 1998; Kanwisher & Wojciulik, 2000). Likewise, Ikkai & Curtis (2011) showed that the same areas, with the addition of dorsolateral prefrontal cortex (DLPFC) and superior temporal sulcus, were identified in a conjunction analysis across three tasks: maintaining spatial representation, covert attention and motor intention (during the stimuli-absence period). Thus, these different processes were all associated with

modulations within a shared neural network and supported by persistent activity in subregions of the frontoparietal network. Similar patterns of neural activity are indicated using object-based WM, raising the possibility that the common processes between SA and WM is beyond just spatial rehearsal (Pollmann & von Cramon, 2000). LaBar et al. (1999) observed IPS, FEF, precentral sulcus and middle temporal gyrus co-activated by a verbal n-back task and a spatial attention task.

So far, the studies included here have adopted relatively simplistic tasks to operationalise SA and WM (e.g. cue-directed attention to left or right visual field). In comparison, a study investigated activation patterns associated with visual search and memory search, both of which were carefully matched on the visual presentation and time course (Makino et al., 2004). Participants were trained prior to scanning to memorize four abstract shapes. Both tasks involve first presenting an arrow pointing at one of the corner locations followed by five items similar to the memory targets (four on each corner and one in the centre). In a visual search condition, participants were asked to make a judgment about whether the central item matched one of the corner items, starting with the item cued by the arrow. Whereas in the memory search condition the judgment was about whether the cued item match one of the memory targets. Incorporating a control condition, the authors were able to isolate activity uniquely related to searching within the perceptual domain and that within the memory domain. The results indicated that left DLPFC, right FEF and the right precuneus, but not the IPS exhibited conjunctive activity, suggesting a similar but not identical set of regions involved in common processes to both tasks in contrast to those previously reported with simpler SA and WM paradigms.

It seems clear that there is plenty of positive evidence for overlapping neural correlates for SA and WM, and that this has been broadly taken as strong evidence for shared processes. However, the overlapping neural substrates are in essence anatomical, which is not necessarily equivalent to shared neural, computational or cognitive mechanisms. Functional specification may exist in subregions beyond the resolution of functional magnetic resonance imaging (Nieder, 2004), and/or the same population of neurons may carry out task-specific adaptive functions depending on the task demands (Duncan & Owen, 2000; Rao et al., 1997), giving the potentially misleading impression that distinct constructs – in this case WM and SA – are mediated by the same neural mechanisms. In an attempt to circumvent these issues and demonstrate genuine mechanistic overlap, a series of experiments using spatial and object-based stimuli combined SA and WM encoding into one single task, but allowing demands within each domain to be manipulated independently (Fusser et al., 2011; Mayer et al., 2007). Participants performed easy vs. difficult visual search tasks in order to encode the targets into WM. If SA and WM share a common capacity-limited neural resource, certain regions should show an interaction effect when both WM and SA demands are high. Previous work had shown that blood oxygen-level-dependent (BOLD) signals in the posterior parietal regions track memory load and plateau when WM is filled (Todd & Marois, 2004; Xu & Chun, 2006). The question is whether this effect would show an interaction effect when both SA and WM demands are high, when the neural resource is presumably

exhausted. This pattern is observed in IPS, FEF and occipital cortex, whereas an additive increase in activation responsive to SA and WM demands was found in right PFC and bilateral insula. This suggests that the former set of areas reflect a shared processing bottleneck common to both SA and WM, and the latter may support separate mechanisms related to increased WM and attentional demands relatively independently. Overall, the results illustrate that activity within the frontoparietal network including IPS and FEF mediates a common neural mechanism shared by both SA and WM.

What is the shared mechanism? One possibility is that it is ‘top-down modulation’, as briefly mentioned previously. Both SA and WM require top-down control mechanisms because they depend on goal-directed processes resulting from previous knowledge and representations of task rules (D’Esposito & Postle, 2015). PFC has been implicated as a potential source of biasing signals that can influence processing in downstream cortical regions that receive anatomical projections from PFC (Braver et al., 2009; Duncan, 2001). These signals are thought to bias lower-level processing areas by enhancing task-relevant information and suppressing task-irrelevant information, thereby facilitating the processing of information necessary to achieve current goals (Pessoa et al., 2003; Reynolds & Chelazzi, 2004). Within WM similar mechanisms have been identified, during the encoding and maintenance of memoranda, in brain areas overlapping with those identified in SA. A study by Gazzaley and colleagues (2005) used both fMRI and ERPs in separate sessions to measure top-down modulation in WM encoding. Their participants were shown sequences of two faces and two scenes in random order and were instructed to selectively remember faces, scenes or view all items passively in separate blocks. fMRI activity was concurrently measured in the fusiform area (FFA, known to process faces) and parahippocampal place area (PPA, known to process scenes), and an ERP component thought to reflect face-selective activity was identified. Since the visual input for all conditions was identical, and there was a passive baseline condition, the relative facilitation and suppression of stimulus processing as a consequence of top-down signals could be investigated. The data revealed enhancement of FFA activity as well as an earlier occurrence of the face-related ERP component, relative to baseline when faces were to be remembered, and vice versa when scenes were to be remembered. The same group showed that the process of top-down suppression of irrelevant information may be selectively weakened in older adults. Interestingly for the current purpose, the degree to which this weakening occurred was directly related to impairment in WM performance, whereas the top-down enhancement process was intact (Gazzaley et al., 2005). A parsimonious account that has gained some popularity in the field (e.g. Gazzaley & Nobre, 2012) is that WM reflects top-down biases refreshing representations held in sensory cortex (Harrison & Tong, 2009; Serences et al., 2009). Indeed, there has been some support for top-down and stimulus-driven mechanisms activating shared neural representations, for example, Stokes et al (2009) used multivariate pattern analysis (MVPA) for neural activations in high-level visual cortex associated with imagining or viewing different shapes. Their findings revealed decoders trained on activity of visual perception can

differentiate between imagined stimuli above chance, suggesting that top-down and stimulus-driven mechanisms rely on shared neural representations within this region.

Casual evidence for the involvement of PFC in generating top-down modulation is provided by lesion studies and repeated transcranial magnetic stimulation (rTMS). Miller et al (2011) employed both methods to explore the effect of PFC disruption on the selectivity of category representations in the temporal cortex. After an rTMS session, or in patients with focal frontal stroke, stimulus-evoked activity patterns in the temporal cortex associated with distinct categories of stimuli were less distinguishable, implying decreased category selectivity. This is consistent with studies that have used TMS to perturb PFC activity, and observed reduced selectivity of categories and features within sensory processing areas, and reduced WM accuracy (Feredoes et al., 2011; Zanto et al., 2011). Aside from the encoding stage, PFC is also suggested to play a similar modulatory role in the preparatory (i.e. expectation) and the maintenance phases of WM (Bentler, 1990; Bollinger et al., 2010; Clapp et al., 2010; McNab & Klingberg, 2008; Sreenivasan & Jha, 2007). Thus, accumulating evidence supports the existence of a top-down mechanism that can serve to bias representations in WM via both neural facilitation and suppression in a manner comparable to that during perceptual processing (SA). The potential role for PFC in generating top-down biasing signals should all be caveated with the side note that top-down signals originating from subregions of PFC are likely to represent *different task-related variables* which could have different influences and behavioural consequences, depending on the receiving area (D'Esposito & Postle, 2015). In reality PFC covers a large and heterogeneous set of regions, with a hierarchical organization, making it likely that they support different levels of representational abstraction (for detailed discussion, Badre, 2008; Badre & D'Esposito, 2007; Koechlin & Summerfield, 2007; Rottschy et al., 2012).

1.6 Summary and foreshadowing the rest of the thesis

This chapter has provided a comprehensive, if not exhaustive, summary of what is known about the interplay between SA and WM. Since the first theoretical accounts of WM, some kind of vaguely describe attentional resources has been incorporated. More recent theoretical accounts postulate SA as a critical process within the short-term maintenance of information in WM. Behavioural studies have broadly shown that these two processes can go hand-in-hand, using dual-task, individual differences and cueing paradigms. These behavioural findings have been mirrored by a growing neuroimaging evidence base. Over the past decade this literature has gone beyond demonstrating overlapping neural correlates for SA and WM, to demonstrate that they share capacity-limiting neural resources. One candidate mechanism is top-down modulation. The goal-oriented biasing of lower-level representational activity within sensory processing areas, either in the service of selection or short-term maintenance, may provide some unifying framework to understanding how these two processes share

neural architecture. Areas within the PFC seem like good candidates for the origins of these biasing signals.

A key assumption of all the work described above is that the relationship between SA and WM, in terms of either behaviour or neurophysiology, is constant and ubiquitous. Experimental studies test whether SA and WM interact or not, neuroimaging studies test whether they share resources or not. But what if the answer to these questions is variable across participants? Could it be that for some participants largely shared mechanisms are deployed for both SA and WM, whereas for other participants these processes are more independent? These questions are largely unanswered within the field to date and provide the starting point for the empirical work in this thesis. This notion of mechanistic heterogeneity becomes more likely as tasks become more complex. A complex task like maintaining information for brief periods of time may be achievable via multiple mechanistic routes. This has not been considered, with most studies relying heavily on group average performance or neural activity.

The first empirical chapter of the thesis focuses on the development of carefully matched behavioural paradigms for measuring SA and WM, and individual differences in the performance of these tasks. The second empirical chapter identifies behavioural subgroups, who subsequently undergo an fMRI study to test whether the process-critical neural correlates couple differently depending upon participants' behavioural profiles. The third chapter uses a form of computational modelling – recurrent neural networks (RNNs) – to explore the overlap between SA and WM within an artificial system, and then relate these processes to neural activity from the previous chapter. The final chapter is a cognitive training experiment, in which participants with different behavioural profiles are trained in order to use patterns of transfer to test for causal relationships between SA and WM.

Chapter 2: Task development and behavioural validation

2.1 Introduction

This chapter focusses on the measurement of two different cognitive processes. The first is the goal-directed selection of information relevant to a task at hand, at the exclusion of irrelevant information, namely selective attention (SA). The second is the maintenance, and in some cases the additional manipulation of, information for brief periods of time, namely working memory (WM). It is widely acknowledged that these two processes are highly interdependent (Awh & Jonides, 2001; Chun, 2011; Gazzaley & Nobre, 2012). They may interact as we anticipate, encode, maintain and retrieve memory contents (Bollinger et al., 2010; Gazzaley et al., 2004; Kuo et al., 2009; Rutman et al., 2010). One process may act as a capacity limit for the other (Cowan, 2001; Wheeler & Treisman, 2002). And to take a more extreme position, the content in WM may reflect the limited capacity of an actively controlled attentional system for objects or events critical to ongoing tasks (Cowan, 2001). These various options for the nature of the SA / WM interdependence, and a series of others, are outlined in the previous chapter, and will not be repeated here.

Research in this area typically relies on a set of classic tasks, such as Stroop, attentional capture and visual search to measure SA and complex span, n-back and serial-recall tasks for WM. These complex paradigms can also span multiple modalities of sensory input (e.g. verbal material, visual objects and spatial locations, etc.), and multiple different response types (e.g. speeded reaction times, span measures, accuracy differences etc.) adding to the heterogeneity of tasks. This makes it very difficult to draw direct inferences about the nature of, and the variance shared by, SA and WM. In other words, it is really hard to know how interrelated these processes are behaviourally let alone mechanistically, because the paradigms we use to measure them are so different. In this chapter we developed a pair of tasks based on traditional SA and WM measures, in an attempt to remedy this.

To date, there are broadly three experimental approaches for exploring the interplay between SA and WM. Firstly, dual-task setups, in which participants perform carefully designed and calibrated tasks, in search of specific or general interference effects that would speak to the presence and nature of any shared processes (e.g. Burnham et al., 2014; Lavie & De Fockert, 2005) Secondly, by examining individual differences in performance, testing for whether one ability correlates with another, or whether subgroups formed with one measure will demonstrate performance differences on another task (e.g. Lustig et al., 2001; Sobel et al., 2007; Turner & Engle, 1989). And thirdly, cueing paradigms, in which attention can be oriented to locations, features or items at different moments in WM processing (Griffin & Nobre, 2003; Landman et al., 2003).

Taken together, these main strands of research suggest that there is substantial cognitive and behavioural overlap between WM and SA. However, in each case there is a core underpinning

assumption – that the degree of overlap is common for all participants. That is, there is an assumption that SA and WM are either related or unrelated, and not that these domains can be variably related across participants. This is because findings are almost always interpreted at the group level. This view is in stark contrast with the developmental perspective on cognition, in which the organization of the cognitive architecture changes in a dynamic fashion across life span (Baltes et al., 1999). For example, the strength of coupling among intellectual abilities and cognitive processes is more pronounced during childhood and old age, whilst is lower in the middle stage (Li et al., 2004), and is accompanied by changes in neurobiology, such as fine-tuning of brain regions associated with a given cognitive task as age increases (Casey et al., 2005; Tamm et al., 2002). Similarly, it seems plausible that there could be variance in the strength of coupling between cognitive functions across individuals, regardless of age. This is important because it speaks to the potential mechanistic heterogeneity of complex cognitive processes like SA and WM. Almost everything, we know about the relationship between these two constructs is derived using group-average performance. Whereas in reality some individuals may rely on shared sub-processes to perform tasks more so than others. Whilst this has not been explored directly, some studies probing age-related effects on attentional modulation in the context of WM have shown that there is a deficit in suppressing irrelevant information in older adults, but not in enhancing relevant information (Gazzaley et al., 2005; 2008). This suggests that there can be substantial variability in the cognitive processes that participants use to support WM. But does this happen in the wider population? As cognitive neuroscience advances towards understanding brain functions and neurocognitive impairments associated with SA and WM, the heterogeneity in the relationships between the processes should be explored.

The present study

One caveat to the individual differences approach reviewed earlier is that other than the core construct measured by the tasks, other aspects could also be the source of performance variance. For example, Kane et al., (2006) examined the relationship between WM capacity (WMC) and visual search, using an operation span task to screen for high and low capacity individuals. In addition to a latent construct related to WMC, individual differences were also dominated by task-specific variance particular to the paradigms (i.e. the procedure of performing a memory task with a secondary task) and the contents (i.e. the use of words as memory targets and mathematical equations in secondary task). Since the SA task used in the study did not share these task-specific features, it is possible that variance was misattributed to WMC. As a result, in three subsequent experiments using visual search paradigms with different task features, no effect of WMC was found. Moreover, although this correlational method has identified a group-average relationship between individual WMC and SA performance, it cannot be used to study variance in the strength of the SA-WM relationship. This is because by chopping the WM distribution

in order to produce the subgroupings, the potential variability (both behavioural and mechanistic) in WM performance has been artificially constrained.

The purpose of this initial study is to extend the previous work in two ways. First, we aimed to develop paradigms tapping SA and WM that have carefully matched stimuli and timing. As will unfold across this thesis, these paradigms would subsequently be used for neuroimaging, computational modelling and training experiments. Second, we wanted to examine the variability in how coupled SA and WM performance is. Namely whether SA and WM are strongly coupled in some individuals, and only weakly coupled in others.

To foreshadow subsequent chapters in the thesis. Because the intention was to combine these behavioural paradigms with neuroimaging and cognitive training it is necessary that they have the same content material and motor responses (especially important for the scanner), in addition to both being independently scalable in terms of difficulty (especially important for the training study), and, of course, both being psychometrically robust. In summary, there are multiple constraints that we were attempting to satisfy when designing these tasks. For this reason we first spent substantial time piloting multiple variants of the two paradigms. This culminated in a formal in-house pilot study ($n = 30$), before moving to a large online behavioural study ($n = 664$). Owing to the disruption of the pandemic, there was not time to write up the various precursor experiments in the thesis. For simplicity we just report the final large online study here.

2.2 Methods

2.2.1 Participants

Six hundred and sixty-four participants (315 females; mean age = 29.71 year; SD age = 8.60 year) were recruited through the online psychological study platform Prolific. The inclusion requirements included being aged between 18-40 and a previous Prolific approval rate higher than 90%, to ensure only participants with a good track record of participation were recruited. Forty-seven of them were recruited for a second experiment designed to estimate the new measures' test-retest reliability, which involved completing the SA and WM tasks twice on two different days. Their initial data were combined with the others for the general analyses, and their subsequent data are used for the test-retest measures.

2.2.2 Stimuli, tasks and procedures

To achieve the goal of examining individual differences in the relationship between SA and WM we adapted classic tasks while matching visual presentation and motor responses as closely as possible.

SA task

The SA task was adapted from visual search experiments, which have been widely used to measure visual SA. They typically involve an active scan of a visual array for a particular object or feature (the target) among other irrelevant but similar objects or features (distractors) (Treisman & Gelade, 1980; Wolfe & Horowitz, 2004). Traditionally, the set size (i.e. total number of stimuli) is varied in order to estimate the slope of reaction time (RT) \times set size function as a metric of search efficiency, for a particular target-distractor combination. However, we opted for fixing the set size and instead changed the search difficulty in a different way. This is because one of our requirements was to keep the visual input consistent across both tasks thereby making it suitable for neuroimaging. Accordingly, changing set size would complicate the neuroimaging subtraction method, since the amount of visual information would be confounded with search difficulty. Therefore, we varied difficulty based upon the attentional engagement theory from Duncan & Humphreys (1992, 1989), who argued that search efficiency is the result of target-distractor (TD) and distractor-distractor (DD) similarities. According to their theory, the difference between the target and nontargets can be considered as a signal that must be found amongst the noise of the surrounding distractors. Consequently, search difficulty decreases as a function of TD difference (signal) and increases as a function of DD difference (noise). They also noted an interaction between the two axes: when distractors were highly similar to each other (or in extreme case there is only one type of distractor, DD difference is 0), TD difference will have little effect on search time (Duncan & Humphreys, 1989, Figure 2.3). Their own experiments as well as others have supported the claim, at least on dimensions of colour (Wolfe & Horowitz, 2004) and orientation (Duncan & Humphreys, 1989; Foster & Ward, 1991). In our new SA task, we decided to use a set of T-shaped stimuli and their orientations as the main axis for manipulation.

Participants were presented with an array of 3 by 3 items on a black background and had to determine whether a target was present or absent within 3 seconds, while ignoring the other nontargets. All stimuli were drawn in white colour (RGB value: 255 255 255) and the two perpendicular lines in the shape were of exact same length. The 3 by 3 array of stimuli were displayed around the centre of the screen, with the central one appearing exactly at the centre. The height and width of each shape was 50 CSS pixels and they were arranged so that the distance between any two adjacent shapes were 180 CSS pixels. Because this was an online experiment, screen size and monitor type were not within our control, but we specifically asked all potential participants to use either a desktop PC or laptop as their device, and they can only proceed with the experiment if keyboard input was detected. Otherwise we did not record the screen size, resolution or pixel per inch (PPI) therefore it was not possible to calculate the exact dimensions of the stimuli as appeared on their screens, or the view angle as the distance between head and screen were not under control. However, in practice these properties should not change drastically across people as most common monitors should have fairly similar PPIs. The target was set to be a capital letter T (with the vertical line extended from the midpoint of the horizontal line), and the distractors were visually similar shapes, but the vertical lines started from off-centre positions of the

horizontal line. Each trial always contained two types of distractors, differing from the target in both the horizontal bar position and orientations (Figure 1). It was designed so that if bar positions of Type A distractor were on the left to the midpoint, that of Type B distractor would be on the right. This was to ensure they were sufficiently different shape-wise so that search was difficult. Other than that, bar positions of distractors were jittered across trials so that it did not confound the main target orientation.

To vary search difficulty in a systematic manner, we altered the orientation-wise DD difference, whilst keeping the mean TD difference constant across trials. For example, if the orientation of a target was set to 0 degrees, one pair of distractors could have orientations of $(-20^\circ, 20^\circ)$ whereas another pair could have $(30^\circ, 10^\circ)$. Their mean TD difference was identical but the DD difference was different. This way search difficulty theoretically should increase with DD difference alone and not affected by TD difference on average. Importantly, the target was randomly oriented so that successful search depended only on identifying the shape. Thus, the procedure allowed us to change the search difficulty continuously without varying the set size (Treisman & Gelade, 1980; Treisman & Souther, 1985). The more homogenous the distractors (i.e. the lower the DD difference) the easier the search, the more heterogeneous the distractors (i.e. the higher the DD difference) the harder the search. In total, there were 273 trials in this task, which were divided into three equal length blocks between which participants were given a 1-minute break and they could choose to end it prematurely. 168 trials contained target and the rest did not. The DD difference levels were set these following values: from 9 to 63 degrees with 9 degrees increments. From our pilot it was decided that beyond this range increase of DD did not affect the search difficulty any further. This amounted to 39 trials per DD difference level with 24 trials having a target. Trials would end when key press was detected or at 3000 ms if there was no response.

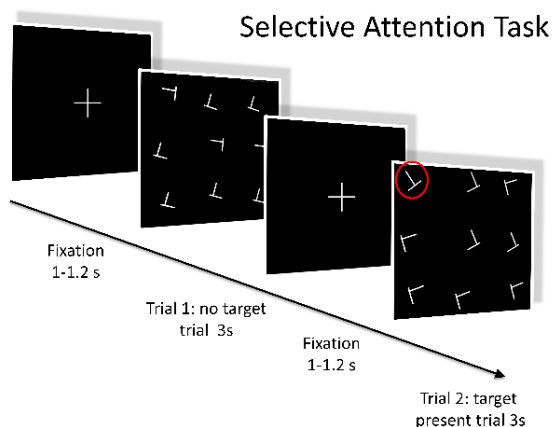


Figure 1. Schematics for SA task

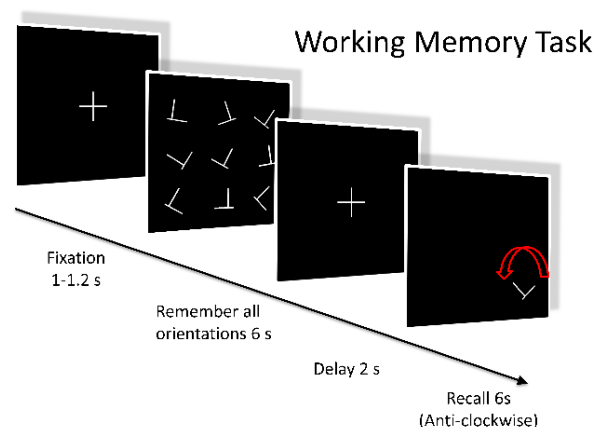


Figure 2. Schematic for WM task

WM task

The memory task was based upon change detection, a WM measure popularized by Luck and Vogel (1997), and a memory precision task popularised by Bays and Husain (2008). To match the SA task visually, participants were first presented a 3 by 3 array consisting of items drawn from the set of distractors used in the SA paradigm (Figure 2). They were asked to memorize all items' orientations in 6s. These orientations were determined randomly. After a 2s delay of a blank screen, a randomly chosen item is re-displayed in its previous location, rotated with an angle Δ chosen from (45 °, 30 °, 15 °, 5°), counter clockwise and clockwise. Participants reported the direction of rotation with two corresponding buttons for counter clockwise and clockwise. Every rotation angle was repeated 10 times, resulting in a total 80 trials. Based on previous studies (Bays & Husain, 2008; Ma et al., 2014), we expected that the bigger difference from the original orientation, the higher accuracy in the direction judgment. Furthermore, that varying offset should produce a psychometric function that corresponds to WM precision. The present procedure allowed us to keep the visual stimuli matched as closely as possible between the SA and WM tasks, and vary the WM demands orthogonally to the encoding demand (i.e. participants were required to encode all 9 items).

Comparison tasks

In addition to our main tasks we also designed some simple cognitive paradigms that have been variably labelled as SA or WM assessments within the wider literature.

Flanker task: This task required participants to indicate the direction of target arrows flanked by congruent or incongruent arrows on both sides (Eriksen & Eriksen, 1974). Attentional capture of the flankers is detrimental to performance when it is not consistent with target response, therefore attentional control and response inhibition are thought necessary to successful performance. A trial consisted of five white arrows in a line appearing in the centre of black screen. Each arrow had a width of 54 and a height of 58 CSS pixels, the spacing between two nearby items was 8 pixels. Maximal time allowed was 1000 ms per trial. Participants were given 40 trials in total with target direction (left vs. right) and congruency counterbalanced. Interference scores (mean RT for correct incongruent items – mean RT for correct congruent items) were calculated.

Spatial span task: This task assesses visuospatial WM capacity (Kessels et al., 2000). Participants saw a 5 x 5 grid of squares on screen that flashed red one at a time and were instructed to remember the order of the sequence. Each square was 60 CSS pixels in both width and height, and the size of the resulting grid is 334 pixels in both dimensions. Each At the end of a trial, they were asked to repeat the same sequence by clicking on the squares. It started with 3 items and incremented by 1 each time

participants recalled successfully and stopped at 3 consecutive errors. Their performance was measured by the maximum number of locations successful recalled.

Colour change detection: Change detection is a popular test purported to assess visual WM storage capacity (Luck & Vogel, 1997) and is more similar to our own WM task. A set of colour squares ranging from 2,4,6 or 8 was displayed randomly on an invisible circle around the centre of a black screen for 1000 ms and after a 1000 ms delay one probe reappeared at a studied location. Participants must make a judgement about whether the colour of the item was same or different from before. There were 6 candidate colours: red (RGB value 252 13 27), white (255 255 255), blue (16 114 189), pink (252 19 104), green (26 175 84) and yellow (255 253 56). No more than two instances of the same colour could appear in one trial. Each square was 50 CSS pixels long and wide and the invisible circle the shapes appeared on had a radius of 250 pixels. The shapes' locations were randomly decided but they were always equally distributed on the circle. Memory capacity was estimated with Cowan's K (2001) using $K = N * (\text{hit rate} - \text{false alarm rate})$ where N is the set size and the maximum K across set sizes was taken as individual's WM measure.

2.2.3 Analyses

Task performance measures

SA task: For this RT-based task we wanted to derive metrics beyond simple mean RT and response accuracy. First, outliers were excluded if their average RT was below 300ms, which was very unlikely given the required perceptual processes. To confirm the validity of the search difficulty manipulation, we fitted a linear regression model to the average RT, with DD difference level (i.e. search difficulty) as the independent variable. Next, to estimate how much RT was prolonged as a function of increasing search difficulty, a linear mixed effect model with participants as random effects was used to estimate each individual's performance from unaggregated data (i.e. all target-present trials were used instead of the mean). Mixed effect modelling is a framework that can account for uncertainty in behavioural data at both individual and group levels, therefore it is a natural solution to traditional designs where trials/observations are nested within individuals (Haines et al., 2020). Individual slopes can then be extracted from the fitted model, which reflect how each participant's RT was influenced by the increasing DD difference. In short, how much slower do participants get as the distracters become more heterogeneous? The shallower the slope, the less one's ability to suppress distractors was affected by search difficulty, and vice versa. For this step, log-transformed raw RT data were used, as the distribution of raw RT was skewed and log-transformation ensured the model residual would be normally distributed, an assumption of multilevel models (Schielzeth et al., 2020).

WM task: This task tested participants' ability to remember orientations of multiple items and the memory precision associated with it. Exclusion of outliers was based on the same RT threshold as the

SA task. We also decided to drop participants who achieved percent correct below 50% on both practice and actual trials, as this could indicate a failure to understand the task instructions, given that theoretically it should be 50% just based on pure guessing. A mixture model similar to that deployed by Bays and Hussain (2008) was fitted to individual's response data, containing a cumulative Gaussian distribution that corresponded to the distribution of error in the stored representation of the probed item and a uniform distribution that corresponded to the probability of guessing. Maximum likelihood estimation was used to optimize model parameters: standard deviation (σ) of the Gaussian distribution and guess rate, the inverse of the former ($1/\sigma$) can be used as an index of individual's memory precision, since larger μ indicates more uncertainty around an item representation.

As a post-hoc decision, an additional metric was included, because we encountered a problem in the original model fitting procedure: for some participants the proportion of variance explained by their models (R-squared) appeared to be negative. Further investigation suggested that due to the high difficulty of using a 9-item array, these individuals have "failed" a high proportion of the trials, resulting to a high guess rate estimation and an unreliable estimation of μ based on only a few trials. To circumvent this issue of one metric depending on the other, we computed Kullbeck-Leibler divergence (KL divergence; Kullback & Leibler, 1951) which can be estimated by the two parameters from the above mixture model. In essence, it is a measure of how one probability distribution differs from another reference distribution and we used it to quantify the distance between the probability density function that best fitted the response data (i.e. the mixture model) and pure guessing. In other words, it measures the information gain when a participant remembers an item instead of randomly guessing its orientation (Dalmaijer et al., 2018), hence those with high guess rate should yield a KL divergence closer to 0. It is a proxy of combining both standard deviation and guess rate and should lead to a more reliable measure for WM task.

Test-retest reliability

To measure the consistency of the new paradigms over time, 47 participants were recruited online to complete the SA and WM tasks twice on two different days. The test-retest reliability would be computed using intraclass correlation (ICC), of which the version (3, k) was chosen since our situation granted a 2-way fixed-effect model (Weir, 2005). Where multiple measurements can be derived to quantify performance, such as overall accuracy, precision, guessing rate and KL divergence from the WM task, we included all of them to offer a full picture of the psychometric properties of the new tasks, from which the most suitable index was used in subsequent chapters.

Comparison with other SA and WM assessments

The validity of the new tasks was assessed against some well-established behavioural measures, namely the flanker task for SA ability, and spatial span and change detection tasks for WM ability. Pearson's correlations were calculated to assess the between-task relationships. Although even at the outset we recognise that this is challenging, since despite being described as measuring particular constructs, it seems clear that these tasks measure many different things.

Identifying subgroups with distinct SA and WM profiles

SA and WM measures are often correlated at the group level, but are there subgroups whose abilities are not matched proportionally on the two scales? For this purpose, we sought to identify individuals with performance falling within the first or last quantile of the entire sample. Specifically, 4 groups of participants were explored: highSA+highWM, highSA+lowWM, lowSA+high WM and lowSA+lowWM. Having identified these different profiles, we would then explore the relationships between the SA and WM tasks within each subgroup, as well as sub-group differences in performance on the other tasks. Multiple comparisons were corrected using the Bonferroni method. Correlation matrices, or differences therein, cannot be directly interpretable without some formal tools for quantifying the differences in correlation structure between groups. For this purpose, we used multi-group confirmatory factor analysis (MG-CFA; Hirschfeld & Von Brachel, 2014) and examined the measurement invariance of the factor model across groups. Specifically, we first identified the best-fitting model to the entire dataset through CFA, then nested models with increasing parameter constraints were added and each compared to its preceding less restricted model. This step usually involved four levels of progressively restricted models: a) all parameter estimates were free to vary across groups including factor loadings, item intercepts, factor variances, and covariances (*configural invariance*); b) factor loadings from each observed variable to factors were fixed to be equal across groups (*metric invariance*); c) intercepts from observed variable to factors were fixed too (*scalar invariance*) and d) finally residual variance was constrained too. The idea is that if a more restricted model did not change goodness-of-fit compared to the less restricted one, different groups did not vary in terms of the constraints imposed by the more restricted model. To give a concrete example, if the highSA+highWM group showed a good fit of the configural model, it implies that this group and the baseline sample (those who do not belong in any groups) have the same number of factors and the same pattern of connections between individual tasks and factors. Configural invariance is a prerequisite for the other tests (Steinmetz et al., 2009). Passing the metric invariance test would suggest the causal effect of the latent factor on its connected measures is equal across groups and is necessary to infer that constructs have the same meaning. Lastly, scalar invariance and residual invariance would imply same response bias and test reliability across groups. The first two levels of test will be where we focus on because difference in strength of SA and WM coupling will likely alter factor structure or the strength of factor loadings if such differences really exist.

Furthermore, model comparison was conducted with chi-square test and changes in comparative fit index (CFI; a change $>.01$ is taken as evidence for model difference (Cheung & Rensvold, 2002). Decision rules for goodness-of-fit of CFA models are not hardbound, but recommendations set forth in the literature involved examining several fit indices including a chi-square (χ^2) test against a baseline model with $p>.05$ (significant difference between the tested and the baseline model); comparative fit index (CFI) $>.90$ for acceptable fit and $>.95$ for good fit; Root Mean Squared Error of Approximation (RMSEA) $<.08$ (Bentler, 1990; Steiger, 1990).

2.3 Results

2.3.1 Performance on the new SA and WM tasks

SA task: The averaged accuracy was relatively high (84%. Table 1). There was a significant difference in RT between target- present ($M = 1479.86$ ms, $SD = 226$ ms) and -absent trials ($M = 1928.48$ ms, $SD = 342$ ms), $t(663) = 66.04$, $p<.001$, suggesting it took longer to respond when the display only contained distractors.

A regression model was fit to investigate the effect of altering search difficulty on RT. Only correct target-present trials were included. We assumed a linear model in which the independent variable significantly predicted the mean RT ($p <.001$, $\beta = 371$) and explained 89% of variance, confirming the effectiveness of the difficulty manipulation. This suggests a unit change in the independent variable (DD difference) corresponded to a 371 ms increase in search time. Next, a linear mixed-effect model was fitted to the log-transformed raw RT data with search difficulty as fixed effect and included random slopes for search difficulty by participants. Estimates of individual's slope can be extracted, which was an approximation of how much one's RT increased as a function of search difficulty. For fixed effects: Intercept = 3.07, $CI = 3.075-3.080$, $p<.001$; search difficulty $\beta = 0.10$, $CI=0.8-0.133$, $p<.001$; For random effect: $\beta =0.11$, Standard error = 0.04, $p<.001$. To note, the unit of the estimated parameters was not in millisecond anymore as RT has been log-transformed. Most importantly, we obtained individual RT slopes ($M = 0.10$, $SD = 0.33$) from the model, which quantifies an individual's ability to ignore distractions (Figure 3). Plotting the individual slopes for each participant provides one way of representing variability in the impact of our difficulty manipulation. Most participants showed a positive relationship between RT and search difficulty, while a minority showed a negative to zero relationship.

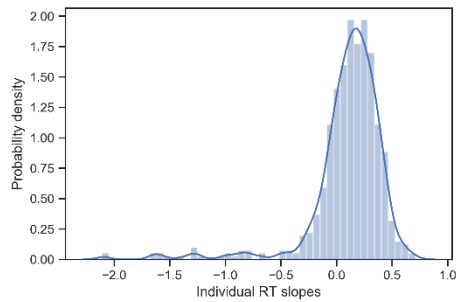


Figure 3. Distribution of search slopes of log-transformed RTs

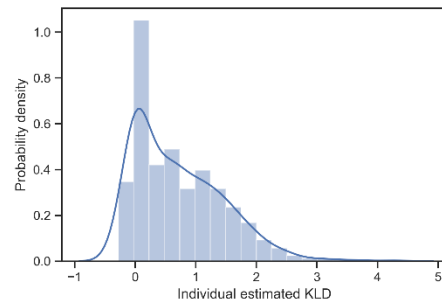


Figure 4. Distribution of KL divergence for WM task.

WM task: Overall accuracy of the WM task was lower, relative to the SA task but with comparable variation (Table 1). Accuracy increased as the degree of change became larger ($M = 0.55, 0.62, 0.69, 0.73$, at $\pm 5^\circ, \pm 15^\circ, \pm 30^\circ$ and $\pm 45^\circ$, respectively). A repeated-measures ANOVA suggested a difference between the angles of rotation, $F(3,1914) = 332.17, p < .001$, and post-hoc tests showed all differences between two given levels were significant (all $ps < .001$).

The fit of the mixture model to data aggregated across all participants was good (Figure 5). However, it should be noted that although for significant number of participants the model explained more than half of the variance in the data, as noted in Methods, when participants had to resort to guessing in a large number of trials, estimation of memory precision ($1/\text{estimated } \mu$) was unreliable and overall model fit (R-squared) was poor (Figure 6). Moreover, guessing rate and precision were positively correlated ($r = .039, p < .001$), suggesting the more participants guessed the more precise their memory was, which is unlikely and counterintuitive. A logical explanation is that as trials are increasingly tagged as part of a guessing distribution, those that are left appear to be part of a tighter response distribution – in short if participants guess a lot then what is left looks like a rather accurate response. Alternatively, if the guessing rate is high, the solution for the precision estimate was so unstable that it tends to be exaggerated. This is reflected in the negative R-squared of individual models, which showed that in 143 participants the model failed to outperform merely using the sample mean. For this reason, we used the alternative metric KL divergence, which negatively correlated with guessing rate after controlling for covariance with precision (partial $r = -0.71, p < .001$), and positively and uniquely correlated with precision (partial $r = 0.40, p < .001$). This indicates KL divergence was a good metric for representing the combined parameters and avoiding the problematic trade-off of parameters that can arise in the standard mixture model. Distribution of the KL divergence was shown in Figure 4.

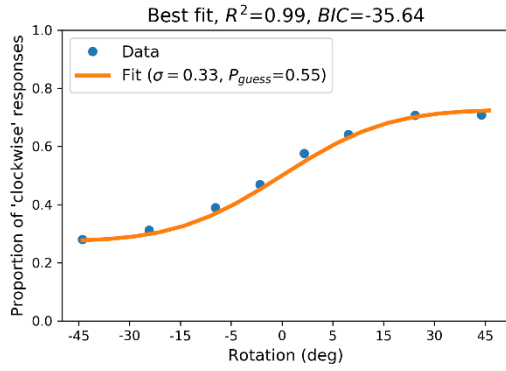


Figure 5. Mixture model fit to all participants' data.

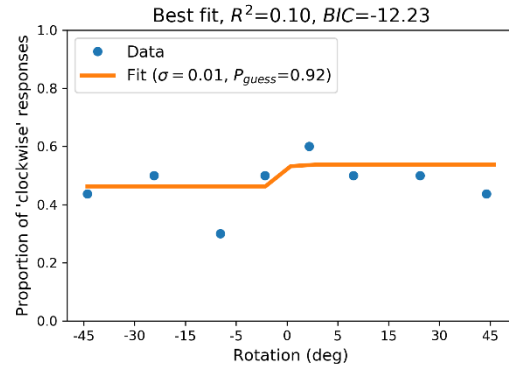


Figure 6. Model fit to an example subject, whose performance was poor. Note that how when guess rate (P_{guess}) was estimated high, precision ($1/\sigma$) was also high.

Table 1. Descriptive statistics of all tasks.

Task	N	Measure	Mean (SD)
1.SA task	647	Accuracy	.84 (.09)
		RT (ms)	1629 (598)
		Search slope	0.10 (0.33)
2.WM task	639	Accuracy	.65 (.12)
		RT	1624 (1017)
		Guess rate	0.48 (0.33)
		Precision	15.71 (31.06)
3.Flanker	514	KL divergence	0.71 (0.74)
		Accuracy	.91 (.15)
		RT cong (ms)	476 (76)
4.Spatial span	539	RT incong (ms)	558 (84)
		Max span length	5.9(1.48)
5.Change detection	522	Accuracy	.81 (.10)
		Cowan's K	2.7 (1.1)

Note. RT = reaction time; cong = congruent trials; incong = incongruent trials

2.3.2 Test-retest reliability

Data from the same group of participants who completed the tasks on two separate days were entered to assess the consistency of the measures. For SA task, ICC (3, k) for individual RT slope was 0.62, $p < .001$, 95% confidence interval (CI) = 0.32-0.79. For precision and guessing rate from the WM task, $ICC_{precision} = 0.11$, $p = .34$ and $ICC_{guess} = 0.71$, $p < .001$, 95% CI = 0.49-0.84; ICC for the combined measure KL divergence was 0.53, $p = .006$, 95% CI = 0.15-0.74. According to Koo & Li (2016), RT slope, guessing rate and KL divergence had moderate strength of agreement whereas memory precision had very poor reliability. This confirms the observation that estimation of precision was unstable as a result of increased guessing. However, KL divergence seemed to perform reasonably well in capturing unique variance of the precision and guessing rate while preserving an acceptable level of test-retest reliability.

2.3.3 Comparison tasks and between-task relationships

For the Flanker Task, participants achieved a high accuracy ($M = 0.91$, $SD = 0.15$); RT in congruent trials ($M = 496$ ms, $SD = 76$ ms) was faster than incongruent trials ($M = 558$ ms, $SD = 84$ ms; $t(529) = 30.64$, $p < .001$). The difference between averaged RTs for correct incongruent and correct congruent trials was computed. For spatial span task, participants can remember an average 5.91 locations ($SD = 1.48$). Lastly, for the change-detection task we adopted Cowan's K (Cowan et al., 2001) to compute each person's WM capacity. It yielded mean K of 2.68 ($SD = 1.07$).

Pearson's r was calculated among all task measures (Table 2). For easier interpretation, all values were firstly normalized across participants to have zero mean and standard deviation of 1, and RT based measures in SA and flanker tasks were sign-flipped so that higher value indicates better performance. SA task was weakly but significantly correlated with the WM task and colour change detection, but surprisingly this relationship was negative. SA and flanker tasks were weakly correlated, but this effect was not statistically significant ($p = .07$), possibly reflecting the fact that they draw upon distinct sets of cognitive processes even though both were claimed to assess "selective attention". The new WM paradigm seemed to relate to other traditional measures moderately, demonstrating that the new tasks at least partly tap into the same memory processes as change detection and spatial span tasks.

Table 2. Pearson's correlations between representative measures of all assessment tasks

	1	2	3	4	5
1.SA task (search slope)					
2.WM task (KL divergence)	-.13*				
3.Flanker (RT incong-cong)	-.08	.05			
4.Spatial span (span length)	.04	.32**	.08		
5.Change detection (K)	-.12*	.28**	.03	.26**	

Note. * $<.01$, ** $<.001$. cong = congruent trials; incong = incongruent trials

2.3.4 Subgroups

From 460 participants who had complete data points across all assessments and using the upper and lower quartiles as a threshold, four groups of individuals with distinct profiles were formed: highSA+highWM (N=20), lowSA+lowWM (N=27), highSA+lowWM(N=40) and lowSA+highWM (N= 28). These groups necessarily differed on our SA and WM measures. But there were also significant differences on the other tasks: performance on change detection ($F(3,111) = 4.09, p<.01$) and spatial span ($F(3,111) = 3.88, p =.01$), significantly differed according to group. But the groups did not differ on the flanker task ($F(3,111) = 1.63, p =.19$). Post-hoc comparisons revealed that for change detection highSA+highWM scored significantly higher than the highSA+lowWM and lowSA+lowWM ($t(58) = 3.19, t(45) = 3.27; p_{corrected} <.01$ for both). LowSA+highWM was marginally higher than the lowSA+lowWM and highSA+lowWM, but the effect was not significant after correction ($t(53) = 2.08; t(66) = 2.13; p_{corrected} =.06$). For spatial span, pairwise comparison tests did not find any significant difference. Further examination of the groups showed that this could be due to the large variance within the highSA+highWM group (Figure 7). Within each group, between-task correlations were also calculated (Table 3). As expected, there were some differences in specific pairs of tasks but they are hard to quantify without a formal statistical tool.

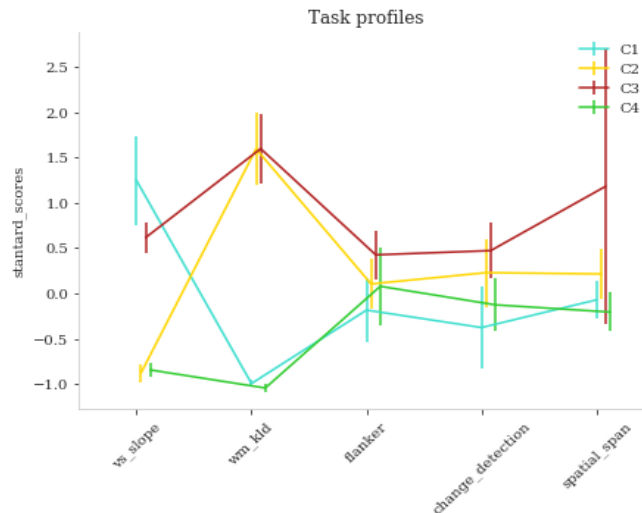


Figure 7. Profiles of subgroups. C1 = highSA+lowWM; C2 = lowSA+highWM; C3 = highSA+highWM; C4 = lowSA+lowWM

We applied MG-CFA to test whether factor structure and parameters are consistent across groups. Firstly we found the model that best fit the entire sample ($N = 460$) as the general model (Hirschfeld & Von Brachel, 2014). This step showed that a 3-factor model with the 3 WM tasks loading onto the same latent variable and the SA and flanker tasks each loading onto a separate variable performed the best ($\chi^2 = 9.18$, $p > .05$; CFI = 0.95; EMSEA = 0.053; Figure 8). Next, measurement invariance of each identified subgroup was tested against the baseline sample who did not belong to any groups ($N = 345$). If the subgroups were indeed part of the baseline sample, a restricted model in which model parameters were fixed to be equal across groups should fit as well as the non-restricted model. The results were presented in Table 4. HighSA+highWM and highSA+lowWM showed a good configural invariance but did not pass the test for metric invariance, indicating that the subgroups shared the same factor structure as the entire sample, but the factor loadings diverge. This would suggest that the tasks measure the same thing in each subgroup, but that the loadings vary significantly within subgroups relative to the full baseline sample. A closer look at the Modification Indices revealed that change detection and spatial span tasks loaded onto the SA factor in highSA+highWM group – in other words, when participants are good at both SA and WM tasks, their underlying factor loadings are somewhat different to the overall sample, with ‘memory and ‘attention’ tasks all loading onto a common factor. Table 5 shows modification indices for constrained parameters in the metric invariance model for the highSA+highWM group and the increase of model fit if they are free to vary across groups. It suggests that if the model allows separate estimations between this group and baseline sample for the factor loading of SA task to the latent variable SA and that of WM task to the latent variable WM, model fit would be improved. Next, LowSA+lowWM and lowSA+highWM individuals also showed a good metric invariance, meaning both the factor structure and the strength of factor loadings were equal to

those in the baseline sample, however the intercepts (i.e. group means) differ. All in all, the MG-CFA results suggest that all subgroups manifested the same factor structure but strayed on either the loading strength or the intercepts, compared to the baseline sample. This implies the difference could come from task relationships in the former, or group means in the latter.

Table 3. Pearson's correlation between tasks within each subgroup.

	1	2	3	4	5	1	2	3	4	5
1.SA task										
2.WM task	-0.04					.21				
3.Flanker	0.1	-.05				.03	.08			
4.Spatial span	0.01	.06	-.17			-.01	.02	.06		
5.Change detection	-.05	-.10	.33*	.13		.23	.34	.20	.31	
1.SA task										
2.WM task	-.20					-.20				
3.Flanker	.03	-.01				-.44*	.26			
4.Spatial span	.04	.07	.36			.37*	-.23	-.26		
5.Change detection	-.14	.72*	.13	.20		.20	.01	-.32	.29	

Note. Between-task correlations by group. Blue= highSA+lowWM; yellow=lowSA+highWM; red = highSA+highWM; green = lowSA+lowWM. * $p < .05$

Table 4. Multi-group CFA model fit indices and model comparison statistics

Group	N	<i>Nested model</i>	χ^2 (df)	RMSEA (90% CI)	CFI	Model comparison	<i>P</i> values
highSA+highWM	20	Configural	6.66(8)	0.0 (0-.07)	1.0	NA	NA
		Metric	32.40(10)	.11 (.07-.15)	.74	Configural vs. Metric	$p < .001^{**}$
LowSA+lowWM	27	Configural	10.22(8)	.04 (0-.10)	.97	NA	NA
		Metric	14.50(10)	.05 (0-.10)	.95	Configural vs. Metric	$p = .11$
		Scaler	31.23(12)	.09 (.05-.13)	.77	Metric scaler vs.	$p < .001^{**}$
highSA+lowWM	40	Configural	3.56(8)	0.0 (0-.03)	1.0	NA	NA
		Metric	11.56(10)	0.03 (0-.08)	.98	Configural vs. metric	$p < .05^*$
lowSA+highWM	28	Configural	4.04 (8)	0.0 (0-.05)	1.0	NA	NA
		Metric	6.17(10)	0.0 (0-.05)	1.0	Configural vs. metric	$p = .34$
		Scaler	34.52(12)	.11(.08-.15)	.62	Metric scaler vs.	$p < .001^{**}$

Note. CFI: comparative fit index; RMSEA: Root Mean Square Error of Approximation

Table 5. Modification index for metric invariance model in highSA+highWM group

lhs	Op	Rhs	Group	MI	epc	Spec.lv	Spec.all
VS slope	~~	VS slope	highSA+highWM	11.984	-0.930	0	0
SA	=~	VS slope	highSA+highWM	11.90	-0.469	-0.127	-0.469
WM	=~	WM KLD	Baseline	7.492	-14.592	-5.591	-6.662
WM KLD	~~	change_detection	highSA+highWM	6.654	-0.392	-0.392	-2.353

Note. Lhs = left hand side. Op= operation. Rhs=right hand side. MI= modification index (improvement of model fit if the particular parameter is freed). Epc=expected parameter change. Spec.lv = only standardizing the latent variables. Spec.all= standardizing all variables; VS slope = SA task search slope; WM KLD = WM task KL divergence; SA= selective attention latent variable; WM= working memory latent variable. ~~ = error covariance; =~ = factor loading;

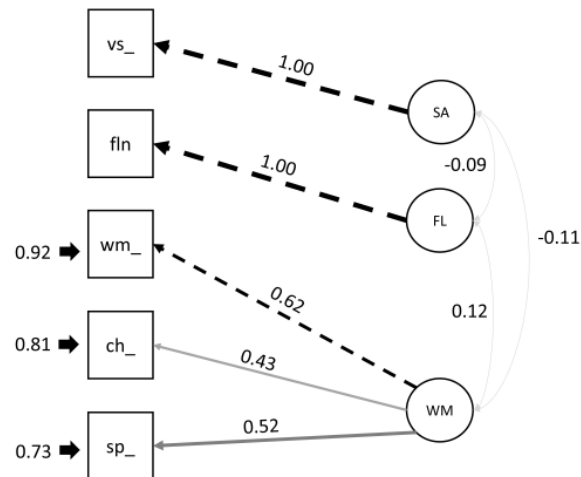


Figure 8. CFA baseline model.

Square = observed variables (tasks); circle = latent factors; vs = SA task; fln = Flanker; wm = WM task; ch = change detection; sp = spatial span; SA = SA factor; FL = flanker factor; WM = WM factor;

2.4 Discussion

In this first empirical chapter we tested the psychometric properties of the newly adapted tasks for assessing SA and WM, while attempting to match the paradigms on task-specific features. The SA task was based on traditional visual search, but task difficulty was achieved through varying distractor dissimilarity but fixing set size. Our modelling confirmed the effectiveness of this manipulation.

Furthermore, the search slope metric computed from individual's raw RT data appeared to be of moderate reliability and was assumed to capture the amount of change in the rate of processing relative to a unit change in search difficulty (Wolfe & Horowitz, 2017). This is generally consistent with other studies showing the test-retest reliability of various search tasks to be in the range of 0.5 to 0.7 (Pearson's r ; Van Wert et al., 2008), though the specific choice for metrics can affect this. The correlation result suggested no relatedness between the SA measure and flanker interference across individuals, indicating that the tasks tap different processes albeit both attentional. The flanker paradigm has often been used to capture "executive attention", which consists of both target selection and conflict resolution for successful performance (Posner & DiGirolamo, 1998). The incompatible flanking stimuli produce an automatic interfering response that must be suppressed by some putative control process, whereas during visual search it is less clear whether such response inhibition is really needed (Vandierendonck, 2014). In addition, the metric used in Flanker task (RT cost) has low test-retest reliability around 0.4 (ICC; Hedge et al., 2017), and this itself may contribute to the lack of correlation we observed. In retrospect, choosing another task that assesses a more similar construct to the SA task or using multiple tests for a latent variable approach would have been a better solution for a validity check. That said, such is the heterogeneous nature of so-called attention tasks it is possible that no paradigm would really count as a 'gold-standard'.

The WM task showed a moderate consistency and a good construct validity compared to other common paradigms. The results are broadly similar to those from the WM precision literature; the larger the change from the original items the more accurate response (Bays & Husain, 2008; W. Zhang & Luck, 2008); at high set sizes memory representations are noisy and target items remembered less precisely resulting in more difficulty differentiating between probe and memoranda (Luck & Vogel, 2013). We first attempted to use a mixture model to estimate precision and guess rate of each individual. Such a model relies on the assumption that overall response variances can be decomposed into several separate sources, including correctly remembered and forgotten items, which reflect distinct psychological processes. However, a problem arose because model fitting was unsuccessful in a subset of participants, particularly those with very few correct trials. One plausible reason could be that if guessing rate was estimated to be high, most trials were ascribed to the "forgotten items" distribution, leaving only a few trials attributed to the "remembered items" distribution and consequently a less reliable estimation of precision. An alternative metric (KL divergence) was used instead to measure memory performance. This makes more sense from a psychometric perspective, as well as theoretically, according to recent debate within the WM precision field about the validity of the mixture model. Schurgin and colleagues (2020) demonstrated that when perceptual similarity of feature space (e.g. orientation or colour) is taken into consideration, precision and guess rate were no longer distinct contributions to memory errors and rather can be estimated by a single memory-related parameter. They also showed that guessing rate and precision, as estimated by mixture model across various manipulations, change together. This suggests

a unitary process affecting both values, which bears resemblance to our results showing the positive correlation between the individual estimates. Regardless of which explanation provides the better account for our pattern of findings, using KL divergence as the final measure is sufficient to capture variance in task performance. As for the other WM tasks, it is not uncommon for different WM tasks to show a low or even zero correlation with one another (Redick & Lindsey, 2013). Despite this, Schmiedek et al. (2014) have shown that they can still measure the uniform construct on the latent level. In our case, the highest correlation was 0.32 between the span task and the new WM task. Our factor analysis suggested a single factor explaining a significant portion of variance in all three ‘WM’ measures. The results are also consistent with studies reporting similar strength of correlation between span-like and change-detection paradigms (Table 2, Unsworth et al., 2014).

There was no correlation between the SA and any of the WM measures, except a negligible one with new WM task, although this is tiny and negative. As to whether WMC is predictive of visual search performance, or vice versa, the findings are inconsistent at best. As mentioned in the literature review at the beginning of this thesis, studies adopting the high vs. low WMC participants report some evidence for differences in visual search, depending on the specific type of search (e.g. similarity between target and distractors; Williams & Drew, 2018). Several factors may play a role in moderating the relationship of these two sets of measures. First, according to Schwark et al., (2013), high WMC individuals are better at identifying low-prevalence targets, relative to low WMC counterparts, suggesting WM may be important for maintaining a representation of the target in the absence of a constant reminder. We used a high prevalence target design which may have reduced the need to hold the target template in mind. A second, not mutually incompatible explanation could be the choice of metric to quantify performance in the SA task. Williams and colleagues have shown that irrespective of target type, WMC only correlates with overall search accuracy, and not with mean RT or search slope (2018). Those authors argued WMC only predicts how well an individual can encode and maintain a target representation, not the scaling difficulty of the search process itself. A closer look into our own data revealed that visual search accuracy indeed was moderately correlated with the WM measure. Aside from these factors, there are other potentially complicating factors within the literature: CDA amplitude, a marker sensitive to number of items currently held in mind, is independently related to individual WMC and search difficulty (Gaspar et al., 2016; Luria & Vogel, 2011). Individuals with higher WMC encode fewer items into WM during visual search, variably so depending upon the difficulty of the search, potentially reflecting two separate processes that impact visual search performance (Luria & Vogel, 2011). This suggests that search difficulty may modulate the relationship between SA and WM, such that in easy conditions both low and high-capacity individuals are able to reject irrelevant information from WM, whereas group differences emerge when search gets more difficult. In short, the absence of a correlation between the SA and WM performances could be a product of multiple elements such as target prevalence, the sensitivity of the outcome measure, or genuinely independent processes.

Next, subgroups with differential performance on the SA and WM tasks, were identified using a hard cut-off score. Generally speaking, the occurrences of the profiles among the sample were roughly equal, except for the highSA+lowWM group which was bigger. This artificial “carving” of the distribution could result in genuinely interesting individual differences, or just an arbitrary slicing of measurement error. In subsequent chapters of this thesis you will see that some of the groups were invited back for further experiments. This gave us the opportunity to repeat the tasks and suffice to say these group differences still hold at a second time of testing. This is good evidence that this grouping is reflective of some genuine underlying difference in ability and not just randomly distributed noise.

We explored whether the relationship between SA and WM varies systematically across our subgroups. According to the CFA result, the highSA+highWM group showed a different variable-factor relationship from that which best fitted the baseline sample. This can be seen in the large drop in model fit index from configural invariance model to metric invariance model (the highSA+lowWM group also had significant decrease but effectively only a small one). This implies at least one factor loading is not equivalent between this group and the baseline sample. A closer look at the modification indices suggested that highSA+highWM and the baseline sample diverged mainly on the strength of association between SA and WM tasks and their respective latent variables, indicating that the amount of variance the tasks captured on the constructs was dissimilar across groups (Putnick & Bornstein, 2016). These measures are more closely related to the constructs in one group than in the other, perhaps a result of this groups relying on different sets of processes to perform tasks (or relying on the same set of processes to a varying degree). In contrast, the other groups appeared to only differ in the averaged latent score (intercept), which was likely to be caused by the difference in basic task performance. All in all, although this is not an ideal set-up for a latent-variable analysis due to the lack of multiple measures per factor, and the groups are not massive, the result nevertheless provides some insights into the underlying differences associated with the distinct subgroups. There seems to be a differentiation in the degree to which the SA and WM tasks reflect the latent constructs in highSA+highWM and highSA+lowWM, compared to baseline.

A valid point was raised about the influence of visual sensory ability on performance of the tasks, which has been shown to contribute to age-related changes in various cognitive function (Salthouse et al., 1996). The pertinent question here is whether visual acuity affected the SA and WM tasks differently. Considering that both presented highly similar display of stimuli, individuals’ ability to process sensory information should not be unevenly impacted. One small distinction between the tasks related to perceptual processing is that in SA task detecting the target from distractors depends on distinguishing the location of the vertical line in the T-like shapes (i.e. whether it was in the middle or with an offset), whilst in WM task this is not necessary. We designed the distractors in a way that the offset distance of the vertical line from the middle point of the horizontal line (i.e. target) was at least 20% of the entire

line length, therefore, detection of the stimuli identity should in principle fairly immediate. In addition, all participants throughout experiments reported in this thesis had normal or corrected-to-normal vision, which should further eliminate variance due to different levels of sensory ability.

There are several ways in which this study could be improved in hindsight. Firstly, more appropriate validation tasks for SA could have been included to better capture the processes tapped by the visual search paradigm. The flanker task is known to rely on selection as well as other cognitive control processes (Keye et al., 2009). Secondly, and relatedly, using multiple validation tests for a single cognitive construct would also work better for the MG-CFA, as it generally requires more than a single assessment to identify the shared variance across indicators and is less susceptible to measurement error with more tasks (Schmiedek et al., 2014). This would allow us to establish a better baseline model of task structure against which group differences could be tested. Another improvement regarding the design of the task is to balance the difficulty of the SA and WM paradigms better, so that more participants have a higher likelihood of remembering WM items instead of random guessing. Another point related to the design of the new tasks is whether they involve different oculomotor patterns, which may itself explain some differences in what each task measures. It is very likely that eye movements are not identical between the SA and WM tasks, with previous studies having shown that in the former participants usually make a few saccades before locating the target in serial search (Williams et al., 1997), which are guided by top-down and bottom-up processes (Wolfe, 2015). There is certainly no advantage to any particular eye movement strategy as target location is unknown to participants. In the WM task however, it is expected that all stimuli would be fixated in order to process and encode orientations, and that some participants have reported using grouping strategies if adjacent items faced similar directions. Therefore, whether one employed such strategy may have contributed to some variances in the WM task performance.

To sum up: this chapter describes two new SA and WM paradigms, based upon classic measures usually deployed within this field. There were multiple competing constraints upon the design of these tasks, including the overall visual presentation, response modality and difficulty. The large online behavioural sample reflects the end of a large amount of pilot testing. Both new tasks produce a moderately good test-retest reliability, although the correlations with other tasks was patchy. Within the sample we identified subgroups and used MG-CFA to test whether the relationships between the measures vary systematically. To some extent this was true – where participants are good at both novel tasks, their underlying factor loading suggests a differential task-construct relationship, relative to the overall sample, possibly reflecting dependencies on different subprocesses. In subsequent chapters we will explore how the mechanisms underpinning these two tasks overlap, and individual differences in that overlap, using neuroimaging, computational modelling and cognitive training.

Chapter 3: Individual differences in the shared neural substrate of selective attention and working memory

3.1 Introduction

The interplay between selective attention (SA) and working memory (WM) have been repeatedly observed in behavioural terms. As outlined previously, this can be seen in dual-task setups, group comparisons, cueing studies and correlational work (Anderson et al., 2008; Griffin & Nobre, 2003; Heitz & Engle, 2007). In the previous chapter we used the new paradigms to test whether individuals who differed in their SA-WM profile (e.g. achieving high scores on both tasks vs. high on one task and low on the other) also differed in the factor structures for these tasks. Different profiles of performance related at least partially to distinct underlying task relationships. The purpose of this next chapter is to understand the neural mechanisms that underpin individual differences in the interplay between SA and WM using functional magnetic resonance imaging (fMRI). There are vast neuroimaging literatures examining the neural correlates of WM and SA, respectively. There is also a substantial neuroimaging literature explicitly exploring their overlap (Gazzaley & Nobre, 2012). Much of this was outlined in the General Introduction, so this is only briefly summarised here.

So far the evidence suggests that an overlapping network comprising frontal, parietal and temporal regions, basal ganglia and the cerebellum, is involved in both processes (Corbetta et al., 2002; Jha & McCarthy, 2000; Tomasi et al., 2007). Both fMRI and electroencephalography (EEG) studies have shown that activity in the intraparietal sulcus (IPS) mirrors individual differences in visual WMC. Using a change detection paradigm with a secondary task to prevent subvocal rehearsal, Todd and Marois (2004) identified neural correlates with the maximum number of objects an individual could store (K). Activity within the left IPS and the intraoccipital sulcus reflected this individual variability with a moderate to strong effect size. Interestingly, the same analysis did not reveal this relationship in dorsolateral prefrontal cortex (DLPFC) or anterior cingulate, which are often implicated in WMC (Curtis & D'Esposito, 2003; D'Esposito & Postle, 1999). Xu and Chun (2006) further dissociated the function of posterior parietal cortex (PPC) by observing that inferior part reflected fixed capacity of four items regardless of object complexity, whereas the superior region although also had a similar storage capacity, represented less number when objects were complex. Other researchers have built upon this finding, for example by showing the preparatory activation (i.e. difference between the no-distraction and distraction conditions during instruction phase) in basal ganglia and middle frontal gyrus (MFG) was predictive of individual's WMC (McNab & Klingberg, 2008). McNab & Klingberg also observed a negative correlation between this preparatory signal in basal ganglia and the extent to which participants failed to suppress distractors, measured by activity in the load-sensitive IPS (Vogel & Machizawa, 2004). This implies some kind of subdivision within the neural architecture supporting WM, with some areas such as the basal ganglia exerting top-down modulation towards more load-

sensitive areas like the IPS. Across a range of studies (see also Edin et al., 2009)), it is postulated that mechanisms underpinned by various frontoparietal and subcortical structures are indexing the gatekeeping of WM encoding (i.e. prefrontal cortex and basal ganglia) and storage (i.e. IPS) may contribute towards variability in WM performance.

In parallel, individual differences in SA have been associated with similar brain networks to those identified in the WM literature. One approach is to correlate neural activity and participants' visual search times (e.g. Nobre et al., 2003). While posterior parietal regions including IPS is sensitive to both increasing frequency of distractor items and baseline difficulty, prefrontal regions only show correspondence to baseline difficulty, pointing to a differential involvements of brain areas in the visual search process. Anderson et al (2007) reported the selective involvement of right prefrontal cortex (PFC) in search condition which was difficult (inefficient) and with more distractors, in addition to activation in the frontal eye field (FEF) which was present across all search conditions. In general, SA in visual search activate extensive regions in the frontal, parietal and occipital cortices, proposed to underlie differential functions, e.g. allocation of spatial attention is controlled by priority maps in the IPS (Silver & Kastner, 2009) the FEF (Hagler & Sereno, 2006); whilst activity in the superior subregion of the parietal lobule was linked to shift of feature-based attention (Serences et al., 2004). Taken together, there seems to be some indirect evidence for common neural substrates mediating processing in SA and WM performance. However, this is somewhat circumstantial. Seldom do studies adopting this approach also consider both processes simultaneously within the same participants. The present study builds upon the previous behavioural findings in this thesis and uses fMRI to examine brain activity that co-varies with different properties of SA and WM. In Chapter 2 subgroups were identified whose ability to perform the SA task either matched or did not match with their ability to perform the WM task. Furthermore, task relationships were not uniform across these behavioural profiles. One possibility is that in some individuals SA and WM are more strongly functionally coupled or in other words, rely more on the same underlying cognitive processes (e.g. top-down modulation (Gazzaley & Nobre, 2012), or spatial rehearsal (Awh & Jonides, 2001)), relative to other individuals. That is, not only is there inter-personal variance in both cognitive functions, there could also be variance in how much they rely on the same set of cognitive/neural processes in an individual. This becomes more likely as tasks become more complex and requiring the coordination between many sub-processes. For example, Cokely and colleagues (2006) argued that variation in strategy use drives the allocation of executive resources, such as forming helpful associations between to-be-remembered items, and that this variability explains some apparent individual differences in WMC. Of course, this is just one instance where conscious decisions as to how to perform a task could change the relative importance of a basic process. In short, in the context of the tasks used in this thesis, it is possible that in some individuals SA could account for more variance in WM but less so in others, depending upon the degree to which the participant is drawing upon attentional control to support WM, implicitly or explicitly.

The aim of the next study is therefore two-fold: we intended to identify the overlapping neural correlates of SA and WM, using our behavioural paradigms; second, we wanted to determine whether the degree of overlap differs between a ‘coupled-process group’, who perform both tasks equally well, and the ‘dissociated-process group’ whose performance on one task does not mirror that on the second task. We also manipulated the WM and SA difficulties parametrically in order to determine whether the regions highlighted in both tasks truly reflect the varying demands.

In addition to a traditional group-level conjunction analyses, which allow inferences about the range of functions a brain area is involved with, we employed an individual subject-level technique called group-constrained subject-specific analysis (GCSS; Fedorenko et al., 2010; Julian et al., 2012). The purpose of the latter analysis is to reveal conjunctive activation in the same individuals, while accounting for the functional and anatomical heterogeneity across participants (Frost & Goebel, 2012). This will enable a more detailed analysis of the neural activity from the comparable brain region across tasks, relative to methods based simply on activation magnitude.

In summary, the purpose of this study is to reveal the neural substrates of the functional overlap between visual SA and WM, and then to test whether the degree and nature of this functional overlap varies across individuals according to their behavioural profile. We predicted that evoked activity in certain regions would be more similar in the coupled-process group than the dissociated-process group. Specifically, we expected that areas in the frontoparietal network outlined earlier in this chapter would show this most markedly, because they are widely implicated in both SA and WM.

3.2 Methods

3.2.1 Participants

140 healthy, MRI-eligible adults between the age of 18 and 40 (92 female/47 male/1 other gender), mean age=23.41 years (SD = 3.36 years) were recruited from the MRC Cognition and Brain Sciences Unit research participant panel. They were first screened online using behavioural tasks described in detail in the Chapter 2. Participants reported normal or corrected-to-normal eyesight, right-handedness and no history of neurological or psychiatric disorder. The study was approved by the Cambridge Psychology Research Ethics Committee (Reference number: Pre.2018.132).

Forty-two of them (24 female, mean age = 27.3 year, SD age = 5.6 year) were selected to take part in the fMRI study. Fifteen belonged to Group 1 (G1; 6 female, mean age = 26.3 year, SD age = 3.4 year) with poor SA and good WM performance. Nineteen belonged in Group 2 (G2; 14 female, mean age = 28.2 year, SD age = 6.8 year) with good performance on both measures. The eight remaining participants fell into neither group, and were only included in the analysis of common neural substrates in order to increase statistical power.

N.B., there were two possible behavioural profiles that we could have used for the dissociated-process group, but we chose to use the low SA/high WM (rather than the high SA/low WM counterpart) because the WM task is more difficult, and we worried that we may not get enough correct WM trials for the fMRI analysis if we choose participants that are relatively poor at this task.

3.2.2 Experimental design and procedure

We used the same stimuli for both the online screening tasks and the fMRI tasks, which are also the same as in the previous chapter, though the tasks were adjusted slightly for fMRI compatibility.

Experiments were implemented using JavaScript-based JsPsych package (De Leeuw, 2015) and JATOS server (Lange et al., 2015) for online screening. Participants were sent a link to access the experiments and instructed to complete the study on their own personal computers/laptops in a quiet environment with no external distractions. Of all participants who successfully finished the study, none reported any technical difficulty or problem with understanding the tasks.

We adapted the paradigms in order for both tasks to have similar trial lengths, stimuli presentations and motor response that are suitable for subsequent analysis. The major deviations from the previous version were that: 1) the presentation array changed from a 3 x 3 square into all stimuli equally distributed to appear on an invisible circle around the centre of the screen to equalize the probability of saccadic movements; 2) coloured borders were introduced to each stimulus to serve as a cue for memory target in the WM task (i.e. blue border indicated the item inside needed to be remembered and red indicated otherwise), which allowed us to vary memory load; and 3) a dummy retention and probe phase were added to the stimuli array phase of the SA task in order to match the presentation time and required movements in the WM task. Each trial began with the presentation of the stimuli array for 6 s (Figure 9), followed by a delay interval (retention phase) jittered between 2, 3 or 4 s showing a fixation cross and 6-s probe phase when one of the original items was shown again with a randomly determined new orientation. The total trial duration thus fell between 15 and 17 s.

In the WM task, participants were instructed to remember the orientations of items within the blue border during stimuli presentation phase, and to rotate the probe item until its orientation matched the one they remembered. The differences between their response orientation and the true answer were recorded as the memory precision (Bays & Husain, 2008). The memory load (3, 6 or 9) was manipulated by changing the number of items in blue borders. The initial orientations of the array were sampled from a list of angles between 5 and 350 degree with 15° step (i.e. 5, 20, 35, etc.), and the degree of rotation for probe item during response phase was drawn from a uniform distribution between 0 and 360 degree, in 1 degree steps.

In the SA task, participants firstly responded during the stimuli presentation phase with a right-hand index or middle finger button press to indicate whether a target shape was present or not while ignoring

the coloured borders. They had only 4 s to make a response although the array stayed on the screen for another 2 s afterwards. RT and accuracy were recorded. The central fixation cross would turn red once response time was up. After a short delay of retention phase (2-4 s), they were asked to use the same buttons to rotate the reappearing item back to a canonical orientation of 90° (horizontal bar pointing upward) within 6 s so that the motor response was consistent across both tasks but minimal memory load was imposed to the SA task. We varied the difficulty of this task on three levels of distractor-distractor orientation difference corresponding to 9, 36 and 56 degrees in the search array.

To summarise, both paradigms were closely matched with regards to stimuli presentation and motor response and at the same time the difficulty level was adjusted independently for each task. The experiment consisted of six runs of 30 trials each. Each task was presented in separate runs in an interleaved fashion with difficulty level pseudorandomised within each run, resulting in 30 repetitions per difficulty per task (i.e. SA level 1,2 and 3 and WM load 3,6 and 9) in total. At the start of a new run, the type of task ('selective attention' or 'working memory') was shown on screen for 10 s to notify participants, and afterwards the trials appeared continuously with a 2000 ms interval in between. A full run lasted approximately 600 s. Other aspects of the tasks, such as border colour in the SA task and location of probe items were counterbalanced in order to avoid confounds and keep consistency across tasks. Prior to scanning, participants performed two practice blocks of 20 trials, one for each paradigm.

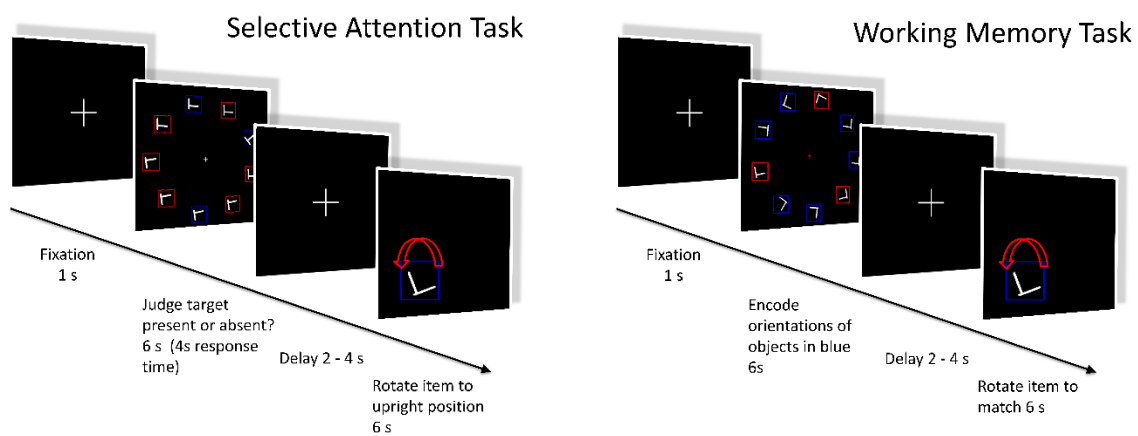


Figure 9. Schematics for SA and WM tasks in fMRI study

3.2.3 Behavioural analysis

To ensure the task difficulty manipulation was effective in the SA task, the same linear mixed effect model described in Chapter 2 was fitted to the unaggregated reaction times from target-present trials with search difficulty as predictor and individuals as random effects. If search difficulty successfully increased the time required to identify the target, parameter estimation for this effect should be significant.

Next, it is important to confirm that the SA/WM profiles based on which participants were selected were stable and still persisted in the in-scanner task, performances were compared between group 1 (low SA/high WM group) and group 2 (high SA/high WM group). Specifically, a 2 by 3 mixed ANOVA with difficulty as within factor and group as between participants factor, was performed on the RTs from target-present trials for the SA task. The same analysis but with memory load as the within factor was conducted for the orientation error for the WM task. Post-hoc tests will be used if the result suggests an interaction. If the group profiles were stable, we expect to observe a group difference on the SA task whereas none for the WM task.

3.2.4 fMRI acquisition, pre-processing and analysis

fMRI data were acquired at the MRC Cognition and Brain Sciences Unit in Cambridge, on the Siemens 3T Prisma system (Siemens Healthcare) using a 32-channel head coil. T1-weighted volume scans were acquired using a whole brain coverage 3D Magnetization Prepared Rapid Acquisition Gradient Echo (MPRAGE) sequence acquired with voxel size $1.0 \times 1.0 \times 1.0 \text{ mm}^3$. Echo time was 3.0 ms and repetition time was 2,250 ms. Functional images were acquired using a standard gradient-echo echo-planar imaging (EPI) pulse sequence (TR = 2000 ms, TE = 30 ms, flip angle = 78° , slice thickness = 3 mm, 25% gap, voxel size $3 \times 3 \times 3 \text{ mm}^3$, 32 axial slices covering the entire brain, matrix size = 64×64 , field of view (FOV) = $192 \times 192 \text{ mm}$). The first 5 volumes served as dummy scans and were discarded to avoid T1 equilibrium effects. Field maps were collected at the end of the experiment (TR = 400 ms, TE = 3.19 ms / 7.65 ms, flip angle = 60° , 64×64 matrices, slice thickness = 3 mm, 25% gap, resolution 3 mm isotropic, 32 axial slices). Stimuli were presented using a monitor viewed through a mirror by participants in the MR scanner.

The pre-processing procedure was implemented using the default stream in fMRIPrep 1.1.8 (Esteban et al., 2019) which is based on NiPype 1.1.3 (Gorgolewski et al., 2011). The T1-weighted image was corrected for intensity non-uniformity and skull-stripped. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and grey-matter (GM) was performed on the brain-extracted T1. Pre-processing for functional runs included skull-stripping, correction for susceptibility distortion using field maps, slice-timing correction and rigid body realignment. This was followed by co-registration to the corresponding anatomical image with nine degrees of freedom. Motion correcting transformations, BOLD-to-T1 transformation and T1-to-template (MNI) warp were concatenated and applied in a single step. The resulting pre-processed T1 and BOLD images were all in standard MNI152 space. Spatial smoothing was applied before single-subject level analysis with FWHM of 6mm.

A general linear model (GLM) was implemented using the *modelgen* algorithm in NiPype using the Statistical Parametric Mapping (SPM) backend. The GLM included task regressors for task types (SA or WM) during stimuli presentation phase (6 s) and its parametric modulator (mean-centred) denoting the trial difficulty level, one for the remainder of the trial length, the six motion parameters, and a

nuisance regressor for trials participants failed to respond to. Only correct trials were specifically modelled and for the WM task this was defined as responses within 90 degrees of the correct orientation (Richter et al., 2016). Duration of the SA task regressor was set as the reaction time of the trial for better capturing the variability in the BOLD response (Grinband et al., 2008), while for WM task it remained fixed to 6 s. Each task was modelled separately; three runs were put into a single design matrix containing the above regressors as well as three constant terms associated with each run. The resulting time course was convolved with a canonical hemodynamic response function to generate the hypothesized blood oxygenation level dependent (BOLD) response. At this step, we also specified a high-pass filter of 128 Hz. Contrasts of interest included each task type compared to the implicit baseline and those related to the parametric modulators, in order to identify regions whose activity varied monotonically with difficulty in SA task or load in WM task.

3.2.5 Group-level univariate analysis and conjunction

Group maps of task vs. baseline contrasts for each task (SA and WM), SA difficulty and WM load were computed using one-sample t-tests on the individual participants' contrast images. NiPype's *OneSampleTTestDesign*, which interfaced with SPM's group test module was used to estimate group level statistics. Group statistical maps were then thresholded voxelwise at $p < 0.001$ and corrected for multiple comparisons by controlling the cluster-level false discovery rate (FDR) at 0.05.

To identify regions commonly activated by both task contrasts and both parametric contrasts, respectively, we conducted conjunction analysis using the group level results. According to the conjunction null hypothesis (Nichols et al., 2005), the valid test should control for the worse-case scenario when all but one individual effect is significant. Hence the minimum statistics across all contrasts of interest should be compared against a single t distribution to calculate significance level. Un-thresholded images were conjoined using minimum statistics (Friston, 1999) before being thresholded at voxel-wise $p < 0.01$ uncorrected for multiple comparisons, and effects were considered to be significant if they passed an FDR corrected cluster-wise level of $p < 0.05$.

3.2.6 Group-constrained subject-specific (GCSS) analyses

In addition to the standard group-mean approach, we also employed the group-constrained subject-specific analysis to identify regions that were activated systematically by both tasks across participants (Fedorenko et al., 2010). A set of brain "parcels" were generated using an algorithm similar to that in Fedorenko et al., (2010) and Julian et al., (2012). The steps for discovering these parcels are depicted in Figure 10. First, each individual's uncorrected statistical maps of the task vs. baseline contrasts (2 for each person) were thresholded at $p < 0.001$ uncorrected and binarized. Next, the binarized maps were intersected to find areas co-activated by both tasks within each individual. Third, all binarized conjunction maps were overlaid on top of one another to create a probability map where the value of

each voxel corresponds to the percentage of participants that have activation at that specific voxel. Fourth, local maximal values were identified and parcels were grown around them using a Watershed segmentation algorithm (Beucher & Meyer, 1993) implemented in the SPM-SS toolbox (Nieto-Castañón & Fedorenko, 2012); https://www.nitrc.org/projects/spm_ss. To create this parcellation, voxels containing local maxima are assigned unique labels which are iteratively propagated to neighbouring voxels until a local minima or zero-valued voxel is reached. The result is a map of labelled voxels, with each label representing a volume (i.e. a parcel) where multiple people exhibited suprathreshold activity, without this conjunction necessarily occurring within the exact same voxels across participants. Lastly, because we wanted to focus on regions where a majority of participants showed intrasubject overlap, parcels containing conjunctive voxels from more than 80% of the participants (Fedorenko et al., 2010) were entered for further examination to ensure consistency of the identified regions of interest across all individuals. The GCSS analysis is advantageous over the traditional one-sample group t-test as it accounts for anatomical variability across participants by allowing for nearby voxels from different participants to be analysed together. This approach also guards against the possibility that a few participants with highly significant voxels drive mean activation in a region that is not representative of the population.

It is still possible that anatomical overlaps do not reflect functional overlaps, because regions could be recruited for distinct neural processes during different tasks. To determine whether the activities in each parcel reflects similar neural computations during SA and WM tasks, beta values for the SA parametric regressor (reflecting difficulty-dependent activity) were correlated with those for the WM parametric regressor (reflecting load-dependent activity) across the top 20% most significant voxels (based on minimum t statistics across contrasts) in each parcel, within individual participants. In other words, the same set of voxels for an individual within a parcel were used to calculate the correlation. The significance of these correlations from all participants was tested using z-transformed coefficients in a one-sample t-test to evaluate the null hypothesis that correlation would be zero if the voxels were engaged in different processes. Finally, we also assessed whether G1 (low SA/high WM) and G2 (high SA/high WM) differed in the extent to which the same set of voxels tracked similar processes within each parcel. A significance level $\alpha = .05$ was chosen and adjusted for multiple comparison using Bonferroni method.

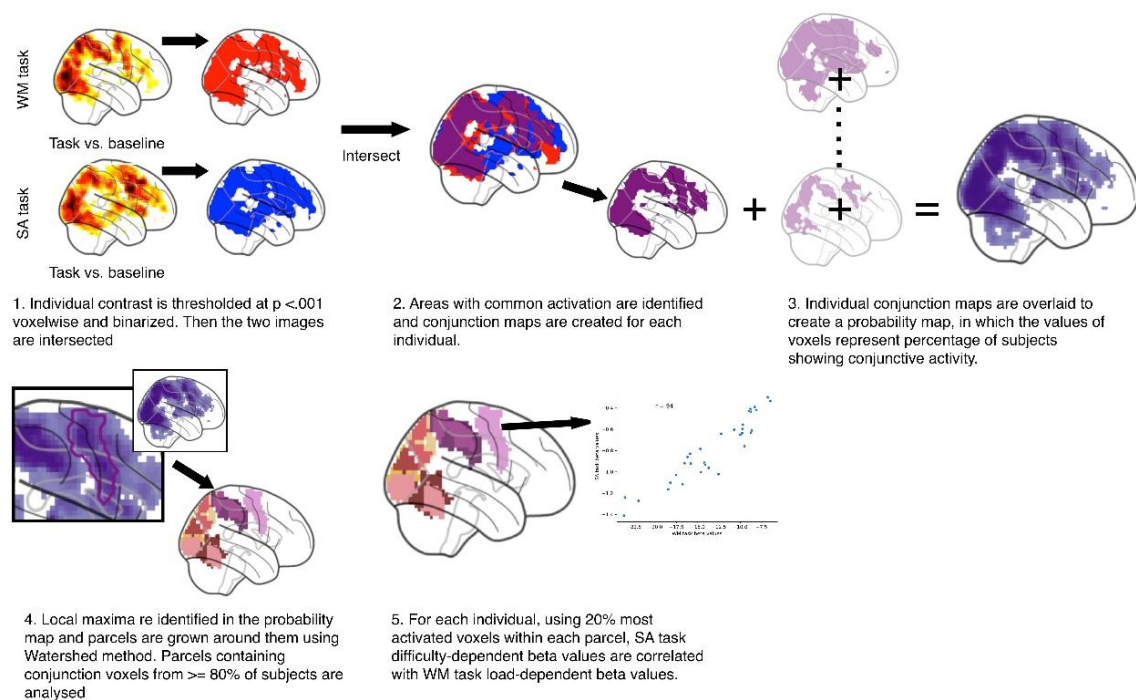


Figure 10. GCSS pipeline

3.3 Results

3.3.1 Behavioural results of in-scanner tasks

For the SA task, all data from one subject from G2 was discarded as they did not make a response in 2 out of 3 runs of the SA task due to misunderstanding the instructions. Mean percent correct for the SA task across participants was 0.84 (SD=0.11) and the average RT was 2491 ms (SD= 205 ms). Participants were in general more accurate ($M_{dif} = 0.09$, $t(40) = 2.44$, $p = .02$) and quicker on trials with a target than without ($M_{dif} = 947$ ms, $t(40) = 22.49$, $p < .001$), in accordance with behavioural results reported in the previous chapter. Difficulty levels (distractor-distractor similarity) predicted RT for target present trials ($\beta = 370.74$, $p = .004$; 95% confidence interval for $\beta = 121.00 - 620.49$) using linear mixed effect model with random effects including individual intercepts and slopes. This effect did not exist for target absent trials ($\beta = 50.85$; $p = 0.70$; parameter estimate confidence interval = -207.33-309.02). We tested whether the two groups selected still differed in the in-scanner SA performance (RT) with a mixed-design ANOVA, with difficulty as the within factor and group as the between-subject factor. There was a significant main effect of group ($F(1,32) = 6.10$, $p = .02$) and significant main effect of trial difficulty ($F(2,64) = 4.30$, $p = .02$), but no interaction ($F(2,64) = 0.86$, $p = .43$). Figure 11 presented the result by group and trial difficulty. Furthermore, to replicate analysis conducted in the

previous chapter, the same mixed linear model was fitted to the log-transformed raw RT data for correct target-present trials from the in-scanner task. Individual slopes of the RT by distractor-distractor similarity function were estimated and an independent t-test confirmed that the group difference was still present for the fMRI task ($M_{G1}=0.141$, $M_{G2} = 0.003$, $t(30.2)= 2.71$, $p = .01$).

For the WM task, participants' mean error (difference between correct orientation and response) across memory loads was 23.45° (SD = 11.90°), 50.14° (SD = 15.03°), 67.26° (SD = 11.90°), suggesting memory precision decreased as number of items to be remembered increased. ANOVA showed that there was no main effect of group across memory loads ($F(1,32) = 0.37$, $p = .55$), or interaction ($F(2,64) = .73$, $p = .48$) but a significant effect of load ($F(2,64) = 198.58$, $p < .001$). Figure 12 showed the averaged error in degree by group and memory load. This result confirms that the two groups were matched on the WM performance.

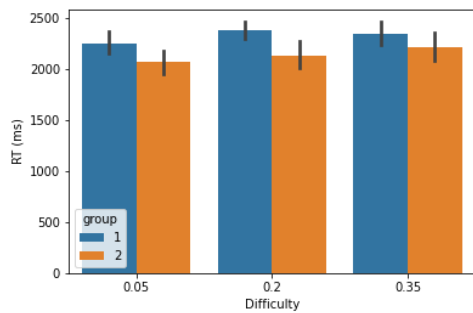


Figure 11. SA task. RT on target present trials by difficulty level

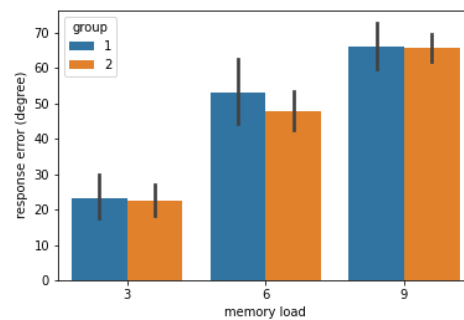


Figure 12. WM task. Orientation error by memory load

3.3.2 Group-level univariate results

Task vs. baseline activity

The whole-brain group analyses tested the contrasts of task vs. baseline for each condition. For SA task vs. baseline, significant clusters of task-related activation in bilateral occipital cortices, superior parietal lobes including the IPS, precentral (PreCG) and postcentral gyri (PostCG), right inferior temporal gyrus (ITG), left middle temporal gyrus (MTG), bilateral middle frontal (MFG), left superior frontal sulcus (SFS) as well as left insula, bilateral thalamus and cerebellum (Figure 13). For WM task vs. baseline contrast, activations patterns were highly similar albeit varying in strength (Figure 14).

Difficulty-dependent activity in SA and load-dependent activity in WM

Load-dependent and difficulty-dependent activity were identified using the contrasts where the corresponding parametric modulators were set to one and the rest of regressors to zero. For the SA task, no cluster relating to search difficulty passed the threshold, suggesting the behavioural effect reflected

in RT was perhaps too small to induce substantial changes in BOLD responses. For consistency, a thresholded image at voxel level ($p < .001$) without cluster-level FDR correction was included (Figure 15) which showed a small activation at left IFS. Significant WM load-dependent activation was largely observed in LOC, precuneus, IPS, medial PFC (BA8/9), anterior cingulate cortex (ACC) and a small cluster in the right MTG (Figure 16). The pattern was mostly bilateral and less widespread relative to the WM vs. baseline contrast.

Conjunction results

In the conjunction of group maps from SA and WM task vs. baseline contrasts we found significant overlap between the two tasks in 13 clusters including the lateral occipital cortex (LOC), SPL, ITG, PostCG and premotor area including FEF, ventrolateral PFC (vIPFC; BA 44) and insula cortex (Figure 17). The biggest cluster expanded from primary visual cortex, SPL and part of the temporal cortex, hence we reported multiple peak coordinates for large regions identified in the analysis (Table 6).

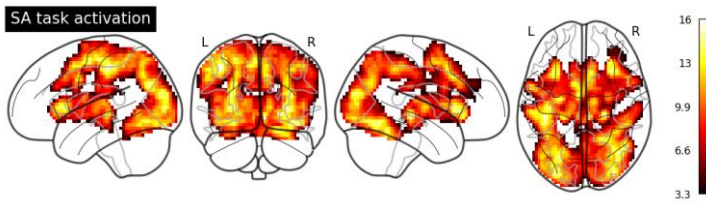


Figure 13. SA task vs. baseline contrast

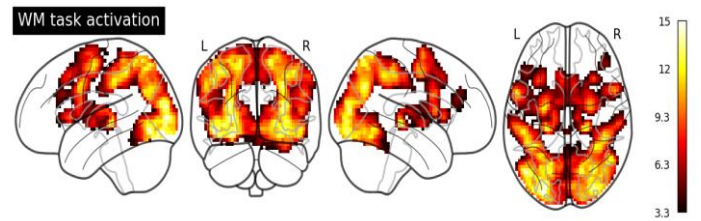


Figure 14. WM task vs. baseline contrast

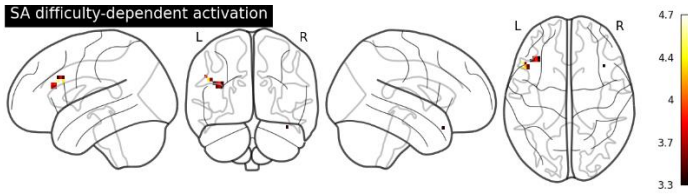


Figure 15. SA difficulty-dependent activity. This image was thresholded at voxelwise $p < .001$. No cluster passed FDR correction.

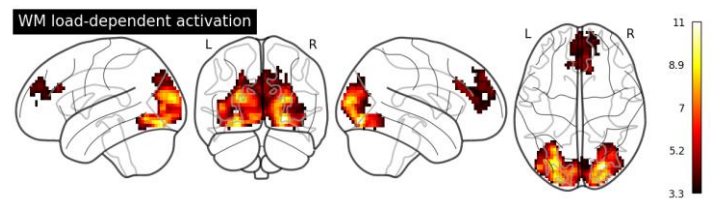


Figure 16. WM load-dependent activity.

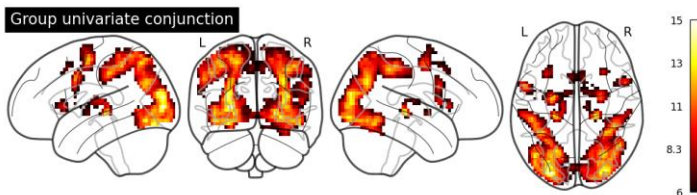


Figure 17. Conjunction of the two task vs. baseline contrasts.

Table 6. Conjunction results from group-level task vs. baseline contrasts

Brain regions (AAL)	BA	Peak MNI coordinates		
		x	y	z
Occipital_Mid_L	18	-27	-90	5
Occipital_Mid_R	19	33	-87	20
Fusiform_L	19	-30	-81	-14
Occipital_Inf_R	18	27	-90	-7
Occipital_Inf_L	19	-45	-72	-10
Parietal_Sup_L	7	-21	-66	53
Parietal_Inf_R	40	42	-39	50
Parietal_Inf_L	7	-36	-45	50
Temporal_Inf_R	37	48	-60	-14
Frontal_Sup_2_L	6	-27	-9	57
Frontal_Inf_Oper_R	44	54	9	27
Precentral_R	6	27	-6	50
Thalamus_L	50	-15	-15	12
Precentral_L	6	-57	6	35
Thalamus_R	50	12	-18	12
Supp_Motor_Area_R	8	9	15	46
Supp_Motor_Area_L	8	-6	15	46
Insula_L	13	-30	18	5
Frontal_Sup_2_R	6	21	0	68

Note. L= left; R=right; Sup=superior; Mid = middle; Inf = inferior; AAL = Automated anatomical labelling

3.3.3 GCSS results

In order to account for the variability between participants in our group-level analysis, GCSS technique was used to identify brain regions systematically and repeatedly co-activated in both contrasts of interest in individual maps. From the parcellation of the conjunctive probability maps generated through combining each subject's data, 20 parcels emerged in which 80% or more of participants showed consistent overlapping activation of SA and WM tasks: bilateral MOG, SPL including the IPS, PostCG, right ITG, left PreCG, and bilateral cerebellum (Figure 18). It was highly similar to the group conjunction map identified using traditional method but in particular differed in the lack of frontal regions.

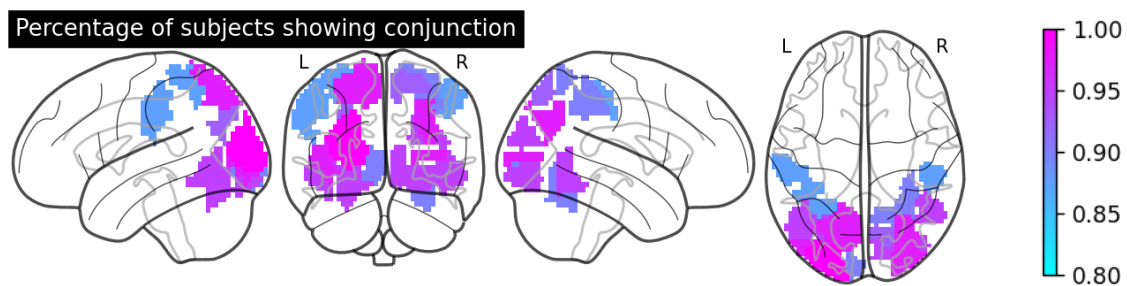


Figure 18. Percentage of participants showing conjunctive activation at each parcel.

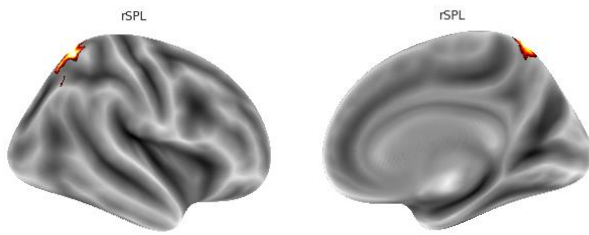


Figure 19. Group difference in correlation strength was found in right superior parietal lobule (SPL)

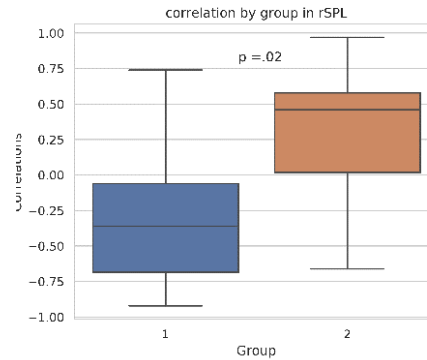


Figure 20. Correlation coefficients between SA difficulty-related and WM load-related activity by group in right SPL

Table 7. Conjunction results from GCSS analysis. (voxels containing overlapping activations from > 80% of total participants were included)

Brain regions (AAL)	BA	Parcel centre MNI coordinates		
		x	y	z
Parietal_Sup_L	7	-18	-69	53.25
Temporal_Inf_L	37/19/18	-48	-69	-10.5
Frontal_Sup_2_L	6	-27	-9	57
Occipital_Mid_R	37/19/18	30	-87	19.5
Occipital_Mid_R	19/39	30	-72	30.75
Parietal_Inf_R	7	33	-45	45.75
SupraMarginal_R	40	48	-33	45.75
Calcarine_L	18	-12	-96	-6.75
Parietal_Inf_L	7	-33	-57	57
Parietal_Sup_R	7	18	-60	60.75
Occipital_Mid_R	17/18	30	-90	4.5
Occipital_Mid_L	39	-30	-75	30.75

The subsequent analysis focused on whether individual subject’s activation patterns within each parcel were similar between SA and WM paradigms, as well as whether the relationship differentiated groups of participants. Given that significance tests were performed for each of the 20 brain parcels independently, a Bonferroni-corrected threshold $\alpha = 0.0025$ was adopted. We calculated the correlations between the beta values in each subject’s contrast maps for SA difficulty and WM load using the top 20% most activated voxels within each parcel of each participant (Figure 19). The Fisher-transformed correlation coefficients for each parcel from all participants were not significantly different from zero after correcting for multiple tests (Table 7). However, there was a significant difference in the correlations between G1 ($r = -0.33$) and G2 ($r = 0.31$), in right SPL including part of precuneus and IPS ($t(32) = 3.36$, $p_{corrected} = 0.02$; Figure 19 & 20). However, separately the correlation of each group was not significantly different from zero after correction. This suggests that for participants whose SA and WM performance were well-matched, rSPL was engaged in similar patterns of computation (e.g.

tracking task demand), whereas for G1 who exhibited dissociated behavioural profiles, the activity patterns of rSPL were reversed.

3.4 Discussion

Building on the behavioural findings from the previous chapter, this second chapter sought to identify the overlapping neural substrates for modified versions of the SA and WM tasks. Participants were screened online, assigned to groups, and then recruited for an fMRI session. The behavioural data demonstrated that the two groups – the highSA+highWM versus the lowSA+highWM – are genuinely reflective of different underlying performance profiles. With each participant tested twice, the group differences remained robust within the scanner. Within the fMRI data, a group-constrained conjunction analysis enabled us to identify brain areas in which neural activity scaled with difficulty to a similar extent across both tasks. Finally, we could use this analytical framework to test whether there are brain areas in which this coupling differs across the two groups, thereby mirroring their differing behavioural profiles.

The fMRI data revealed an extensive network of frontoparietal, premotor and occipital areas recruited for both the SA and WM tasks respectively. Subcortical structures such as the thalamus were also implicated in both tasks. The patterns of neural activity in both conditions showed a highly similar pattern activity in the occipital and parietal regions, although less so in the prefrontal regions. Conjunction results confirmed that several clusters showed spatial overlap across tasks, including the SFS in the lateral PFC, premotor areas, IPS extending to the lateral occipital cortex and inferior temporal gyrus. This is consistent with both respective task literatures (for SA: Hopfinger et al., 2000; Nobre et al., 2002; Tomasi et al., 2007; for WM: Rottschy et al., 2012; Wager & Smith, 2003). One particular difference between our results and previous findings is the strong engagement of primary visual cortex and adjacent areas. This is not typical in visual search or delayed-response tasks and is likely because the current study did not use a control condition that accounts for the basic sensory processing; hence the related activity was not subtracted from the overall signal. (NB: this was intentional because our primary aim was to contrast the two tasks, for which the sensory requirements are fully equated).

For the WM condition, neural activity increased monotonically with memory load across multiple regions. These areas were broadly similar to those identified by contrasting the WM task to baseline activity, but more constrained to the posterior portion of the brain, namely the IPS and the LOC. The effect is likely induced by the escalating WM demands, as other aspects of the task were held constant. Of these regions, IPS and LOC have been consistently identified in studies with a low vs. high WM load contrast or a parametric load regressor (Rottschy et al., 2012). This effect could reflect multiple heterogeneous processes including, but not restricted to, increased WM storage or qualitatively different mechanisms required to orchestrate the encoding and maintenance of more information compared to

when memory demands are low. It is unlikely that participants would use perceptual grouping consistently throughout the trials, as orientations of items were drawn from a uniform distribution so that no meaningful pattern should form in the visual display to facilitate chunking (e.g. when nearby items were similar, Peterson & Berryhill, 2013). Xu and Chun (2006) varied number of visual objects and presentation sequence and observed that both regions were sensitive to memory load, but LOC also tracks the complexity of objects, reaching similar level of activation for four simple objects and two complex ones. This is in keeping with findings that visuospatial WM is mediated by the same regions processing perceptual information – the so-called sensory recruitment account of WM (see also Harrison & Tong, 2009; Jonides et al., 2005)). In contrast, Xu and Chun (2006) found that the IPS was sensitive to the *number* of spatial locations; activity did not increase if different objects were presented sequentially in the same location. They argued that the function of IPS was more associated to the allocation of spatial attention rather than object representation, consistent with findings that IPS activity supports goal-directed attention (Culham et al., 2001; Hutchinson et al., 2009), whereas the LOC plays a crucial role in object representation.

The GCSS analysis demonstrated functional differences between groups in a parcel located at right SPL expanding across IPS and precuneus. That is, participants with discordant SA and WM performances (G1) showed decoupled voxel activities between the two tasks whereas those with equivalent task performance (G2) showed positively correlated neural responses in the region. Several experiments investigating the common capacity limits of visuospatial WM and SA using concurrent tasks have identified this area (Fusser et al., 2011; Mayer et al., 2007; Silk et al., 2010). For instance, Fusser and colleagues combined visual search and delayed response into a single paradigm, by using the target stimuli in the former to guide memory encoding and manipulated the attention and memory loads independently. They reported an interaction in BOLD response to increases in attentional and encoding demands in the SPL and FEF, in contrast to frontal regions where an additive effect was observed (i.e. the activation difference between high vs low memory load was equal across different attentional loads). This was taken as support for these areas underlying a shared resource between the two functions which limits processing capacity. Besides this direct demonstration of the interplay between SA and WM within this area, others have theorised that the common function of the SPL is goal-directed allocation of attention, which is recruited for both SA and WM (Awh & Jonides, 2001; Pessoa et al., 2003). Specifically, the spatiotopic patterning of neural activity within the IPS could reflect a spatial map, in which coordinates guide visual selection as well as encoding and retrieval of spatial locations. Alternatively, it has been suggested that SPL subserves attentional modulation of information integration into object representation stored in distributed sensory cortex (Postle et al., 2004; Wheeler & Treisman, 2002), which is consistent with findings that memory contents can be decoded in the same structure (Bettencourt & Xu, 2016).

In addition, one previous study by Nee and Jonides echoed our finding that individual differences exist in the neural overlap between perceptual and mnemonic selections (2009). In one of their analyses, participants were split into two groups based on activation intensity in rSPL during memory selection, a region recruited by both types of tasks but more strongly by perceptual selection. The group with higher memory-related activity in this area also recruited the FEF and part of the premotor cortex to a greater extent, and performed better on the memory condition compared to their counterparts. In contrast, the other group showed stronger activation in the left vlPFC, a region more uniquely associated to memory selection in the study. In other words, individuals who relied on more SA-associated regions for memory selection (i.e. more neural overlap) had a behavioural advantage against those who did not. Our result provides converging evidence that, aside from common activations, functional coupling (the extent that same set of voxels track both attentional and memory demands) in the rSPL also differentiate between individuals with separate behavioural profiles

Over and above the various theoretical accounts for the role of this cortical area, the unique contribution of the current study is its differential association across those with strongly coupled SA and WM performance, and those who seem to be good at one task, but not the other. This provides some neural instantiation of the behavioural pattern identified in previous chapter, and supports the hypothesised notion that there is individual difference in the functional coupling between SA and WM. In short, some participants are relying upon more overlapping neural architectures to perform these two tasks, relative to other participants. Although, it is fair to say that this effect is not widespread. Moreover, the results showed that although overlaps revealed by univariate conjunction analysis were largely present across the cortex, differential levels of inter-subject variability were associated with the regions. Visual processing area was marked by highly robust activation by both tasks across individuals, whereas the consistencies of conjunctive activity seemed to reduce moving from posterior to anterior part of the brain (no frontal regions passed our selected threshold in the GCSS analysis), indicating a varying degree of individual differences in the functional overlaps underlaid by these regions. This is generally in line with various models proposing functional segregations in subregions of lateral PFC, for example, the rostral vs. caudal PFC distinctions reflecting task set vs. memory load effects (Rottschy et al., 2012); Flechter and Henson (2001) attributed manipulation/monitoring to dlPFC and updating/maintenance to vlPFC, etc. That is to say, PFC are broadly related to a number of control processes compared to the relatively uniform function (e.g. control of attention, mental representation) supported in the PPC (Berryhill et al., 2011) and visual cortex. This provides a plausible explanation for the varying degrees of inter-subject consistency in the functional overlap we observed here: while participants need to rely on the same low-level sensory processing mechanism, they could draw on different high-level processes to perform a complex task.

There are a number of limitations that we encountered, which provide areas for improvement in future. Firstly, the design would benefit from a more effective manipulation of SA difficulty, which would

yield a more robust above-threshold activation in the group-averaged analyses. However, it is worth noting that we did still identify a robust behavioural result, consistent with the preceding chapter, showing that the between-distractor similarity manipulation is effective. Moreover, in the GCSS analysis, where inter-subject variability was taken into account, significant correlations between the difficulty-related beta values from both tasks were observed within a subset of individuals. It would be very surprising if the beta coefficients for SA difficulty and the correlation with WM load were a result of random noise. Considering that all other aspects of the SA trials were held constant among the different levels, the parameter estimation was very specific to the changing search difficulty and should not have been confounded by any other factors. A possible explanation is that the neural effect was present but only at a sub-threshold level, which was more sensitive to analysis at an individual level. Secondly, and relatedly, due to limitations of time imposed by the pandemic, the current sample size was small considering it was partly a between-subject study. Recruiting more participants may afford better power for establishing SA-related effects at a group-level, and to detect more widespread underlying differences between the behaviourally defined groups. Thirdly, the SA and WM task may pose different oculomotor demands during the stimulus phase, as eye movements in the SA task potentially consist of a few saccades and fixations guided by internal target template, those in the WM task should in principle depend on number of encoded items. A potential outcome is that brain activity associated with eye movement in one task is stronger in one task compared to the other. However, this should not be an issue to the conjunction result as the analysis identified regions activated in both conditions above a threshold, so that if there are indeed regions specific to saccadic control, they would not be indicated in result unless activations are significant in both tasks. Furthermore, there has been strong evidence that attentional and oculomotor processes recruit highly indistinguishable brain networks (Corbetta et al., 1998), the former of which is fundamental to both tasks. It would be very difficult to disentangle one from the other. The experiment however, could benefit from including an explicit control condition with matched perceptual input and motor response, the activity of which can be subtracted from that in the SA and WM tasks in order to isolate regions specific to attentional and encoding processes.

Chapter 4: Task representations in the brain and artificial neural networks

4.1 Introduction

There is a behavioural and cognitive overlap between working memory (WM) and selective attention (SA) that can be seen in patterns of individual differences (Astle et al., 2014; Engle, et al., 1999; Kane et al., 2001; Vogel & Machizawa, 2004) and dual task performance (Anderson et al., 2008; Burnham et al., 2014; Lavie & De Fockert, 2005). This overlap is thought to reflect a common underlying capacity-limiting process (Cowan, 1998, 2001; Gazzaley & Nobre, 2012; Wheeler & Treisman, 2002). In the preceding two chapters, we developed two new tasks that are carefully matched in multiple respects, but in which the difficulty of selection and maintenance could be independently manipulated. These two processes can be differentially coupled behaviourally across participants and this could reflect potential different underlying latent constructs. A neuroimaging study then demonstrated that a broad network of areas is not only co-activated across the two tasks, their pattern of activity scales with both within-task manipulations. Finally, one particular region showed individual differences, scaling differentially depending upon the participants' behavioural profile. However, despite the spatial overlap between SA and WM in neural correlates, we do not yet know if or what they share in terms of computational processes. That is, whilst an increasing number of researchers are asking whether the behavioural, cognitive and neural correlates of SA and WM overlap, this is not the same as testing whether they share computational processes and what the functional nature of those processes, or indeed that overlap, is. That is the purpose of this next chapter.

Using artificial neural network to study complex cognitive functions

In WM-related research, neural network models have been constructed to test and quantify potential theoretical frameworks associated with specific paradigms (Botvinick & Plaut, 2006; Farrell & Lewandowsky, 2002; Oberauer et al., 2012). One successful example is the study of serial recall. Botvinick & Plaut (2006) trained a recurrent network (a type of artificial neural network whose current activation is determined by both the current input and its output at previous time steps, therefore generating temporal dynamics) to recall the order in which sequence information is encoded through sustained patterns of activation within the network architecture. The model performance replicated well-known characteristics of human data such as the primacy effect and swap errors, and implicated possible neural mechanisms that could simultaneously represent all elements, as well as a mechanism that represents the currently activated for output. Importantly, such hypotheses generated by the model are testable with future behavioural and neuroscience experiments.

A parallel line of research sits at the intersection of computational models and cognitive neuroscience, which heavily emphasises neurobiological accuracy in the modelling (for a review, see Ashby & Helie, 2011). The approach often involves the specific modelling of certain brain regions including neuron

firing patterns, the synaptic strength changes through learning that underlies long-term potentiation/depression, and anatomical connections between areas with detailed directions (i.e. excitatory or inhibitory influence). Such biologically plausible models have achieved moderate success in accounting for behaviours and neurocognitive phenomena in WM maintenance, operationalised by a series of complex interconnections between PFC, posterior parietal cortex (PPC) and basal ganglia structures (Ashby et al., 2005; Frank et al., 2001).

A benefit of an artificial neural network model is that the researcher has complete access to the activity and structure of the circuit, and the ability to probe them. This can provide a convenient substitute for biological circuits and a useful testing ground for theoretical inquiry. Nevertheless, the application of neural network models to understand *multiple* cognitive functions, and the nature of their interrelationship, has only recently become a focus of the field (Yang et al., 2019), and few if any have directly studied WM- and SA-related processes. In their ground-breaking work, Yang and colleagues trained a recurrent neural network (RNN) to perform 20 cognitive tasks simultaneously, thereby mimicking the flexibility of biological brains. The tasks ranged from categorization, WM, inhibition control and decision making, including multiple variants that are mixtures of these elements, such as a delayed decision-making task. Having successfully learned to perform all 20 tasks, the model exhibited functional clustering of computational units (i.e. neurons) that specialised in common processes shared by subsets of the tasks. For example, one cluster is selective in both perceptual decision-making and delayed decision-making tasks, presumably representing the decision-making component common to both, whereas a second cluster is selective to both delayed decision-making and delayed matching-to-sample tasks, presumably underpinning the WM component of the tasks. This computational model allowed the authors to test for multiple different configurations of task relationships. For example, whether the representational units needed to perform a certain task are a nested portion of those needed for another task, or whether two tasks share certain key processing components in common alongside their own unique components. In this chapter we intend to use the same principle to understand the relationship between task representations for our novel SA and WM tasks.

RNNs: a biologically-constrained computational framework

RNNs are networks in which neurons can send and receive feedback to and from each other. As a result, the activity of neurons in the network is affected not only by the current external stimulus, but also by the current state of the network (Barak, 2017). This property makes RNNs ideally suited for computations that unfold over time such as holding information in WM or accumulating evidence for decision-making, such as in the SA task. Using RNNs to model the brain is inspired by anatomical and electrophysiological findings that the majority of cortex is recurrently connected to itself (Douglas & Martin, 1991). Moreover, the dynamic patterns generated by these networks are reminiscent of ongoing activity observed in the real brain (Van Vreeswijk & Sompolinsky, 1996). For these reasons, RNNs

have been applied to model the dynamics of neuronal population in cortex during perceptual, cognitive, and motor tasks (Carnevale et al., 2015; Mante et al., 2013; Masse et al., 2019).

What architecture would an artificial neural network develop in order to perform complex tasks that consist of attentional selection and delayed responding (i.e. processes needed for the SA and WM tasks, respectively)? One possibility is that it may resolve multiple tasks by representing each one with non-overlapping populations of network units. Alternatively, a network could develop subcomponents that are optimized towards several objectives by dedicating a subset of neurons to represent the common computational processes across tasks. This latter possibility is intuitively a feature of multiple cortical areas thought critical for flexible cognitive control (Cole et al., 2013; Woolgar et al., 2015). And if this is true, for a model to perform both SA and WM tasks flexibly, how big must the overlap be between the tasks represented by the network be? And what is the functional nature of their overlapping components? The RNN model may help address these questions. Furthermore, once trained, a RNN model could be related to brain activity by means of computing the representational similarities of experimental conditions (Kriegeskorte et al., 2008), in order to evaluate the extent to which the computational model resembles the way a brain representing the relevant processes. In short, the present work drew inspiration from recent advances in computational model training (Yang et al., 2019), and set out to test: 1) whether an RNN with biologically plausible architecture (Song et al., 2016) can successfully learn WM and SA tasks simultaneously; 2) whether the model trained in this way exhibits comparable properties that resemble actual behavioural and neural data; and 3) what can be learned about the functional nature of the overlap between SA and WM from the RNN model.

We simultaneously trained a single RNN to perform the SA and WM task used throughout the experiments in this thesis. Systematic examination of the resulting network then enabled us to probe the emerging structure, identify clusters of neurons on the basis of their activity patterns, and better understand the computational roles of these clusters. We also attempted to draw a link between our model and brain measurements acquired from real participants performing the same tasks in the scanner. In essence, we wanted to test whether there are similar principles governing how these two types of network – artificial and real – resolve these two tasks. Given the theoretical backdrop, in which multiple researchers have posited shared processes across SA and WM, such as the top-down modulation of attention (e.g. Gazzaley & Nobre, 2012), we expected to find a subset of network units dedicated to functions shared across tasks, whilst the other units potentially serve more task-specific functions. The nature of these processes would then be tested further by comparing the representational similarity matrix of network clusters to that of brain regions, incorporating previous knowledge on their functional roles of these regions in SA and WM.

4.2 Methods

4.2.1 Recurrent neural network model

Recurrent neural networks are a large class of neural network architectures that enable processing of a sequential input over time. Generally speaking, a neural network model learns to predict the task by minimizing the distance between its output and the correct answer provided. We trained a RNN model using the Python package PsychRNN (Ehrlich et al., 2021; Figure 21), a high-level wrapper of the TensorFlow backend that allows flexible application of RNNs and coding of cognitive tasks. We report result of a model containing 128 fully connected recurrent units in the recurrent layer which can either be excitatory or inhibitory (i.e. the unit's connection weight was set either positive or negative), but similar results were also obtained with a model containing 256 recurrent units. This biological constraint allows the network to mimic the neurophysiological set-up of the brain, which specifies that neurons have exclusively excitatory or inhibitory synapses (Rajan & Abbott, 2006; Song et al., 2016). There were additional input and output layers whose number of units matched the corresponding channels required for the network input and output (see below). The model is governed by the following dynamic equation:

$$\tau \frac{dr}{dt} = -r + f(W_{rec}r + W_{in}u + b_{rec} + \sigma_{rec}\sqrt{2\tau}\zeta)$$

where $\tau = 100$ ms is the neuronal time constant. r is the recurrent layer's activity at any given time. W_{in} , W_{rec} , and W_{out} denote the connection weights of the input layer, the recurrent layer, or the output layer. U is the input to the network, b is the bias term associated with the layer, $f()$ is the Softplus activation function which allows nonlinear network activity. ξ are independent Gaussian white noise with zero mean and unit variance and $\sigma_{rec} = 0.05$ is a constant to scale recurrent unit noise. Finally, the output of the model is specified by

$$z = W_{out}r + b_{out}$$

Regarding the regularisers (i.e. penalties added to prevent over-fitting to the training data), weight initialiser and the loss function (i.e. the measure for difference between the model and desired outputs that the training procedure tries to minimize), we did not change the default setting of the package, therefore the detailed implementation can be found in (Ehrlich et al., 2021)). As the present work is largely exploratory, we maintained the parameters settings used by (Yang et al., 2019) in their main text throughout the training procedures, in hope of closely imitating the network structure which has been proven successful in learning to perform multiple cognitive tasks.

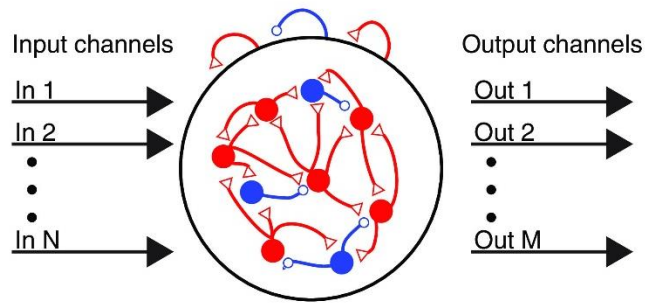


Figure 21. RNN structure from (Ehrlich et al., 2021). The model contains fully connected recurrent neurons and can take and send any number of inputs and outputs. Colours represent excitatory or inhibitory neurons.

4.2.2 Coding the behavioural tasks

Similar to Yang and colleagues' work (2019), we abstracted the relevant stimulus properties from the paradigms used in the fMRI chapter, rather than presenting the exact same visual inputs to the RNN (this would have required an entire RNN of its own to deconstruct the visual input, prior to it being entered into the cognitive RNN). Descriptions of the SA and WM tasks can be found in the methods section of the preceding chapter. In short, the goal in the SA task is to judge whether a target is present among an array of distractors, in which the distracter-distracter similarity can be varied parametrically to adjust the difficulty level. The purpose of the WM task is to retain relevant information about the target stimuli and make a response after a short delay, with the memory load being manipulated. Both paradigms adopted the same set of stimuli which allows the same stimulus features to be abstracted.

Specifically, the input consisted of 8 channels with the first 2 for task rule that informed the network which task was currently activated, and the other channels represented stimuli features. The rule inputs were either (1,0) for the SA task or (0,1) for the WM task, and remained unchanged throughout the whole trial. Each stimulus item was simplified into 2 pertinent properties: its orientation and the colour of its frame border. We chose to simplify the trials by only including 3 items instead of 9 in the original fMRI task, leading to $2 \times 3 = 6$ stimulus channels, thus affording more flexibility in the exploratory stage. Moreover, it is not very likely that the number of items would result in qualitatively different network structure, given that unlike humans and animals, RNNs are not constrained by intrinsic processing limits, at least not in the same way. In the study which inspired the current work, Yang et al (2019) included two stimuli for a set of 20 tasks and identified behavioural hallmarks typically found in real participants. Therefore, we believe reducing the number of items should not hinder our goal of studying the *overlap* between cognitive processes.

Within any given trial, the first 20 timesteps were set as the fixation period during which only the rule input was active. For the SA paradigm (Figure 22), the orientation of the target was always set to be 0, and the distractors' orientations were randomly selected from (-0.3, -0.2, -0.1, 0.1, 0.2, 0.3). The values did not correspond to any real angles as we do not assume the network perceives the relative difference of the stimuli the same way as human participants. What is important is that we added noise to the input signals which was drawn from a Gaussian distribution of zero mean and 0.05 standard deviation to mimic the intrinsic noisy nature of visual processing, resulting in the stimuli being more difficult to discern when distractor orientation became more similar to the target orientation (0) in the "eyes" of the RNN. It is important to note that this manipulation effectively varied both the distractor-distractor and target-distractor differences, unlike the SA task reported before which only varied similarity between distractors, as we were concerned with quickly implementing the model. Although its performance indeed showed decrease the smaller the distractors' values, replicating behavioural finding from human participants, this remains a point for improvement in the future. One of the distractors' orientation was also multiplied by -1 to replicate the behavioural trials where there were always two types of distractors in terms of orientation. Half of the trials had a target while the other half have only distractors to prevent the network from learning alternative strategies such as sticking to one type of response, which in this case would only lead to a maximum of 50% accuracy. As for the frame colour feature, same as in the fMRI paradigm it did not provide any useful information for completing the SA task, therefore were randomly assigned to either 0 or 1. The stimuli presentation period lasted for 140 timesteps and the response period for 40 timesteps. During the response phase, all inputs but the rules were turned off and a readout of the network's output channel was carried out to obtain its predictions. The correct output is 1 for target present trials and 0 for target absent trials. The network's learning process was based on adjusting its connection weights according to the discrepancy between its predictions and the correct answers (Kingma & Ba, 2014)

For the WM task (Figure 23), the values of orientation channels were randomly chosen from the same parameter space as the SA task. Here the memory load was varied by adjusting the number of memory targets during the encoding phase between 1, 2 and 3. Value 1 in the frame colour channel represented a target and 0 a nontarget. Only memory targets were probed later, making the trials easiest at load 1 but most challenging at load 3. After a short retention period when all stimulus channels were set to zero, the probed target channels were turned on again but on half of the trials with an altered orientation value (always ± 0.4). The task goal was to determine whether the probe's orientation was changed compared to the remembered one. The correct output was 1 for change trials and 0 for no change trials. Notably, this particular setting is a slight deviation from the WM task employed in the fMRI study, as instead of rotating the probe to match the remembered orientation, the network only needs to make a binary judgment, resulting in an easier and more suitable starting point for our modelling effort.

Additionally, the recurrent layer projects to an output channel where response range is limited between 0 and 1. In the SA task, the option for target present is coded 1 and target absent 0. In the memory task, change is coded 1 and no change 0. The prediction is considered correct if the averaged result over the response period is closer to the correct choice than the incorrect one.

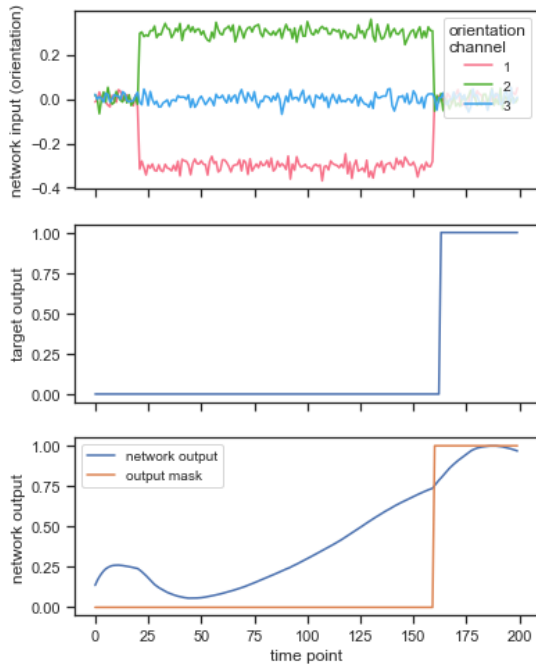


Figure 22. Network input and output on the SA task.

The first subplot shows input into the three orientation channels during a trial. This trial had a target (blue) and 2 distractors. Frame colour channels were not shown as their inputs were randomly assigned and resembled the WM task. Second subplot showed the correct output for this trial, which only requires outputs during the response phase. The last plot illustrated the actual network output (blue) and a temporal mask applied to it in order to mask out outputs during the non-response period.

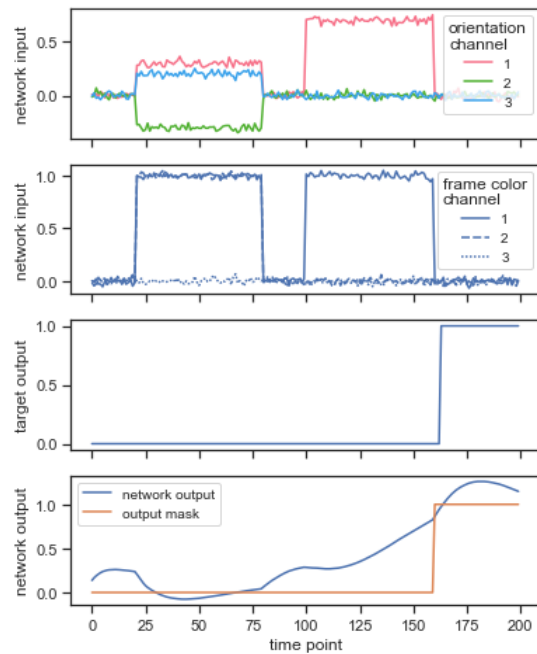


Figure 23. Network input and output on WM task.

This particular trial had a memory load of 2, as shown by two frame colour channels with values of 1. Stimulus 1 was re-presented during the probe phase (100-160 timestep) and its orientation was changed compared to the encoding phase (20-80 timestep). Therefore the correct output for this trial should be one (3rd subplot). The output mask is shown overlaid on the plot of network output to illustrate that only network output during the response phase (160-200 timestep) was considered for model training.

4.2.3 Dissecting the RNN model

Firstly, we tested whether individual units in the network were selective to different tasks, or whether units tended to be similarly recruited for both. To quantify how selective a recurrent neuron is to one task, Yang et al. (2019) defined a measure for task variance (TV) by feeding the trained model a new batch of trials from a given task that utilizes the entire task parameter space. Taking the SA task as an

example, the new trials covered all combinations of stimuli orientation, frame colour and target presence/absence. For a given neuron of the network, the activity variance across trials at each time step was first calculated then averaged across all time steps to obtain the final TV for this neuron. Next, TVs were normalized by the peak value across tasks for each neuron. In essence, if one neuron is functionally involved in a task, its activity should differentiate between the task parameters necessary for performance. Conversely, if a neuron or set of neurons was not selectively recruited for a task then its activity pattern should be relatively invariant across this set of inputs.

Next, to quantify the selectivity of a RNN unit i to any given task A relative to task B, we used a metric *fractional task variance (FTV)*, which is specified as

$$FTV_i(A, B) = \frac{TV_i(A) - TV_i(B)}{TV_i(A) + TV_i(B)}$$

FTV ranges between -1 and $+1$. The more positive the *FTV* the more selective a given neuron is towards task A compared to task B.

We also conducted clustering of the recurrent neurons based on their task variance using k -mean clustering from the Python package Scikit-learn to investigate whether the model has reorganized into several functional modules. Since the choice of the hyperparameter K is predefined and subjective, this could potentially cause varying results. To assess how well a clustering solution fits, we computed silhouette scores for all possible K s (Rennie et al., 2019). The silhouette score measures the ratio between the average distance of a unit with other units in the same cluster, and the average distance between this unit and units in the nearest cluster. The silhouette score of a clustering solution is the average silhouette score of all units. A higher silhouette score means a better configuration with distinct clusters. We computed the silhouette for the number of clusters ranging from 2 to 10. The optimal number of clusters is determined by choosing the k with the highest silhouette score.

To understand the functional roles of the clusters, neurons were lesioned, one cluster at a time. This was achieved by setting neurons' connection weights to all recurrent and output units to zero. We compared the network's accuracy after inactivation of a cluster on a new batch of SA and WM trials, relative to that after inactivation of the same number of randomly chosen neurons. This comparison is designed to control for a performance drop purely due to losing a set of neurons. All reported values were based on the results of 640 trials that equally represented the entire task parameter space. The random lesioning step was permuted 500 times in order to achieve a representative accuracy. Additionally, it may be possible that no difference between clusters would be observed if each contains a significant number of units, the deactivation of which would lead to catastrophic consequence in performance. To solve this issue, we also applied lesioning in an incremental manner. Specifically, within each cluster, gradually more units were disabled over 5 steps with equal increments with each

step repeated for 100 times. In this way not only the drop in accuracy but its slope will provide information regarding the sensitivity of a task to the loss of a particular group of units.

4.2.4 Representational similarity analysis

Here we addressed the question whether the clusters of neurons differentially represent the tasks and stimulus conditions. To do so, we used representational similarity analysis (RSA; Kriegeskorte et al., 2008) which is a framework to abstract activity patterns into representational dissimilarity matrices (RDMs) and directly compare information carried by various modalities such as a brain or model. Since it is necessary for RSA to have multiple experimental conditions, we extracted the mean activity from each recurrent unit across all time points but the fixation and response period for each difficulty level separately. This can be seen as an averaged representation of the conditions by the model, producing 6 conditions in total as there were 3 levels in each task.

Having extracted the information, the 1 minus Pearson's correlation coefficients of all pairs of conditions were computed before they were assembled in to a 6 by 6 RDM that is symmetric about a diagonal of zeros. This procedure was repeated for all clusters. RDMs can be further examined to reveal the relationships between clusters of neurons by calculating their relatedness. This is achieved by computing the Spearman's correlation between the two sets of dissimilarity values in the upper (or equivalently lower) triangular region and the significance of the relatedness can be tested by creating a null distribution using 1,000 permutation technique on the condition labels.

4.2.5 fMRI data pre-processing and comparisons to RNN using RSA

The use of RSA allows us to relate the task representations of multiple modalities and to understand to what extent the neural mechanisms implemented by computational models and brains overlap. Here we re-analysed the data acquired for the fMRI experiment that has been described in Chapter 3, where more experimental details can be found. Essentially, we scanned participants while they performed the SA and WM tasks, each for three runs. Additionally, there were some notable deviations in the task implementation between fMRI experiment and the modelled tasks, especially those related to the size of the stimuli array and the way of manipulating task difficulty. For instance, in the fMRI WM task, memory load was varied between 3,6 and 9 items, however we only modelled 1,2 and 3 items for RNN. Overall, however, we maintained as much similarity between the modalities as possible.

We repeated the same pre-processing pipeline for the neural data, but fitted a new first-level general linear model suitable for extracting blood oxygenation level dependent (BOLD) signal associated with the experimental conditions (Kriegeskorte et al., 2008). The model included three task regressors, each correspond to one difficulty level of a task with duration of 6 second from onset (encoding phase), as well as a regressor for the rest of the trial. This was to isolate the hemodynamic response for stimulus

presentation/encoding phase from the later phase of the trials, separately for each difficulty level to extract the unique representational patterns. The major difference from the fMRI analysis was that in that study, only one encoding regressor with a parametric modulator corresponding to the difficulty levels/memory loads was used. There were also six motion parameters and a linear trend predictor per run as nuisance regressors. The resulting time course were convoluted with a canonical hemodynamic response function to generate the hypothesized BOLD response. All runs of a task were fitted with the design matrix separately before averaging. For contrasts of interest, we simply set the weight of the specific regressor (e.g. SA task level 1) to one and the rest to zero to extract the beta estimates for that condition. For example, to attain beta values associated with the easiest level in the WM task, we set the weight of the load 1 regressor in the design matrix to one and the others to zero.

Once the individual beta images have been estimated, we parcellated the cerebral cortex according to the 100-region brain atlas created by (Schaefer et al., 2018) before computing the RDMs within each parcel for each subject. Similar to the neural network RDMs, this was achieved by extracting the activity patterns (i.e. voxel beta estimates) associated to each experimental condition and calculating the dissimilarity values between pairs of conditions, resulting in 100 regional matrices across the brain. Then, every brain RDM was compared to every RNN cluster RDM to derive a correlational value using Spearman's r . To compute the significance level associated with the correlations, condition labels of the cluster RDM were randomly shuffled and a correlation calculated for 100 times creating a distribution for the null hypothesis that the two RDMs were unrelated. The resulting single-subject brain correlation maps could be transformed using Fisher's z transformation and entered in a paired group comparison to test for any regions that were differentially associated to one cluster relative to another (Nili et al., 2014). Since there are 100 brain areas in the atlas, FDR correction was applied to adjust for the significant level ($\alpha = .05$). This method has certain implications in generalizability, as pointed out by Kriegeskorte et al. (2008). This is essentially a random-effect analysis in contrast to a fixed-effect approach of computing a brain RDM based on averaged group-level result, which affords greater statistical sensitivity. However, our approach allows for inference at the population level while the alternative depends on the assumption that the brain function of interest has a neuronal mechanism consistent across the population, which is reasonable for basic functions like visual processing but uncertain for those under study in the present analysis.

4.3 Results

4.3.1 Training RNNs on the attention and short-term memory tasks

To study how the same SA and WM paradigms could be learned by an artificial neural circuit, we trained an RNN model to predict correct answers for trials with binary choices. Tasks were coded in a way that stimuli attributes and task rules were abstracted into values, and responses options were

represented by ones and zeros. Before training, a network is incapable of performing any task. It is trained with supervised learning which modifies all connection weights (i.e. input, recurrent and output) to minimize the difference between the network output and the correct output. Here we report results from a RNN consisting of 128 recurrent neurons trained on 50,000 iterations, however it did not vary significantly with the size of the model. After training, the network achieved very high behavioural accuracy across both tasks with a new batch of trials: 93% for WM task and 100% for SA task. Since the model already performed extremely well with relatively few iterations (30,000), we also plotted the psychometric functions using a midway model (i.e. before the training was complete) to demonstrate that it displays behavioural features similar to the result acquired from human participants (Figure 24). For example, in the SA task, the RNN achieved the best performance at the easiest level of difficulty (level 1), which decreased as the trials became harder – i.e. as the distracters became closer to target's orientation value due to noises added to the inputs. Likewise, accuracy was highest when WM load was 1 and became lower at load 2 and 3. Due to the fashion in which these tasks are coded, it is difficult to directly compare reaction time data from the model to that from actual human. The subsequent results were from analysis of the complete model.

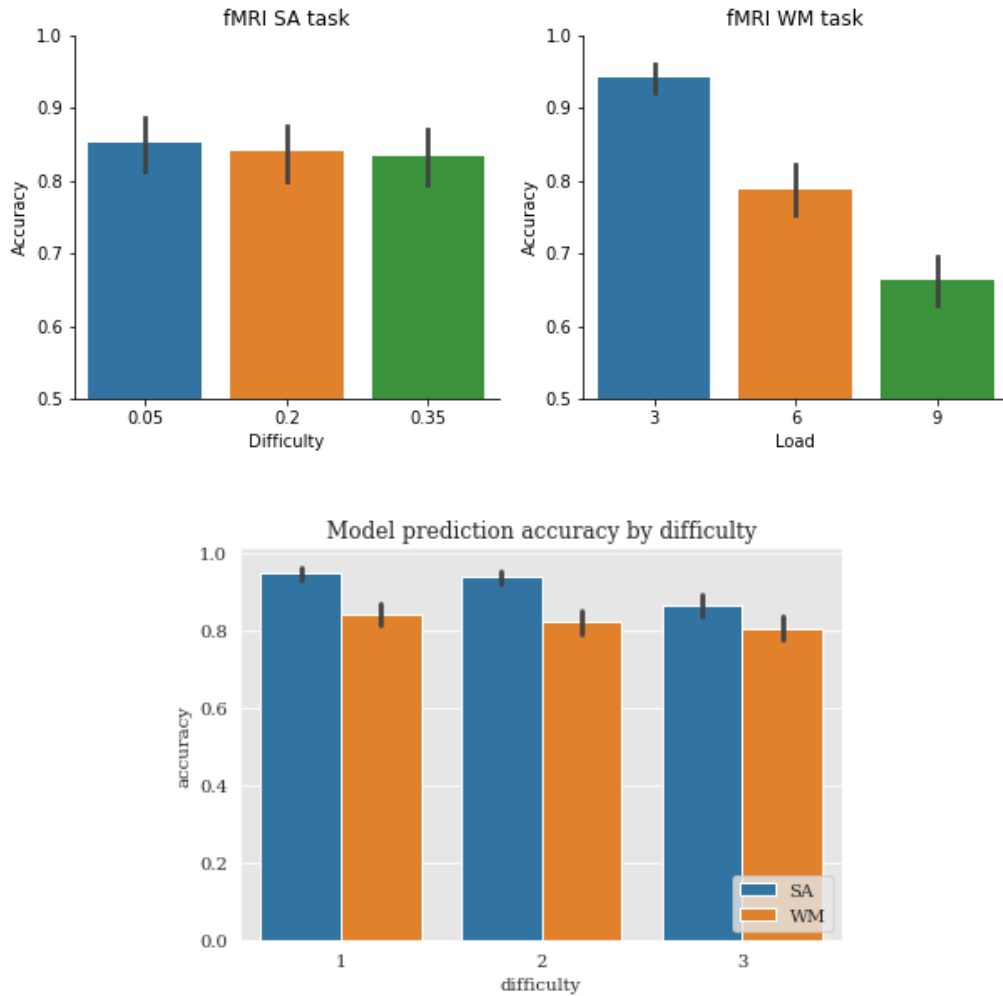


Figure 24. Behavioural accuracy in each task from fMRI study (top), and RNN (bottom). Error bar denotes 95% confidence interval. Note that for fMRI WM task, participants used a continuous response. Similar to the standard adopted in the imaging analysis, correct trials were defined as response error within 90 degrees from the true answer. For RNN plot, the difficulty level was rearranged so that in both tasks higher value represents harder conditions.

4.3.2 Relationships between representations of tasks by RNN

To understand the neural representations of tasks developed through the training process, a new batch of trials spanning the condition space was presented to the trained network for predictions and the activation variances of all recurrent units were recorded. In order to quantify each neuron's selectivity to one task relative to the other, a measure based on these activation variances were calculated: the fractional task variance (FTV).

Figure 25 shows the FTV distribution of all RNN neurons, which was unimodal that was centred around zero. According to (Yang et al., 2019) this kind of distribution suggested a mixed relationship between the pair of tasks, implying that both could share a subset of neural circuitry. On the other hand, the small

tails on both ends could be evidence for additional task-specific neurons existing in the network, which were functionally more dedicated to one task than the other.

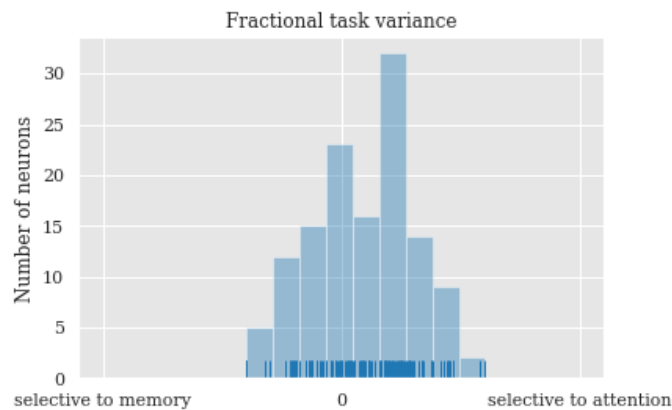


Figure 25. Selectivity of RNN neurons to SA and WM tasks.

4.3.3 Clustering produced two functionally separable subsets of network neurons

Using K-mean clustering on task variances (Yang et al., 2019), we explored whether the model organized itself into functional clusters that serve for different computational mechanisms. Figure 26 shows the results for K ranging from 2 to 9. The two-group solution generated the best silhouette score and therefore was used in subsequent analysis, which was not surprising as there were only two tasks. Cluster 1 contained 71 recurrent units and cluster 2 contained 57 units.

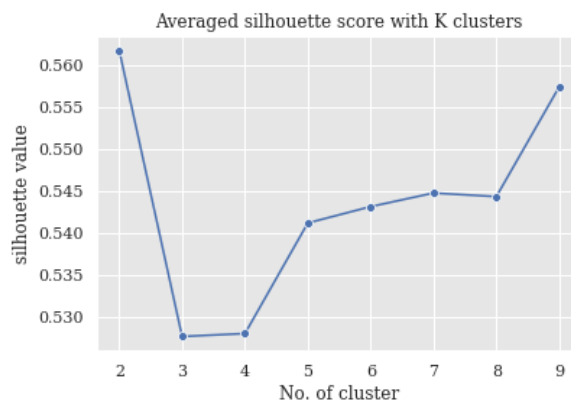


Figure 26. Mean silhouette score with each K-number solution. Higher values represent clearer separation of clusters.

To further probe whether these two subgroups of neurons serve distinct mechanisms, we inactivated one cluster at a time and monitored the decrease in performance. When lesioning cluster 1, accuracy of SA task dropped to 53% from 100% when the network was intact, whereas the accuracy of WM task was reduced to 45% from 93%. Similarly, SA accuracy went down to 48% with cluster 2 inactivated compared to 57% for WM accuracy. In short, lesioning these two clusters has a subtly different impact on the relative performance of the two tasks. Figure 27 shows the difference in task accuracy between lesioning an entire cluster versus the same number of random units. Losing cluster 2 had a significant effect on SA performance, as accuracy drop was consistently lower than randomly inactivating neurons (mean difference = -0.02, $t(98) = -8.0$, $p < .001$). Similarly, lesioning cluster 1 significantly affected WM performance (mean difference = -0.05, $t(98) = -22.83$, $p < .001$). By contrast, the performance decrease for SA task when cluster 1 was lesioned (mean difference = 0.03, $t(98) = 15.00$, $p < .001$, and that for WM task when cluster 2 was lesioned (mean difference = 0.07, $t(98) = 35.60$, $p < .001$) seemed to be more a consequence of losing a large number of units, since the accuracies were actually higher when disabling an entire cluster, relative to lesioning random units.

To tease apart the above results more, we gradually increased the number of units being inactivated within a cluster and measured the effect on model performance. SA task accuracy as a function of number of lesioned units did not differ between the two clusters (Figure. 28), indicating that neither made particular larger contributions to the computational processes underlying the attention task. On the other hand, for the WM task, inactivating cluster 2 neurons led to more catastrophic result than inactivating the same amount in cluster 1. This implies that the WM task was more sensitive to disruption of cluster 2 in contrast to cluster 1, while SA relied equally on both clusters.

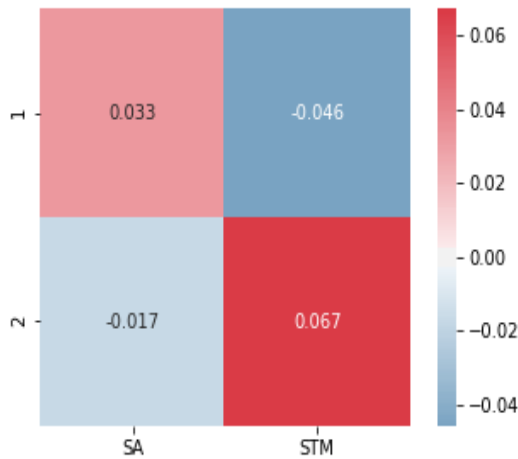


Figure 27. Accuracy difference between lesioning a cluster and same number of random units. Positive value indicates that inactivating a cluster had better performance than inactivating random units, vice versa.

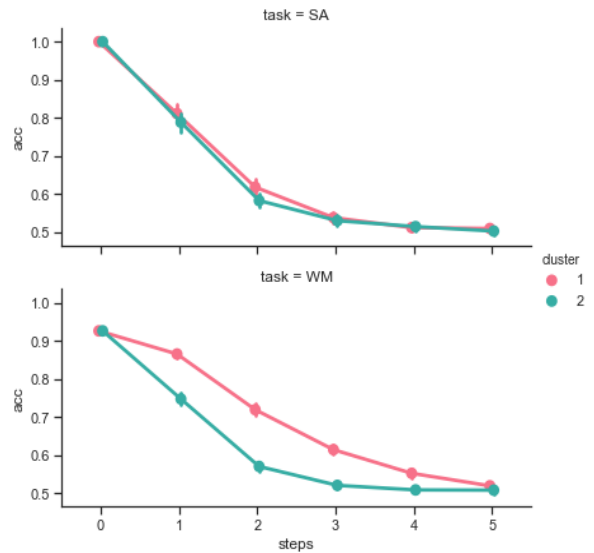


Figure 28. Results of sequential lesioning. At step 0 no units were disabled therefore the network remained intact.

4.3.4 Representational similarity analysis of the clustered neurons

Having shown that the network model develops different functional modules as it learns the tasks, we subsequently investigated whether they also exhibit distinct representations of the tasks using representational similarity analysis (RSA). Important context for this section: a value in a RDM reflects the *dissimilarity* (1-correlation) between the activity patterns associated with a given pair of conditions.

Cluster 2 neurons showed more distinct between-task representations compared to cluster 1 as highlighted by the higher values in the lower left corner (Figure. 29&30) (mean difference = -0.21, $t(16) = -11.44$, $p < .001$). Cluster 2 contained some more nuanced pairwise relationships, such as the gradient of dissimilarities from attention level 1 to 3 relative to the WM conditions, which seemed to suggest that the more difficult the SA task became, the more differently represented it was to the memory task by this cluster (Figure. 30). Altogether, this could indicate that cluster 2 neurons served somewhat more flexible functions within both tasks, as their activity patterns were able to differentiate between the conditions. In contrast, cluster 1 was more equally involved in both tasks, evidenced by the lower between-task dissimilarity. Finally, we can corroborate this speculation by plotting the FTV distributions of each cluster separately (Figure. 31). Cluster 1 neurons were more active towards SA task compared to WM, combined with the low between-task dissimilarity, which suggests that the computational process subserved by this subset of units remained relatively the same across tasks, and the process was potentially attentional in nature. On the other hand, cluster 2 neurons were more selective to WM task. Although it also played important role in SA, it was more essential to the success

of WM. Indeed, in the lesioning analysis we saw that losing cluster 2 had more serious effect on WM performance compared to losing cluster 1, but losing cluster 2 had a comparable effect on attention accuracy as well. It is noteworthy that it is unlikely cluster 2 was involved in similar processes in both tasks as between-task representations were highly distinct (Figure. 30)

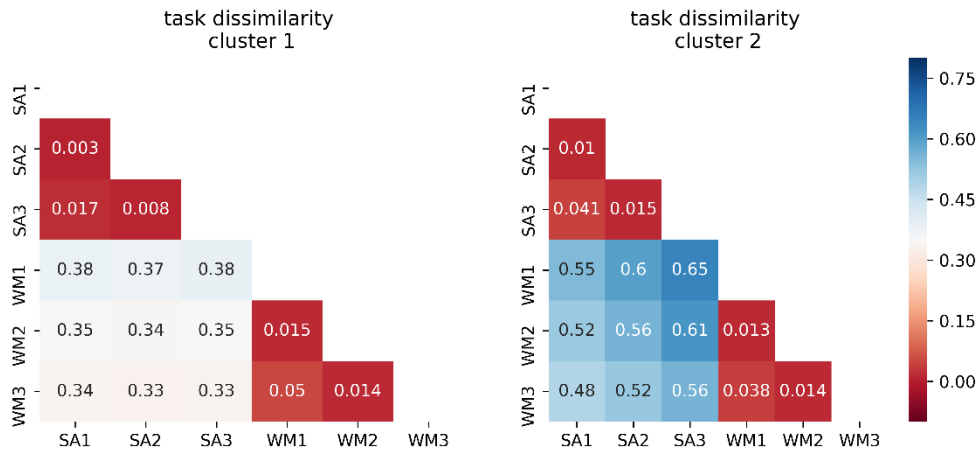


Figure 29. representational dissimilarity matrix of cluster 1 neurons between task conditions. SA1 = SA task level 1; WM1 = WM task load 1

Figure 30. Representational dissimilarity matrix of cluster 2 neurons between task conditions.

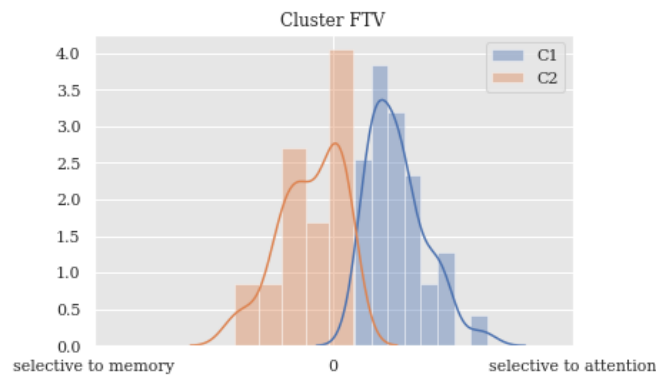


Figure 31. Neuron selectivity by clusters

4.3.5 Relating the RDMs of the network model to brain activity

In the following section, we tried to understand whether the neural representations of the tasks as captured by neuroimaging data share any common features with that of the RNN. The cerebral cortex was partitioned into 100 regions of interests (ROIs), within which a brain RDM was assembled using individual's contrast maps. Next, we quantified the match between the brain RDMs with those of the

RNN clusters, resulting in two brain maps per subject that denotes how similar each brain region's representation was to cluster 1 and 2, respectively (Figure. 32 shows the group averaged maps across participants). Permutation testing was performed to determine the significance level for each ROI but was not reflected on the map as all correlations were highly significant. It is perhaps due to the limited number of conditions that even the lowest correlation throughout the cortex is still larger than 0.5. However, certain areas exhibited *differences* in their correlational strength to the two cluster RDMs.

Figure 33 shows the areas that were more similar to cluster 1 than cluster 2, including the bilateral visual cortex, left anterior insula, left anterior superior frontal gyrus (SFG), left posterior middle frontal gyrus (MFG; premotor cortex), intraparietal sulcus (IPS), superior parietal lobule (SPL) and inferior temporal gyrus (ITG), all in the left hemisphere. (Table 8). We did not find any region that was more cluster 2-like than cluster 1.

Table 8. Regions of interest more similar to neurons in cluster 1 compared to cluster 2.

Brain region (BA)	Schaefer atlas	MNI coordinates of ROI centre			<i>p</i> -values (FDR-corrected)
		x	y	z	
L Fusiform gyrus (17/18)	LH_Vis_2	-26	-76	-14	2.58×10^{-4}
L MOG (17/18)	LH_Vis_4	-26	-96	-4	2.58×10^{-4}
L Calcarine fissure (17/18)	LH_Vis_5	-6	-92	-2	2.73×10^{-5}
L Calcarine fissure (18 /19)	LH_Vis_6	-12	-66	6	6.09×10^{-3}
L MOG(18/19/37)	LH_Vis_8	-26	-88	20	1.27×10^{-3}
L ITG (21/22)	LH_DorsAttn_Post_1	-46	-58	-12	1.75×10^{-2}
L SPG (7)	LH_DorsAttn_Post_2	-24	-68	50	1.45×10^{-3}
L IPS (39/40/7)	LH_Cont_Par_1	-38	-52	46	2.32×10^{-2}
L Insula cortex (13)	LH_Default_PFC_1	-34	22	-10	3.89×10^{-3}
L SFG (10/9/8)	LH_Default_PFC_4	-24	60	-2	2.32×10^{-2}
L MFG (6/45/46)	LH_Default_PFC_6	-40	14	48	6.09×10^{-3}
R Fusiform gyrus (17/18)	RH_Vis_2	28	-66	-12	2.58×10^{-4}
R IOG (17/18)	RH_Vis_4	22	-96	-4	2.16×10^{-5}
R MOG (18/19/37)	RH_Vis_7	36	-82	16	2.16×10^{-5}
R Cuneus (18/19)	RH_Vis_8	12	-86	30	1.20×10^{-2}

Note. SFG = superior frontal gyrus; MFG = middle frontal gyrus; SPG = superior parietal gyrus; ITG = inferior temporal gyrus; LH= left hemisphere; RH=right hemisphere; Vis= visual network; DorsAttn = dorsal attention network; Con_par=frontoparietal network; Default = default mode network;

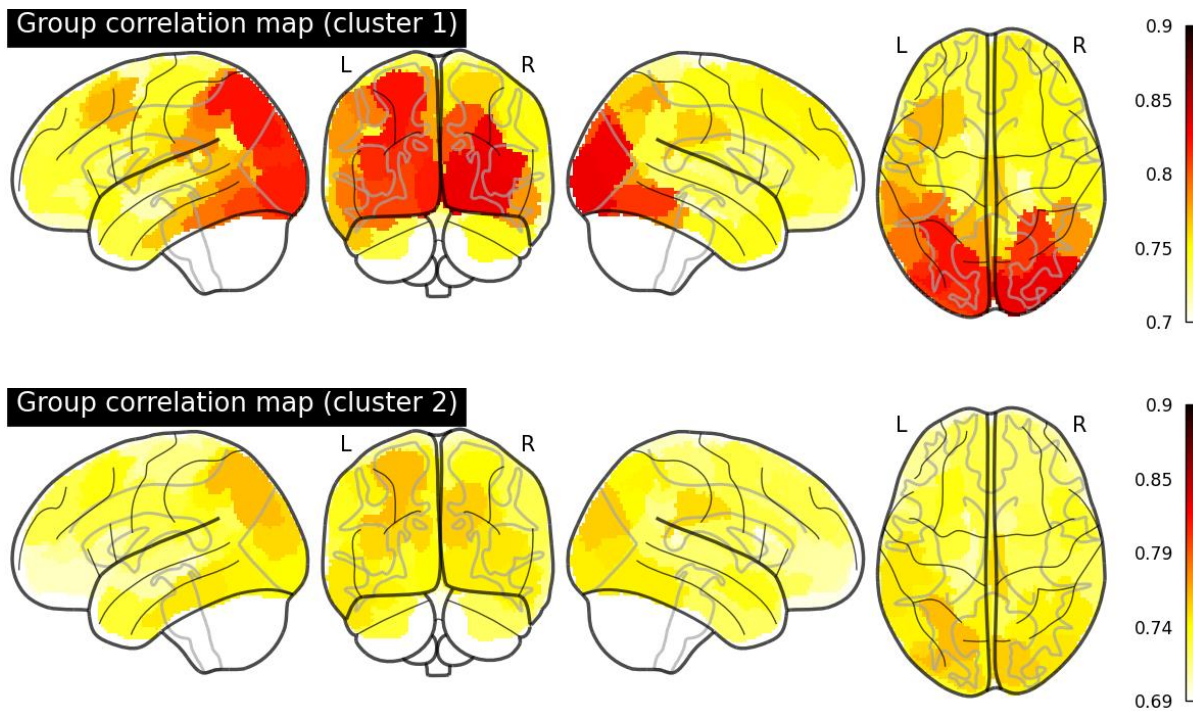


Figure 32. Strength of averaged group correlation between brain and RNN cluster RDMs

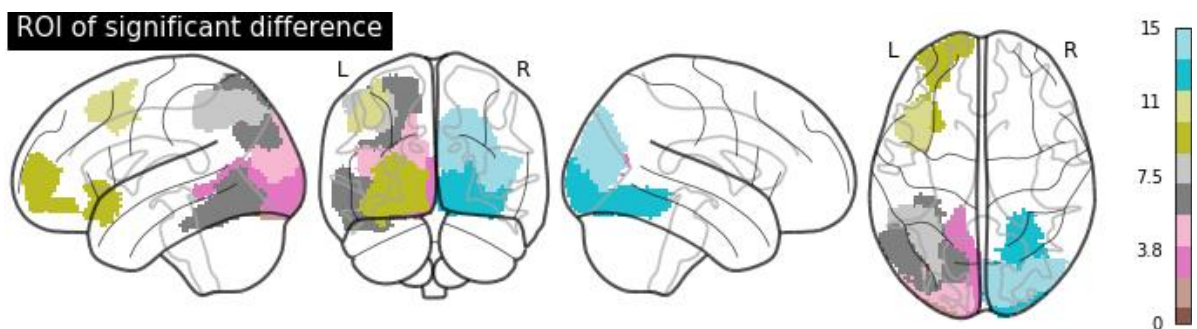


Figure 33. ROIs that showed significant difference in brain-RNN similarity (cluster 1 > cluster 2. Colours only denote distinct parcels defined by the Schaefer atlas.

4.4 Discussion

This chapter used a computational modelling tool and sought to address whether artificial neural networks could be trained to perform two cognitive tasks and subsequently used to explore the relationship between the underlying representations for the two tasks. A simple biologically constrained model with recurrent neurons was able to perform both the SA and WM tasks at the same time. Neurons in the recurrent layer developed different selectivity towards each task but overall exhibited a pattern

of an overlapping neural circuit in service of both tasks. It is possible that there could have been multiple underlying clusters which had specific roles, but we found the best evidence for two broad sets of neurons within the RNN. These were involved in both tasks, but systematically lesioning one of the clusters did have a disproportionately large impact on WM performance, whereas as both clusters seemed to be similarly crucial for SA performance. Next, we used a representational dissimilarity analysis to explore the similarities in activity patterns across human fMRI and the RNN. This revealed a network consisting of prefrontal, parietal regions and the visual cortex that are typically included in a top-down control network, the activity profile of these regions mirrored that for one of the clusters in the RNN – the one that contributes equally to both tasks.

4.4.1 The emerging network structure suggests overlapping neural circuits supporting SA and WM

A majority of the network neurons were equally involved in the computational processes of SA and WM, suggesting that both tasks shared the same neural circuits. It is in theory still possible that the same subset of neurons contributed to completely distinct functions between both tasks but happened to be similarly active to each. However, since the variance measure we chose captured the amount of stimulus information a unit conveyed during a task (Yang et al., 2019), a more probable interpretation is that the neurons were responsible for certain aspect of the processing necessary to perform both paradigms. This is analogous to the defining feature of WM, namely its limited capacity, a property shared by attentional mechanisms that select and enhance visual information (Awh & Jonides, 2001; Chun, 2011; Gazzaley & Nobre, 2012). One theory of WM proposes that it can be viewed as limited-capacity attentional focus that operates across areas of activated long-term memory (Cowan, 1998, 2001; Cowan et al., 2005; Rerko & Oberauer, 2013). In line with this view, we would expect a system optimized to perform both tasks to develop neural circuits that underlie such mechanistic overlap.

On the other hand, the network structure also suggests some subtle functional separation, at least for the WM task, as evidenced by the diverging consequences of sequential lesioning. Cluster 2 neurons seem to play a more important role in the success of WM performance relative to the other cluster and to the SA task. They were also more sensitive to the distinction between the tasks as evidenced by the representation matrix showing higher level of between-task dissimilarity, suggesting that the representations of SA and WM by this cluster were further distinguished compared to neurons in cluster 1. Although unable to determine the exact computations they are involved with, this is reminiscent of the differences between SA and WM in terms of the distinction between perceptual and post-perceptual stages. There is a general consensus regarding the mechanisms at work when perceptual input is present, which requires the higher-order control regions such as the frontoparietal network to exert top-down signals that are projected to the posterior sensory areas. Activities of task-relevant information are thus enhanced while irrelevant information are suppressed, a mechanism having been found for both

encoding stage (Gazzaley et al., 2007) and visual selective attention (Couperus & Mangun, 2010; Serences et al., 2009). As for mechanisms underlying maintaining information during the delayed period of WM, there may be multiple neural mechanisms in addition to the top-down modulation at work in parallel (Masse et al., 2020). For example, it is postulated that PFC neurons generating persistent activity in the absence of external stimuli is likely the fundamental basis of WM (Rainer et al., 1998; Romo et al., 1999). Recent studies posited that the function of the PFC was more diverse, also representing higher-order information such as task rules and goals (D'Esposito & Postle, 2015; Harrison & Tong, 2009). Moreover, stimulus information during delay phase could be detected via multivariate pattern decoding in sensory regions whereas could not be found using univariate statistical approach (Serences et al., 2009), which is in contrast to the case in SA/encoding phase and possibly suggests a qualitative difference between SA and WM maintenance. Likewise, the network model may have developed dedicated circuitry to perform computations related to a WM-specific process, whereas functions common to both tasks (such as attentional mechanisms) were supported by neurons in cluster 1, which showed less between-task representational difference.

4.4.2 Brain regions showing similarity to the overlapping RNN neurons coincides with a frontoparietal network

Our findings showed that the task representation similarities of cluster 1 neurons, which seemed to represent the tasks more similarly, fitted significantly better to regions including the anterior part of SFG (lateral PFC) and premotor cortex, anterior insula, intraparietal sulcus (IPS), ITG and bilateral visual cortex, relative to chance and to neurons in cluster 2. If this particular set of neurons indeed underpin computationally shared mechanisms necessary to both SA and WM, as implied by the lesioning analysis, similar neural substrates would be expected to be found in studies using paradigms that tap into both domains. Consistent with this view, neuroimaging studies assessing the anatomical overlaps between SA and WM have converged to reveal similar regions, for example the frontoparietal control network (Corbetta & Shulman, 2002; Culham et al., 2001; Sestieri et al., 2012; Wager & Smith, 2003; Wallis et al., 2015). LaBar et al. (1999) used a within-subject conjunction analysis between a verbal N-back and a spatial attention tasks, and reported common activations in IPS, precentral sulcus, frontal eye field (FEF), middle temporal gyrus (MTG), and insula. Perhaps even more powerful is evidence of functional interaction between the two sets of processes, since anatomical overlap such as in the above study does not necessitate sharing of neural mechanisms, due to possible regional specialisation that is undetectable with fMRI spatial resolution (Nieder, 2004) or mixed selectivity of neurons that allows them to perform adaptive functions according to specific task demands (Raposo et al., 2014). To this end, several imaging studies have combined visual search and object- and spatial-based delayed-response paradigms into one single task to assess the interactions associated when each process is manipulated independently (Fusser et al., 2012; Mayer et al., 2007). These functional

interactions have appeared in posterior regions including the parietal and visual cortices, which bears partial resemblance to our results.

4.4.3 Functional roles of the identified brain networks

It is possible that the regions we have identified constitute components of several large-scale networks rather than serving one unified function, as SA and WM emerge from interactions among multiple systems that support basic functions such as perception, action and higher-order task coordination (D'Esposito & Postle, 2015; Eriksson et al., 2015). Based on observations across studies and meta-analyses, WM and SA activation closely overlaps with two intrinsic networks: dorsal attention network (DAN) and frontoparietal control network (Rottschy et al., 2012; Schaefer et al., 2018; Vincent et al., 2008).

The SPL, IPS and ITG are often indicated as key sites within the DAN, the activity of which increases during externally directed attention, such as stimulus processing and spatial orienting, suggesting that it underlies a mechanism for orienting attention to the external environment (Corbetta & Shulman, 2002; Sestieri et al., 2012). Specifically, biasing of sensory areas according to current task demand has been established to emerge from higher-level areas in the IPS and FEF. Using TMS and fMRI, Ruff and colleagues demonstrated significant neural modulation in visual cortex activity after perturbation of DAN regions (Ruff et al., 2008, 2009). It has also been indicated that DAN may play similar roles during memory encoding, “gating” what information would enter WM (Rutman et al., 2010) as well as shifting attention towards mental representation during retention as a result of post-perceptual cuing (Wallis et al., 2015). Additionally, varying complexity and location of to-be-remembered items, Xu & Chun (2006) also found that activity in IPS tracked memory load, reflecting objects in mind from up to four spatial locations irrespective of complexity. The IPS was also shown in our fMRI experiment to reflect difference in neural overlap between SA and WM. Taken together, it supports the notion that the DAN is engaged in deploying attentional resources in various cognitive domains, which may explain why IPS and ITG were identified in our analysis as areas where between-task representations were similar. Aside from that, although ITG was labelled as part of the DAN according to (Schaefer et al., 2018) and (Yeo et al., 2011), its exact functions remains unclear. It could be that ITG truly drives some attentional mechanism that work across the SA and WM domains, as indicated by patient study in which people with ITG damage showing object-centred neglect (Corbetta & Shulman, 2011) and WM-related studies in which the same regions are associated with increasing memory demand (Rottschy et al., 2012). Others have purported that ITG could be activated due to its involvement in the ventral pathway for object recognition. Indeed, WM studies using more complex stimuli such as faces and scenes often observed activity in this region (Jenness et al., 2018; Landau et al., 2004) compared to those using simple stimuli. Notably, we also identified this region in the fMRI study (Chapter 3) for both task vs.

baseline contrasts, therefore implying that although the stimuli we used were in no way complex, it nevertheless could benefit from certain object recognition mechanism.

Within our results lateral PFC had higher similarity to the representational pattern of RNN neurons underlying more common processing, relative to the cluster of neurons that had a more selective pattern. This is not surprising as it is often implicated as substrates for flexible cognitive control (Duncan, 2001; Stokes et al., 2013). In line with the theory, areas within frontoparietal network(s) generate top-down signals to modulate posterior sensory regions at the presence of competing stimuli or mental representations. fMRI studies assessing dorsolateral PFC regions involved in modulation during WM encoding have found it serves in processing relevant stimuli as well as in filtering distractors through functional coupling with the posterior sensory areas (Gazzaley et al., 2007; Zanto et al., 2011). The nature of the signals of PFC is likely to be distinguished from that of the DAN, by addressing the kind of information encoded during WM tasks. Multivariate pattern decoding reveals that neuronal populations in PFC initially code for the physical properties of the stimulus during encoding, and afterwards switch to other relevant information, i.e. whether it was a target or a distractor. Consequently, PFC is not necessarily responsible for storing sensory representation per se, but instead has access to that stored information and shows preference for other task-related contents, such as ongoing rules (Stokes et al., 2013). Other studies have since corroborated this finding, with multiple studies showing that PFC is involved in representing higher-order information including task rules, goals or abstract representations of categories, instead of item-specific information (e.g. Sreenivasan et al., 2014). This is also consistent with the notion that coding of task variables in PFC neurons can adapt to accommodate changes in behavioural context, which is ultimately critical for exerting cognitive control and flexibility (Miller & Cohen, 2001; Stokes et al., 2013). Thus, it is possible that a subset of RNN neurons in cluster 1 reflected representation of more abstract task-related information, such as task rule and/or goals that could be shared across both tasks despite some apparent different task components.

4.4.4 Limitations and future directions

The current work only represents an initial attempt to computationally model the task relationships between WM and SA, and perhaps inevitably we encountered multiple limitations. We have discussed in detail the various neural mechanisms that could drive activity patterns in certain frontoparietal areas to exhibit similarity structures akin to a cluster of units in the RNN, which is characterized by a closer representation between the two tasks. However, due to the fact that there were merely six conditions across both paradigms giving only 15 different pairwise dissimilarity values in total, the resulting representational matrix may not effectively sample the dimensions on which all relevant processes supporting the cognitive functions vary. And as a result, it is challenging to further differentiate the nature of the computations within these distributed regions. Our work therefore would benefit from a condition-rich design (Kriegeskorte et al., 2008) in which the number of effective experimental

conditions (i.e. brain states to be distinguished) is large. To this end, variations could be introduced to each of the current paradigms, such as enriching the stimuli types, modalities or even the task designs to include more measurements for SA or WM. Secondly, we have demonstrated that RNNs are capable of performing complex cognitive tasks in a manner that is behaviourally and neutrally comparable to human participants, yet we have not examined the detailed mechanisms (i.e. patterns of neuronal connectivity or the temporal dynamics of network activity) by which the model achieves the task goals. A systematic dissection of trained networks remains one of the most important goals if RNNs are to provide useful insights into the operation of biological neural circuits. Thirdly, as detailed in the Method section, there were some deviations between how we simulated the tasks and those implemented in the fMRI study. This was done to try and simplify things because of our time constraints, but may have hindered the comparability of task representations between the two modalities. Furthermore, many assumptions about the model parameters had to be made in order to complete the training process. Although we had based these decisions in the previous literature which had heavily inspired the current study. So we think this was a reasonable starting point, but in the future we aim to fully explore the effect of parameter choices on model performance and structure. Lastly, we chose to continue the cluster analysis with $K=2$, as it resulted in the cleanest division of neurons according to the pre-defined criterion (the Silhouette Coefficient). However, the difference in coefficient between different clustering solutions is marginal, and all are above 0.5 (which suggests a reasonable division), and using a solution with more clusters could have provide additional granularity needed to reveal more idiosyncratic computational processes, specific to either task, which may have been masked under the current coarse solution.

In conclusion, as a platform for theoretical investigation, trained RNNs provide a unified setting in which diverse cognitive computations and mechanisms can be studied. The present results lay a valuable foundation by showing that relatively simple implementation of computational models can serve as a convenient proxy for biological brains. With the complete access to the network dynamics and the ability to manipulate them systematically, RNNs may prove to be a promising tool for bridging machine learning, neuroscience and behaviour. For the purposes of the current thesis, this approach broadly indicated that different neurons had relatively subtle preferences for SA and WM, with one cluster being more crucial for WM and the other underlying both equally. This broadly shared computational architecture mirrored activity patterns across a range of frontoparietal areas, derived from human fMRI data.

Chapter 5: Does training selective attention improve working memory?

5.1 Introduction

To explore the relationship between selective attention (SA) and working memory (WM) we have so far adopted neuroimaging and computational modelling to study potentially overlapping mechanisms. In this next chapter we investigated the malleability of these mechanisms. A crucial part of understanding the interplay between two related systems is to test how changing one system influences the other, and, if so, what kinds of individual differences moderate the degree of this change. In short, are the two systems causally related, such that altering one will change the other?

As outlined earlier in this thesis, the role of SA within WM has been a focus of the field since the earliest theories of WM. Information held in mind can capture attention, and vice versa (for reviews see Awh & Jonides, 2001; Kiyonaga & Egner, 2013). The developmental literature has also demonstrated a close relationship between these two systems, with highly correlating developmental trajectories across domains, and frequency co-occurring neurodevelopmental difficulties in both SA and WM. Roome et al. (2014) reported an age-related increase in WM capacity alongside improvements in attentional selection, directly relating these constructs' developmental trajectories. Similarly, there has been accumulating evidence that older adults often experience deficits in WM and SA simultaneously (Craik & Salthouse, 2011) and that SA difficulties may underlie much of the WM deficits, with consequences for multiple stages of memory processing. Specifically, during pre-encoding phase older adults have been shown to exhibit reduced top-down modulation in category-selective areas (e.g. fusiform area for face stimuli) and lack the subsequent memory benefits that would be expected for this preparation (Bollinger et al., 2011). Likewise, there are also age-related deficits in the ability to encode in the context of distractors, which was in turn associated with the degree of the deficit in overall WM performance (Gazzaley et al., 2005b). Finally, older individuals also experience greater WM interference when they switch between tasks during memory retention. This increased interference is related to reduced capacity to disengage from the interrupting information and to re-establish functional connectivity supporting the primary memory task (Clapp et al., 2011). Similarly, WM-related impairments are typically identified in children and adults with diagnoses of attention deficit hyperactivity disorder (ADHD) (Martinussen et al., 2005; Willcutt et al., 2005), leading some to propose that this condition reflects a broader executive function difficulty (Rappaport et al., 2001). Although the causality remains unclear, the effect sizes tend to be small, and this reductive approach to ADHD has been roundly critiqued in the literature.

Given the co-occurrence of WM and SA-related difficulties across various populations – during childhood, typical ageing and neurodevelopmental disorders – studies employing cognitive training have explored whether repeated practice on SA tasks will generalise to WM tasks (Greenwood &

Parasuraman, 2016; Greig et al., 2007; Haut et al., 2010; López-Luengo & Vázquez, 2003; Mishra et al., 2014; Wass et al., 2012). The most common approach is to use a mixed diet of attentional training, with adaptive tasks capturing attention control, sustained attention, selective attention, task switching and inhibition. The thinking is that this varied battery approach will maximise potential gains and boost the chances of transfer. Some studies have found improvements in more distally related tasks, such as WM (Greig et al., 2007; Haut et al., 2010) and cognitive flexibility (Greig et al., 2007; López-Luengo & Vázquez, 2003). However, these studies are in the minority, relative to those that focus on training WM itself, which has been shown to generalise to measures of SA, such as Stroop (Klingberg et al., 2005; Morrison & Chein, 2011), although sometimes showing negative results (Thorell et al., 2009; Van der Molen et al., 2010; Westerberg et al., 2007), as well as to measures of sustained attention (Lundqvist et al., 2010; Thorell et al., 2009; Westerberg et al., 2007) and task switching (Lundqvist et al., 2010). Despite these tantalising individual studies, many of them have been widely critiqued on methodological grounds – with prominent critiques including no active control groups, small sample sizes and biased allocation of trainees to group (Redick et al., 2013). By contrast, meta-analyses have been more sobering. Melby-Lervag and Hulme (2013) concluded from a pool of 23 randomized controlled studies that WM training can produce sustained *but limited* improvements in visuospatial WM, with no generalization to SA measures. Likewise, Shipstead et al. (2012) showed highly consistent findings, suggesting that idiosyncrasies in the methodological choices, such as the number of measurements used for an ability of interest, the type of control group (e.g. no-contact vs. active), etc. have given the appearance of far transfer, whereas in reality there is little evidence for gains on tasks beyond those with paradigm structures highly similar to the training activities themselves.

There are a number of competing theories attempting to explain the mechanisms underlying training effects and the (limited) scope of transfer. One interesting candidate, specifically relevant to the role of SA, proposed that perceptual or WM training protocols that yield transfer to other tasks, particularly general fluid reasoning, all tax mechanisms of distractor suppression and recruit a dorsal attention network (Greenwood & Parasuraman, 2016). Some converging evidence came from perceptual discrimination practice, which led to a decreased ERP component associated with distractor suppression and correlated accuracy gains on a WM task (Berry et al., 2010; Mishra et al., 2015). This was mirrored by reduced functional connectivity between the superior parietal cortex (part of the dorsal attention network) and sensory processing regions (Lewis et al., 2009; Strenziok et al., 2014), perhaps reflecting better efficiency on top-down control of attention after training. It is consistent with the work by Barnes et al. (2016) examining neurophysiological changes following WM training in children. In their data, the magnitude of gains on an unpractised WM task was related to alterations in phase-amplitude coupling between alpha-band oscillations in frontoparietal areas and gamma-band activity in inferior parietal, temporal and high-order visual processing areas. This relationship possibly reflects the top-down regulation of attentional demands and task-set information (Duncan & Owen, 2000; Sauseng &

Klimesch, 2008). Considered together, it is possible that training-induced generalization can happen between the domains of SA and WM. However, there are multiple methodological pitfalls to be wary of, including control groups and carefully matched tasks.

One source of variance that might contribute to the apparent inconsistency in training findings is individual differences. Various factors such as age, motivational level and pre-existing ability, over and above differences in study design, could impact training gains (Morrison & Chein, 2011; Wass et al., 2012; Zinke et al., 2014). This issue was partly addressed in part of my PhD work, which is not included in the thesis (see Zhang et al., 2020) in which we reported clusters of children who benefited differentially from WM training, and these group differences seemed to be at least partially associated with baseline non-verbal reasoning performance. Our interpretation was that this may reflect the ability of participants to generate novel strategies, which in turn drive training gains and transfer. The behavioural experiment in Chapter 2 expanded upon the idea that tasks could be differentially associated across individuals – in short, that some participants deploy more shared resources for SA and WM than do other participants. Consequently, the fMRI study detailed in Chapter 3 identified differences in the degree of overlapping neural mechanisms, depending on the behavioural profile of the individual. One possibility is that despite the relatively close relationship between these two domains at a group level, these inter-task relationships are variable. In sum, there are two potentially related pieces to the puzzle – preceding work highlighting the substantial individual differences in training effects, and that highlighting inter-subject differences (both in performance and neural substrates) in the extent to which SA and WM share underlying resources. This final empirical chapter brings these two pieces together, by testing whether these individual differences in task overlap will moderate transfer effects. In other words, whether individuals who have stronger task overlap show better transfer between those tasks.

This study uses repeated training to enhance SA in order to: 1) probe whether it would induce changes in WM performance, and 2) whether the relationship between baseline abilities (correlated vs. dissociated) moderates the extent of transfer. We adapted our own SA task, developed in Chapter 2, as the training paradigm and kept the WM task as one of the assessments to keep the task-specific features consistent (e.g. stimuli presentation). Additionally, multiple transfer tasks drawing on SA, WM and Gf were included to offer a more comprehensive assessment battery (Morrison & Chein, 2011; Shipstead et al., 2012). We selected individuals with closely related SA and WM performance, versus a group who are good at SA but not at WM. We predicted that where these two abilities are more correlated, those participants' WM skills would benefit most from training their SA. Whereas transfer would be more limited in individuals whose WM performance seemed to be less related to their SA skills.

5.2 Methods

5.2.1 Participants and procedures

Five-hundred and seventy participants who were between 18- and 40-year-old were recruited online. The inclusion criteria were that at the time of study they had an approval rate (i.e. how often their online test results were accepted by experimenters) above 90%. Participants were subsequently pre-screened on Prolific using the newly-developed SA and WM tasks deployed in Chapter 2, 435 of which were those who participated in the online behavioural study described in that chapter. From this pool of participants, we identified two training groups and one control group. The first training group consisted of 28 participants (13 female, mean age: 30.14) who scored in the top 35 percent on both attention and memory tasks within the reference population recruited in Chapter 2 and received 3 sessions of intense attentional training (the highWM group). The second training group were 37 participants (24 female, mean age: 31.75) scoring within the top 35 percent on attention task and within the bottom 35 percent on the memory task, and they received the same attentional training (the lowWM group). The active control group (20 participants, 5 females, mean age: 31.85) had the same pre-screening profile as the highWM group but instead went through 3 sessions of phonological sensitivity training (“listen-up” task). Attrition rate was relatively low as of 101 people who had entered the training phase, 85 continued to finish the entire study.

We aimed for approximately 30 participants per group, as this meets the requirement of a power calculation set for 80% power in a mixed design, with groups as between-subject variable and time as within-subject variable. The power analysis was done using G-power for a within-between interaction and effect size (η^2) was set to .06 (medium). Due to time constraint owing to the pandemic and the rarity of these profiles, the size of the control group was slightly smaller in comparison.

Before and after the training sessions, participants completed pre- and post-assessments that consisted of 1-1.5 h of cognitive assessments. The maximum time allowed between pre- and post-assessments was 10 days and there was a mandatory 12-hour gap between each training session and the post-training test. Participants with the highWM profile were randomly allocated to either the training or control group. They were not told about the other conditions. Each of the three training sessions included 8 blocks of 54 SA trials or 20 listen-up task trials (which are much slower), amounting to 20-25 minutes per session. To keep the task constantly challenging, we used a simple adaptive procedure of moving up one difficulty level when block accuracy is above 80% and moving down when it is below 70%. Participants started at the same level and returned to the one they previously achieved at the beginning of the next session.

5.2.2 Training tasks

SA task. In the SA task, the goal is to identify whether a target stimulus is present among an array of distractors. Participants indicated by pressing “Z” or “M” on the keyboard within 2 seconds. Seven levels of difficulty were used, as the trainees progressed to a higher level if they achieved 80% accuracy in a block. Levels were made more difficult by making the distractors progressively more diverse in orientation, thus moving the search gradually further away from a pop-out effect, and towards a conjunction search (Duncan & Humphreys, 1992). Specifically, the between-distractor difference ($D_{between}$) is defined by $D_{between} = x * 180^{\circ}$ where x range from 0.05 to 0.35 with .05 increment, so that the easiest level corresponds to 9° discrepancy between the orientations of both distractors and 63° on the hardest level. This manipulation was validated in Chapter 2, where increased RTs were explained by greater between-distractor difference. Other aspects of the visual display, such as size and colour of the target and distractor shapes were exactly the same as detailed in Chapter 2.

Listen-up task (phonological sensitivity). We selected this active control paradigm on the basis that we thought it would tax separate underlying cognitive abilities, relative to the other training task. A detailed description of the listen-up task can be found in Davis et al., 2020, here only a brief introduction is included. Each trial involves presentation of three spoken syllables, with the first being the referent word and the next two being two potential target syllables for a 2AFC task. Participants first heard an original, unmodified recording of one of the referent words (e.g. bear) spoken in a female voice, accompanied by a photograph of the referenced object to help support accurate perception of the referent word and minimize short term memory demands in retaining that word for the rest of the trial. The referent words in a block were selected at random without replacement from the set of 60 words. It was followed by a pair of acoustically morphed words generated by combining the referent word and a similarly sounding word (e.g. bear-pear). The pair was always constructed so that if word A contained 80% bear and 20% pear, word B would contain 20% bear and 80% pear, resulting in an acoustic difference of 60%. Participants were instructed to choose which of the synthesized word was more similar to the referent word. The task became harder by minimizing this difference (i.e. 55% bear and 45% pear) as participants grew better at judging the correct word.

5.2.3 Transfer tasks

The transfer tasks included six computerised tasks designed to assess SA, WM and fluid intelligence. In particular, as in Chapter 2, our own WM paradigm used identical stimuli to the SA task. Other established STM/WM tasks were chosen to circumvent biases from a single measurement (Moreau et al., 2016), and test for the specificity of any observed transfer effects. Gf was measured by a computerised matrix reasoning task.

Stroop task. The words “red”, “blue” and “yellow” appeared in either the congruent colour or the other two incongruent colours in the middle of the screen. Participants needed to press the corresponding button (“R”, “B” and “Y”) to the font colour and ignore the word meaning. Trial length was 1500 ms and there were 72 trials in total.

WM task. The WM paradigm was a change detection task, and identical to the paradigm developed in Chapter 2. An array of stimuli used in the SA task (except for the target) with random orientations was presented for 6 seconds during the encoding phase. After a 2-second blank screen, a single item from the array was shown again with or without a new rotation. The task was to judge whether a change in orientation happened. Possible degrees of rotations were 55,45,35,30,25,20,15,10,5 in either direction. The difficulty of the task started at the easiest level (rotation = 55 degree) and went up by one if 6 out of every 8 trials were correct and went down if less than 4 were correct; no change was made if 4 to 6 trials were correct. The task would stop if two consecutive no-change or drops occur, so too if participants have reached the highest level or failed the easiest level. The purpose of the staircase procedure was to mirror those used in standardised WM assessments such as the Automated Working Memory Assessment (AWMA; Alloway, 2007).

Forward and backward spatial span. Spatial span is often used as a measure of visuospatial STM/WM (Astle et al., 2015). A sequence of squares lights up one by one within a 5 x 5 grid pad, after the presentation, participants were required to click on the squares in the same order (forward) or the reversed order (backward span). The adaptive procedure and stop rule were the same as in digit span task and performance was indexed by the largest number of spatial locations successfully recalled.

Forward and backward digit span. These measures are among the most widely used tests of verbal simple and complex span (Jones & Macken, 2015). A sequence of numbers was presented and participants were asked to repeat it in the same order (forward span) or reversed order (backward span) by clicking on the number in a number pad provided on the screen. The task always started from 3 digits and increased by 1 each time a correct answer was given. The stop rule was 3 consecutive errors. Performance was scored by the highest length of sequence correctly recalled. Participants were instructed to not note down any of the sequences to aid performance but was not otherwise monitored.

Abstract reasoning task. Due to copyright issue we chose an open-access test designed to assess abstract reasoning. The test was made up of 80 items from the matrix reasoning item bank (MaRs-IB; Chierchia et al., 2019), composed by features varied along four dimensions: shape, colour, size and position in the matrix. Each item consists of an incomplete matrix containing abstract shapes. The administration followed the authors’ original protocol: participants complete the matrices by identifying relationships between the shapes and selecting the correct shape from 4 options. Participants were given maximum of 30 seconds for each trial and 8 minutes in total to complete as many as up to 80 trials.

Two sets of these items (80 each) were counterbalanced between the two time points in order to eliminate familiarity in a test-retest design. The two sets did not differ in accuracy or response time in study conducted by the original authors.

5.2.4 Data analysis

First, given our selection criteria that the groups should be equal on the SA task and differ on their performance of the WM task, we tested for pre-training differences on the transfer tasks using a series of one-way ANOVAs. For instance, the highWM training and control groups should begin with a better score on the WM task compared to the lowWM training group, whereas the SA scores should be equal. Next, the effects of training and baseline ability (i.e. the relationship between SA and WM) were studied separately for each task by testing the post-training differences while controlling for pre-training performance using ANCOVAs, which allowed us to take into account potential aptitude by treatment effects (Karbach et al., 2017). Post-hoc pairwise tests were run to follow up any significant effects within the original ANCOVAs. Any multiple comparisons were adjusted with FDR and Bonferroni methods.

5.3 Results

5.3.1 Pre-training group differences

A series of ANOVAs were used to test for any group differences at pre-training (Table 9). We found evidence for group difference on the WM, forward spatial span and abstract reasoning tasks. Post-hoc tests showed that the highWM training and the control groups performed significantly better than the lowWM group at this time point on all three measures. It is interesting that not only do they outperform the lowWM group on the WM task (consistent with our pre-screening requirement) but convincingly on the other two assessments also, which were not part of our pre-screening.

Chapter 5: Does training selective attention improve working memory?

Table 9. Results of one-way ANOVAs and pairwise comparisons for pre-training group differences on transfer tasks

	highWM+training		lowWM+training		Control		One-way ANOVA	Pairwise comparison		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		highWM vs. control	highWM vs. lowWM	lowWM vs. control
WM task	5.64	2.53	4.57	2.50	6.30	2.15	$F(2,80) = 3.52, p = .03^*$	$t(44.47) = -.97, p = .17$	$t(57.96) = 1.92, p = .04^*$	$t(44.33) = 2.95, p = .01^{**}$
Forward Spatial span	6.36	1.31	5.71	1.20	6.45	1.15	$F(2,80) = 3.15, p = .04^*$	$t(44.05) = -.26, p = .40$	$t(56.00) = 2.05, p = .03^*$	$t(41.29) = 2.30, p = .03^*$
Backward Spatial span	5.79	1.26	5.34	1.26	5.90	.79	$F(2,80) = 1.85, p = .16$	NA	NA	NA
Forward digit span	7.79	1.64	7.26	1.31	7.60	1.10	$F(2,80) = 1.62, p = .21$	NA	NA	NA
Backward digit span	6.86	1.99	6.46	1.62	6.90	1.77	$F(2,80) = 0.86, p = .43$	NA	NA	NA
Stroop task (RT difference)	142.49	56.01	144.12	75.51	180.61	88.61	$F(2,80) = 1.97, p = .15$	NA	NA	NA
Abstract reasoning task(accuracy)	0.76	0.12	0.63	0.14	0.76	0.12	$F(2,70) = 8.90, p < .001^{***}$	$t(39.28) = -.01, p = .50$	$t(52) = 3.69, p < .001^{***}$	$t(41.56) = 3.42, p < .001^{***}$

Note. P-values in multiple comparisons are FDR-corrected. Post-hoc tests were not performed for insignificant ANOVAs. $p < .05^*$, $p < .01^{**}$, $p < .001^{***}$

5.3.2 On-task training effects

For the SA training, two participants were excluded, both happened to be in the lowWM group. The first was due to an unknown technical error resulting in loss of the data of the last training session. The second was due to their reaction time throughout sessions being more than 3 SDs faster than the average across the rest of the sample. A mixed ANOVA showed significant effect of time ($F(2,118) = 52.71, p < .001$) but no effect of group ($F(1,59) = .39, p = .53$) or interaction ($F(2,118) = 1.82, p = .16$), suggesting that the magnitude of improvement was not modulated by groups (Figure 34). The highWM group improved significantly from the first ($M = 4.12, SD = 1.90$), to the last session ($M = 6.73, SD = 0.87; t(25) = 6.87, p_{\text{corrected}} < .001$). So did the lowWM group from first ($M = 4.29, SD = 1.64$) to the last session ($M = 6.09, SD = 1.67; t(34) = 5.83, p_{\text{corrected}} < .001$). Figure 36 shows the distribution of levels by group at the end of first and last sessions. Most trainees reached final level by the end of training.

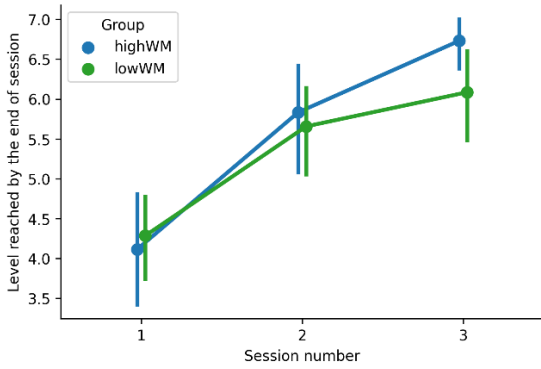


Figure 34. Training trajectories of the highWM and lowWM groups on the SA training task.

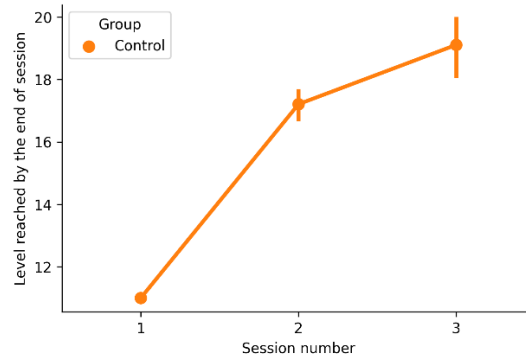


Figure 35. Training trajectories of the control group (highWM) on the control training task.

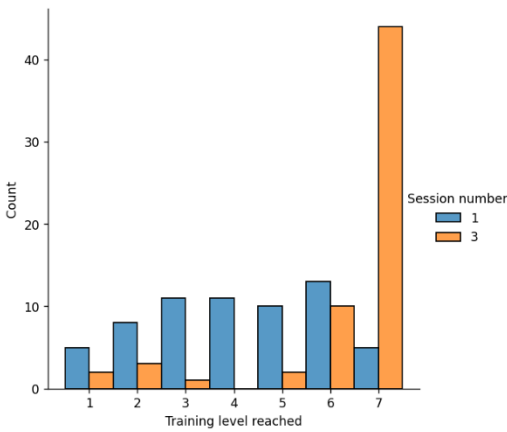


Figure 36. Distribution of level reached in SA training at the end of the first and last sessions

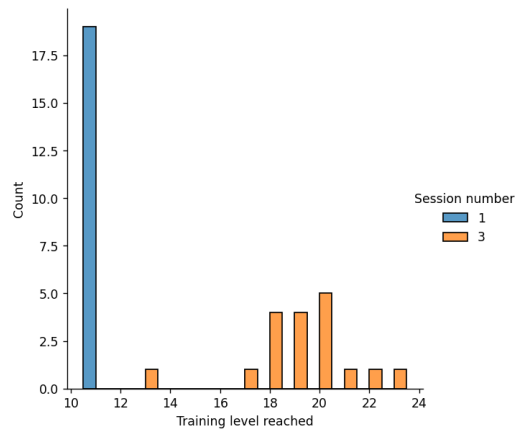


Figure 37. Difficulty level reached in control training at the end of the first and last sessions

The active control group also achieved significant progress on their training task from an average level of 11.0 at the end of the first session to 19.1 when they finished the training, $t(17) = 16.09$, $p < .001$ (Figure 35). The minimum degree of improvement was 9 levels from the starting level and the final distribution of level reached by the end of training suggested no ceiling effect (Figure 37).

5.3.3 Transfer effects

We reported results of ANCOVAs followed by post-hoc tests if justified. Details of summary statistics, test statistics as well as transfer patterns can be seen in Table 10 and Figure 38-44.

WM task. An ANCOVA was used to test for any group difference at post-training while accounting for pre-training difference. The effect of group was significant after covarying for pre-training performance ($F(2,79) = 7.16$, $p = .001$), while performance at pre also significantly and independently predicts post-training scores ($F(1,79) = 17.91$, $p < .001$). Post-hoc pairwise comparisons with pre-training adjusted means showed that the low

WM group achieved lower level compared to the highWM (mean difference = 1.72, $p_{\text{corrected}} = .006$) and control groups after training (mean difference = 1.96, $p_{\text{corrected}} = .006$). HighWM and control groups did not differ (mean difference = 0.23, $p_{\text{corrected}} = 1.00$)

Forward and backward spatial span task. For forward span, effect of group was not significant after controlling for pre-training performance ($F(2,79) = 0.58, p = .56$), whereas pre-training was predictive of post-training span ($F(1,79) = 17.15, p < .001$).

For backward span, the same pattern emerged: effect of group was not significant after controlling for pre-training performance ($F(2,79) = 1.74, p = .18$), whereas pre-training was predictive of post-training span ($F(1,79) = 13.61, p < .001$).

Forward and backward digit span. For forward digit span, effect of group was not significant after controlling for pre-training performance ($F(2,79) = 1.34, p = .37$), whereas pre-training was predictive of post-training span ($F(1,79) = 33.46, p < .001$).

For backward digit span, the pattern repeats: effect of group was not significant after controlling for pre-training performance ($F(2,79) = 0.39, p = .68$), whereas pre-training was predictive of post-training span ($F(1,79) = 37.40, p < .001$).

Stroop task. We used the averaged response time difference between correct congruent and correct incongruent trials as an index of performance. Again, there was no significant difference at post-training adjusted for baseline performance ($F(2,79) = 0.12, p = .89$), but a significant and independent contribution from pre-training performance ($F(1,79) = 11.90, p < .001$)

Abstract reasoning task. Due to technical error in the initial code, ten participants were tested with the same set of stimuli twice. Although the results did not vary with or without these individual's data, we reported findings regarding the abstract reasoning task excluding the said participants. There was no significant difference at post-training adjusted for baseline performance ($F(2,69) = 1.42, p = .24$), but a significant and independent contribution from pre-training performance ($F(1,69) = 69.45, p < .001$)

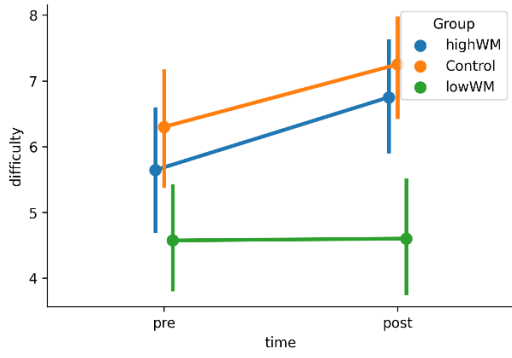


Figure 38. WM task performance before and after training by group. Error bars represent 95% confidence interval.

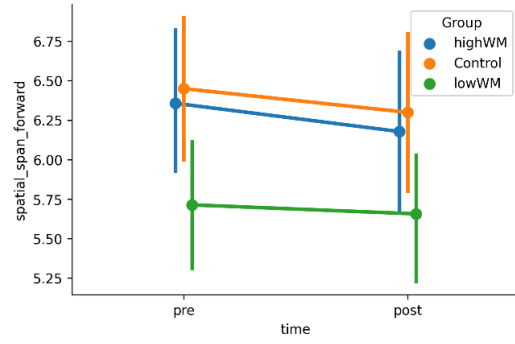


Figure 39. Forward spatial span.

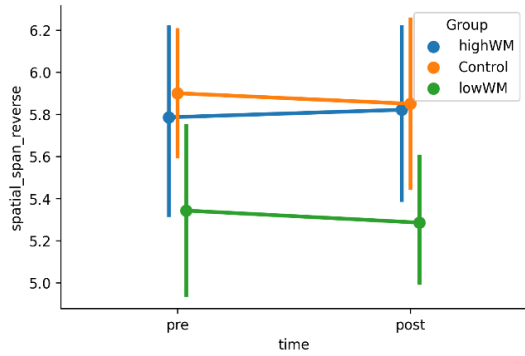


Figure 40. Backward spatial span

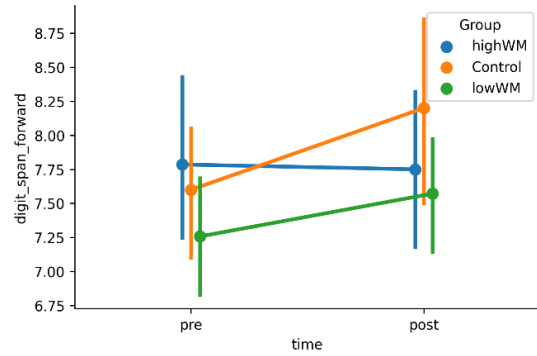


Figure 41. Forward digit span.

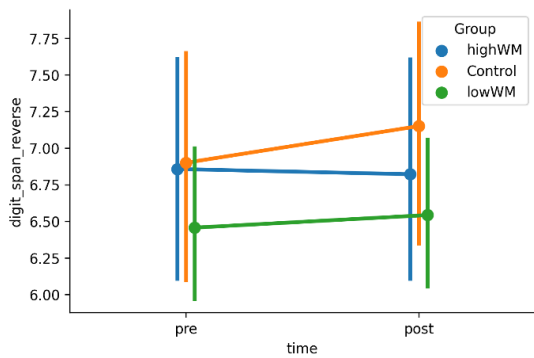


Figure 42. Backward digit span

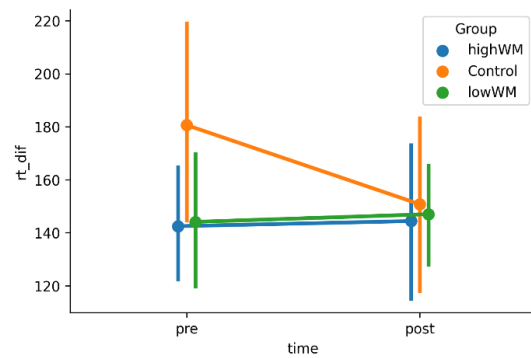


Figure 43. Stroop task (RT difference).

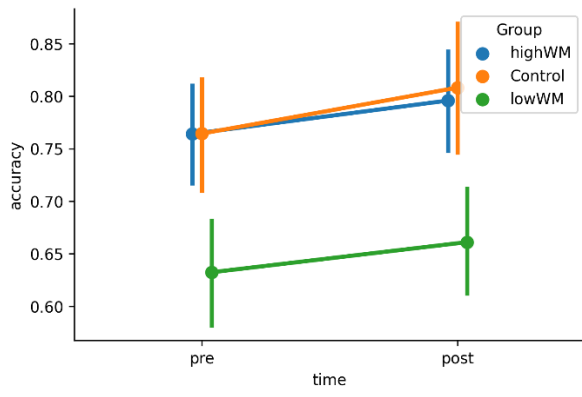


Figure 44. Abstract reasoning task (accuracy).

Table 10. Results of ANCOVA tests for post-training differences with baseline performance as covariate.

	HighWM+ training				lowWM+ training				Control				ANCOVA	Pairwise comparisons		
	Pre		Post		Pre		Post		Pre		Post			highWM vs. control	highWM vs. lowWM	lowWM vs. control
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>				
WM task	5.64	2.53	6.75	2.24	4.57	2.50	4.60	2.56	6.30	2.15	7.25	1.83	$F(2,79) = 7.16, p = .001^{***}$	mean difference = 0.23, $p = 1.00$	mean difference = 1.72, $p = .006^{**}$	(mean difference = 1.96, $p = .006^{**}$)
Forward Spatial span	6.36	1.31	6.18	1.36	5.71	1.20	5.66	1.19	6.45	1.15	6.30	1.22	$F(2,79) = 0.58, p = .56$	NA	NA	NA
Backward Spatial span	5.79	1.26	5.82	1.09	5.34	1.26	5.29	.93	5.90	.79	5.85	.93	$F(2,79) = 1.74, p = .18$	NA	NA	NA
Forward digit span	7.79	1.64	7.75	1.53	7.26	1.31	7.59	1.24	7.60	1.10	8.20	1.54	$F(2,79) = 1.34, p = .37$	NA	NA	NA
Backward digit span	6.86	1.99	6.82	2.04	6.46	1.62	6.54	1.58	6.90	1.77	7.15	1.76	$F(2,79) = 0.39, p = .68$	NA	NA	NA
Stroop task (RT difference)	142.49	56.01	144.42	80.85	144.12	75.51	147.01	56.47	180.61	88.61	150.79	77.44	$F(2,79) = 0.12, p = .89$	NA	NA	NA
Abstract reasoning task (accuracy)	0.76	0.12	0.80	0.13	0.63	0.14	0.66	0.14	0.76	0.12	0.81	0.14	$F(2,69) = 1.42, p = .24$	NA	NA	NA

Note. P-values in multiple comparisons were Bonferroni adjusted. Mean difference were calculated from post-training scores adjusted for baseline performance. $p < .05^*$, $p < .01^{**}$, $p < .001^{***}$

5.4 Discussion

This study used repeated SA training to explore the effects on WM performance, as well as the moderating role of individual baseline abilities. In particular we tested whether the relationship between SA and WM prior to the training would moderate the degree of transfer. We found transfer to the WM task in the highWM and control groups but not in the lowWM group. One interpretation of this could be that where participants are good at both SA and WM, there is a more straightforward transfer between them. This would explain why the relative WM gains are greater for those with good SA and WM before the training began. However, the spanner in the works for this interpretation is the control group. The active controls showed comparable gains on the WM task. This makes the interpretation of the findings more challenging. There are several possible explanations, or a combination therein, to explain the transfer patterns.

A first possibility that we can dismiss is that the lowWM simply did not train properly. The on-task training gains clearly show that they made comparable gains to their highWM counterparts on the SA training. The next possibility is that given both the highWM and the control groups showed transfer gains after their separate training regimes, that these improvements do not reflect training effects at all, but test-retest effects. Though this would not explain why the lowWM group do not show these effects. In fact, given that these participants were selected as the lower-end of the WM distribution, you might expect to see regression to the mean (Smoleń et al., 2018) which would inflate their apparent training improvements (rather than diminish them). Another reason why ‘test-retest’ seems implausible is because our earlier use of this task has shown that it is pretty stable over time.

This still leaves us with a group difference in transfer gains to account for. The crucial finding seems to be that pre-training WM differences strongly moderate the transfer gains – participants with good WM skills before training show better cross-task transfer from their training, relative to participants with poorer WM skills (regardless of the type of training they do). This is in line with our previous findings which showed that the degree of positive transfer was predicted by participants’ baseline fluid intelligence scores (Zhang et al., 2020). In the current study it is possible that we are observing a more primitive manifestation of the magnification effect, with the largest gains shown by the most cognitively efficient to start with (Guye et al., 2017; Smoleń et al., 2018). Indeed, we also observed a baseline difference in the nonverbal reasoning task between the high and low WM individuals. The complication in the current case is that we see this effect for both the SA and control training groups, implying that this advantage of pre-training WM capacity can manifest regardless of the kind of training, or at the least that any kind of generically taxing training will reveal it. There are a few reasons why this seems unlikely. The primary reason is that this is at odds with the rest of the training literature. Most transfer effects happen via strengthening of highly overlapping

task-specific components (Morrison & Chein, 2011). Recent studies have demonstrated feature- and paradigm-specificity of transfer, showing that transfer can be highly specific to the stimuli or paradigms (Holmes et al., 2019; Minear et al., 2016). So the idea that we have managed to induce some transfer from any generically taxing training seems unlikely, moreover, it would not explain the lack of transfer to the other tasks in the battery.

The next alternative to consider is that we are observing two distinct training effects for each of the groups. With some enhanced attentional process through SA practice in the highWM participants driving the improvements on the WM task, whereas those same improvements do not yield equivalent benefits for the lowWM group because they do not draw upon this process to the same extent to support WM. This is consistent with previous work showing transfer to various cognitive abilities from attention training (Greenwood & Parasuraman, 2016; Greig et al., 2007; Haut et al., 2010; López-Luengo & Vázquez, 2003; Mishra et al., 2014; Wass et al., 2012). Similar claims about individual differences in baseline levels reflecting available cognitive resources has been tested and supported (Kliegl et al., 1990; Verhaeghen & Kliegl, 2000), although this was not specific to the idea of task overlap. Foster and colleagues (2017) ran a study to address the question whether gains from WM training differ between low and high span individuals. They chose a visual search task for the active control group and reported some transfer to complex span tasks in high span relative to low span participants, suggesting highWM individuals may translate SA gains to the related WM domain. Considered together with our previous neuroimaging findings, demonstrating common neural substrates in the group with both good SA and WM performance, it is possible that for these individuals there is enhanced coupling and thus the strengthening of one task is more readily shared with the other. Conversely, even though the lowWM group were equivalent to the highWM counterparts on initial SA performance and showed the same magnitude of training-related SA gains, the lesser degree of task overlap constrains the transfer of training gains.

Of course, for this to make any sense, we would have to provide some different account for the improvement shown by the active control group. In this case the perceptual sensitivity training may have altered other underlying processes and induced changes through some perceptual learning process. Stimulus discrimination can become faster and more accurate with repetition. Some have claimed that perceptual training in either visual or auditory domain can enhance other systems such as WM, SA and fluid intelligence by sharpening the ability to ignore distraction (Anderson & Kraus, 2013; Berry et al., 2010; Greenwood & Parasuraman, 2016; Strenziok et al., 2014), accompanied by decreased activity in sensory processing areas and changed connectivity between the dorsal attention network and these sensory areas (Baldassarre et al., 2012; Kelly & Garavan, 2005; Lewis et al., 2009; Strenziok et al., 2014). One research group used a pitch discrimination paradigm to train target-distractor identification and found in both rats

and humans that multimodal neural responses to distractors decreased following training (Mishra et al., 2014). Moreover, the reduction in an ERP component for distractor processing was correlated transfer an auditory WM span task. Although this neural evidence should be taken with caution. Across the literature it is never clear whether training should induce increased or decreased activity or connectivity, and indeed all have been observed (Baldassarre et al., 2012; Chein & Schneider, 2005; Lewis et al., 2009; Strenziok et al., 2014). Coming back to the current set of findings, it is therefore possible that the highWM and control groups showed better transfer because of two distinct sets of mechanisms – one based on training attention, and one based on training some supra-modal process. However, there are a few reasons why this ‘different mechanisms’ account is on somewhat shaky ground. The most obvious of which is: if the control training is really inducing some supra-modal enhancement why does this not show much transfer elsewhere?

One final possibility is that these improvements are not training gains at all. Individuals in the highWM training and control groups may have had an edge over the lowWM counterparts in mastering the WM task, in terms of adapting to the specific configuration of the stimuli sequence, devising relevant memory strategies and so on. The difference in pre-test WM performance could have reflected this. In other words, it may be the case that the improvements were not outcomes of an enhanced domain-general mechanism, but rather people with higher WM capacity (and higher Gf scores) were more efficient at learning task-specific strategies and were able to apply them at the second time of assessment. This is compatible with studies showing that high WM capacity is strongly associated with learning in associative and category-learning tasks (Lövdén et al., 2012; Tamez et al., 2012), which particularly seem to reflect strategy-based as opposed to process-based (e.g. processing speed) changes (e.g. Karbach et al., 2017; Karbach & Verhaeghen, 2014).

Another interesting finding of the present study is the apparent baseline difference in nonverbal reasoning test between the highWM and lowWM individuals. We did not expect this distinction as other tasks we have included which measure more closely related constructs than Gf did not yield the same pattern, with the exception of the forward spatial span. What could be driving the relationship between individual difference in the WM task and that in the Gf test? An explanation could lie in the different types of WM measures that draw on related but not identical processes. Psychometric studies of WM have long made the distinction between WM tasks that require storage and processing, using complex span tasks that have a secondary processing component (e.g. arithmetic operation, reading, etc), and tasks that require passive storage of information (e.g. serial recall). The former has been proposed to rely more on the “executive attention” aspect (or the “central executive” in the language of the multicomponent model) above and beyond passive storage, which refers to the capability of maintaining memory representations in a highly active state in the presence of interference (Engle, Kane, et al., 1999; Kane et al., 2001). WM tasks that

contain this component load on a single latent factor, relative to those that do not and instead load on a separate storage capacity factor, though the correlation between the two factors is usually strong (e.g. around 0.8 in Conway et al., 2002; Kane et al., 2004). Importantly, the ‘executive attention’ latent variable more strongly and uniquely predicts Gf compared to storage capacity, suggesting that the former component may be a critical source of variability in general intelligence ability. Our WM task (change detection), which has been shown to reflect the same underlying structure to complex span (Shipstead et al., 2014), may thus be better positioned to detect individual difference in Gf (Fukuda et al., 2010), relative to the simple serial recall tasks (Cantor et al., 1991).

The present study could benefit from several design modifications to better explain the equivocal findings. Firstly, we noted that majority of participants of the SA training reached the highest level before the end of the program, suggesting a ceiling effect may have occurred. Consequently, the lack of difference between the highWM training group and the controls could arise because SA training was not sufficiently challenging to sharpen the targeted function. Considering that in Chapter 2 further increasing the distractor-distractor difference did not increase the response time beyond the currently adopted range (which makes sense as orientations are circular), an alternative is to decrease the time allowed, forcing trainees to become more efficient to continue progressing. Secondly, the control training may have taxed some general processes through perceptual learning or supra-modal attentional mechanism, and as outlined above this makes the results hard to interpret. Future studies could choose tasks that draw on something very different, like learning factual knowledge, or use a low-intensity version of the SA training.

Chapter 6: General discussion

Studying the interplay between selective attention (SA) and working memory (WM) has a long history and continues to be a focus of experimental cognitive psychology and neuroscience. For at least three decades behavioural studies have broadly shown that these two functions share some common processes. This evidence has come from a mixture of dual-task designs, individual differences studies and cueing paradigms. Over time, neurophysiological and neuroimaging evidence has added to these findings, demonstrating that the neural substrates of SA and WM overlap spatially and that the neural mechanisms underpinning these constructs are similar. Some of the key demonstrations that speak to a shared mechanistic framework include finding that the representation of information in sensory processing regions is highly similar between the perceptual and mnemonic domains (Harrison & Tong, 2009; Serences et al., 2009) and that the top-down attentional signals originating from prefrontal and parietal control areas can influence these representations in a comparable fashion (Gazzaley & Nobre, 2012). However, across that long and rich history of empirical research one core assumption has been that the relationship between the constructs is stable and constant across individuals. In short, the literature has largely explored whether SA and WM share mechanisms (or not) at a group level, with potential individual differences in the relationship between SA and WM remaining largely unexplored. Mechanistic heterogeneity becomes more likely as tasks become more complex: participants may achieve the same task performance through different means. This thesis adds to our knowledge of the SA-WM interplay, by bringing an individual differences perspective. We employed a multifaceted approach, combining behavioural, neural and computational methodologies to address three key research questions:

1. Is there variability in how strongly SA and WM performances are coupled, when low-level processes are carefully matched?
2. What are the neural correlates associated with individual variability in SA-WM overlap, and what is the functional role of the overlap?
3. Do individual differences in the processing overlap moderate the impact of SA training on WM performance?

The implications of our results are discussed in detail in the following sections, with particular reference to these three core questions. To briefly summarise: in the first empirical chapter we tested a large online sample with two carefully matched behavioural paradigms for measuring SA and WM and examined their relationships with other well-established measures. Subgroups with profiles that may suggest differentially coupled processes between SA and WM were identified. Behavioural task loadings were not equivalent across these groups, implying that tasks can load onto constructs variably across participants. The second empirical chapter examined the neural circuits underlying both cognitive functions and highlighted a

network encompassing frontal, parietal and occipital regions. The critical finding was that one area of the right superior parietal lobule, thought to be involved in endogenous control of attention, exhibited differential coupling between the tasks depending on participants' behavioural profiles. The third chapter used recurrent neural networks (RNNs) to explore the functional overlap between SA and WM within a computational model. Neural patterns within the recurrent network were compared to those from human fMRI data. The RNN developed neurons serving subtly differential functions. Some neurons were more crucial for WM performance, while others were broadly shared between both tasks. Furthermore, the shared neurons mirror activity patterns of a frontoparietal network highly similar to that revealed by the fMRI study. The last chapter adopted a cognitive training approach to demonstrate that the casual effect of training SA on WM was modulated by the participants' behavioural profile at baseline. The degree of process overlap may partially determine transfer from SA to WM, although more work is required to understand the nature of the transfer mechanisms.

6.1 Relationship between SA and WM as assessed by our new tasks and individual differences

We designed the new paradigms for measuring SA and WM. One key requirement was to match the visual aspects of the tasks as much as possible, such that extraneous task differences (modality, stimuli etc.) were controlled for. From the large behavioural sample (N = 664), there was very little support for the overall association between the SA and WM measures (the correlation was significant but tiny and negative). In other words, the ability to select a target efficiently among distractors was not predictive of the ability to memorizing items over a brief period of time. This result is somewhat unexpected. But a closer inspection of the behavioural evidence regarding the relationship between SA and WM, and particularly when the former was assessed by visual search, shows that actually the links can be inconsistent. Studies using a dual-task setup often observed conflicting patterns, with some finding interference effect of a secondary visual search task at the cost of WM performance, and others not (Oh & Kim, 2004; Woodman et al., 2001). These studies used a wide range of tasks, stimuli (e.g. letters, colours, shapes) and search type (pop-out vs. conjunction) for the SA task. There is also lots of variability in the WM task (verbal, spatial or visual), adding to the challenge of integrating findings to understand the boundary conditions for SA-WM interactions. Like the dual-task methodology, individual differences studies which typically compare two groups of participants – those with high versus low WMC – have yielded mixed results. In the verbal domain, WMC correlates with the detrimental effect of unexpected, infrequent auditory distractors, but not with continuous and predictable ones (Hughes et al., 2013; Sörqvist, 2010). Whereas in the visual domain, although some researchers have shown that low WMC individuals constantly encode irrelevant information, compared to their high capacity counterparts using neurophysiological measures (McNab & Klingberg,

2008; Vogel & Machizawa, 2004), others have failed to detect such effects behaviourally over a series of visual search tasks (Kane et al., 2006; Poole & Kane, 2009; Sobel et al., 2007). Again, one apparent explanation for the lack of consistent evidence is the task-specific variations across studies. Even within a study, WM is usually assessed by one or more complex span, n-back or change detection tasks, employing drastically different presentation and modality of stimuli from those used in the SA task. Across independent studies it is also difficult to find a subset that use comparable paradigms, with parametrically varied features, that allows for a systematic comparison of the results. Added into this messy mix is interpersonal variability – it could simply be that not all participants recruit shared processes, whereas others do. We tried to address in the current thesis.

Our large online study allowed us to produce normative distributions of performance on each of the new tasks. We could then sample from these distributions systematically, creating subgroups with different profiles of performance on the SA and WM tasks. We then tested whether these groups reflected differences in the degree of shared processes between the functions. In other words, for those whose SA and WM abilities are underpinned by more overlapping mechanisms, we would expect their relative performances for each task to come from a similar point in the respective normative distributions. By contrast, those who are using different processes, their relative performances can come from different points on the respective distributions. We saw some preliminary evidence in Chapter 2, with the relationships between the measures being inconsistent across the subgroups, relative to the broader sample. Specifically, in two of the subgroups (the highSA+highWM and to a lesser degree the highSA +lowWM) the extent to which the corresponding latent factors captured variance in the SA and WM measures differed compared to the baseline sample. One possibility is that these individuals relied upon different sets of processes to perform the same tasks – specifically those who are relatively good at both may be drawing on more shared latent processes. The notion that participants can recruit different processing components to perform a task is by no means a novel idea. Developmental research exploring changes with age on the organization of cognitive functions has repeatedly demonstrated that relationships among cognitive abilities are not constant, but rather become more dissociated from childhood to adulthood and later back to more correlated during older adulthood (Balinsky, 1941; Garrett, 1946). A number of studies have examined the patterns of correlations across a number of psychometric tests and found age-related changes. For example, Gajewski et al. (2018) asked how different functions contributed to n-back task performance in young, middle-aged and old adults and assessed their performance with multiple measures of attention, processing speed, WM and crystallized intelligence. Younger individuals relied mainly on executive functions including interference control to perform the 2-back task, and the middle-aged individuals recruited mainly short-term memory and processing speed, whereas in older participants' performance was distinctively associated with attentional functions. Task switching and updating were associated with n-back performance across ages. Although

the tests used in this study likely capture a complex set of processes, it is a nice demonstration that individuals may draw on the same or different processing components to a greater or lesser degree to execute a complex behaviour. This shift could be the result of age-related changes, as in the example, or presumably due to individual differences, as in our case.

The behavioural chapter addressed the possibility that there were heterogeneous cognitive profiles of SA and WM, underpinned by differential relationships between the measurements. The next chapter took this a step further, using neuroimaging to investigate whether subgroups differed in the extent of functional overlap. Participants with two of the profiles of interest (i.e. one with the coupled SA and WM, and the other with the dissociated SA and WM) were pre-screened against the large normative sample and scanned again with fMRI. The scan also enabled us to confirm that performance differences were relatively stable, rather than a less exciting fluctuation in measurement error. Our neuroimaging results showed that one area among the overlapping neural substrates activated by both tasks, the right superior parietal lobule (rSPL), reflected group differences in the strength of functional coupling between the SA and WM processes. That is, the degree to which the same voxels of rSPL tracked the increasing demands of SA and WM in the coupled processes group was higher compared to that in the dissociated process group. Therefore, we have potentially identified a brain correlate of inter-subject variability in shared processes between these cognitive functions. SPL has been extensively indicated as part of the overlapping activations for attentional and mnemonic representations (Corbetta & Shulman, 2002; Fusser et al., 2011; Labar et al., 1999; Lepsien & Nobre, 2006; Mayer et al., 2007; Nee & Jonides, 2009; Pollmann & von Cramon, 2000). It is also robustly engaged along with frontal eye field (FEF) and the neighbouring intraparietal sulcus (IPS) as the dorsal attention network, responsible for externally directed attention (Brissenden et al., 2016; Corbetta & Shulman, 2002; Kastner et al., 1999). Interestingly, a previous study comparing neural activity associated with selection within perceptual and mnemonic domains showed similar effects in the FEF and SPL in both cases (Nee & Jonides, 2009). Neural activity in these regions was correlated with behavioural measures of SA and WM selection difficulty, respectively. Furthermore, the authors noted individual differences in the rSPL, with some participants exhibited strong commonalities, using SA-related SPL resources to perform memory selection, whereas some did not recruit this area as much, with correspondingly worse behavioural performance relative to their counterparts. The participants were further divided based on rSPL activation intensity in the memory selection condition and illustrated that the group who relied more on this area for memory also recruited the dorsal attention network (FEF and right premotor cortex) to a greater extent. Conversely, those who relied less on rSPL in the memory condition showed greater activation in the left ventrolateral prefrontal cortex (VLPFC), a region more uniquely associated to memory selection. In sum, some individuals showed a similar pattern of activation between SA and WM related search processes,

while the others had more distinct patterns between the domains, consistent with our findings that participants with different behavioural profiles varied in the degree of functional coupling in this region.

Having shown that the neural instantiation of functional coupling between SA and WM, we used a training regime to probe the causal relationship between these functions and whether individual difference modulates it. The preliminary results suggested that training SA can lead to improvement on the perceptually matched WM task, however, these effects were only seen in participants with high baseline WM performance. What mechanism could potentially mediate these gains and why would it differ between baseline behavioural profiles? One possible explanation lies in the training-induced neural changes following repeated practice on attentional control. Mishra et al (2014) adopted an adaptive training program in which a target auditory tone needed to be identified among distractor tones with increasing similarity upon successful responses. Later they found a reduction in distractor-evoked ERP and improved WM span in the training group, and the extent of reduction was negatively correlated with WM benefits. In a related study, the same group measured the N1 component which was sensitive to distraction suppression (Berry et al., 2010; Gazzaley, Cooney, Rissman, et al., 2005a) before and after a visual perception training with adaptive levels of distractors and found similar pattern of results to the previous study (Mishra et al., 2015). That is, N1 amplitude was higher after training and the degree of change was correlated to improvement on the training task as well as on an untrained WM task. These support the notion that improved distraction suppression is accompanied by changes in neuronal processing and can affect related functions such as WM. Moreover, a number of studies and reviews have identified regions of the dorsal attention network (DAN), including SPL as locations of altered neural processing (Chein & Schneider, 2005; Greenwood & Parasuraman, 2016) as a results of many different types of cognitive training. For instance, n-back task training led to increased effective connectivity in the superior parietal cortex (Kundu et al., 2013). Therefore, training-related change in regions associated with attentional control may be one mechanism mediating the transfer effects observed in our study. Importantly, this could explain why such benefit occurred in the highWM group but not the lowWM counterparts: In the former, the neural mechanism supported by the rSPL was enhanced through repeated training and this enhancement was deployed in the untrained task as it shares the same neural mechanism. If individuals already rely on similar neural substrates to perform SA and WM-related tasks, strengthening of the network can be more readily transferred to other tasks that activate the same neural architecture. We are fully aware that this interpretation is limited by several confounding factors in the results (i.e. the control training with highWM individuals also showed transfers) and future work should further dissociate the influences before a strong conclusion can be drawn.

6.2 Common neural correlates for SA and WM and their functional roles

There were some consistent patterns of overlapping activity across the fMRI and computational modelling chapters. Across these studies, we adopted a conventional conjunction approach, a group-constrained single subject analysis (GCSS), as well as a method for identifying brain regions mostly associated with a cluster of neural network units that subserved common computations between SA and WM. Most noticeably, all analyses repeatedly highlighted the inferior temporal gyrus (ITG), SPL (Brodmann area 7; BA 7), intraparietal sulcus (IPS; BA 40) and premotor cortex (BA 6) containing the FEF. As mentioned in discussions in the previous chapters, these regions comprise the dorsal attention network (DAN). This network is associated with externally directed covert and overt shifts of spatial attention, eye movements, and hand-eye coordination (Corbetta & Shulman, 2002). The nodes of the network show robust structural and functional connectivity during rest (Fox et al., 2005; Szczepanski et al., 2013; Vincent et al., 2006). They contain areas with retinotopically organized maps of contralateral space (Silver & Kastner, 2009), which makes them good candidate regions for the maintenance of spatial priority maps supporting covert spatial attention, saccade planning, and visual working memory (Jerde et al., 2012). Moreover, task-based connectivity studies and brain stimulation have demonstrated the role of the DAN in generating top-down biasing signals which modulate activity in the visual cortex (Ruff et al., 2009). Taken together with its consistent activation within WM tasks across different modalities (Konoike et al., 2015; Majerus et al., 2012, 2018), and its overlapping activity in studies comparing SA and WM (Fusser et al., 2011; B.-C. Kuo et al., 2009; Labar et al., 1999; Mayer et al., 2007; Pollmann & von Cramon, 2000), these regions may represent a distributed network underlying processes associated with goal-directed control of attention more generally. The fact that these distinct regions were not only indicated in the conventional group-average conjunction but also in the GCSS analysis with 80% threshold (i.e. more than 80% of all participants had conjunctive activity in these areas), suggests that this network was a stable neural substrate for the SA and WM overlap across individuals. Our computational modelling results echoed this. These regions subserve shared neural computational mechanisms between SA and WM, as opposed to reflecting distinct processes with anatomical overlap. Representation patterns within the DAN were related to neurons in the RNN which under perturbation, affected performance on both tasks equally.

Lateral PFC and insula were also highlighted by the group-level conjunction and the RNN analysis, suggesting that they were also plausible candidates for overlapping process between the SA and WM. However, co-activation in these areas showed higher degree of inter-subject variability, evidenced by the lack of prefrontal and insular indication in the GCSS results. In particular, the lateral prefrontal region revealed by the conjunction analysis, which sits in BA 44, was posterior and contralateral to that revealed by the RNN cluster, which covers parts of BA 10/9 and is more anterior within the PFC. Although the exact

cause for this difference is unclear, as drawing a direct comparison between processing in human neural activity and a computational model remains challenging. But this serves as an interesting contrast to the DAN areas which were far more consistent throughout our analyses. This distinction between lateral PFC and posterior regions is in line with previous studies investigating neural substrates reflecting neural resources shared by visual search and WM encoding (Fusser et al., 2011; Mayer et al., 2007). In their combined paradigms, the authors asked participants to search for targets in order to encode the objects or locations containing them. The attentional and memory demands were manipulated orthogonally by adjusting the similarity of the targets to distractors and the number of targets, respectively, thus enabling them to identify brain regions that would show increased activity alongside each type of demand, as well as those that would show an interaction effect. They found canonical prefrontal, premotor, posterior parietal and visual processing activity, which was sensitive to both SA and WM demands, yet only FEF, SPL (IPS) exhibited interaction effects. This implies that activity reaches a plateau under high attentional and high encoding demands, and that in these areas competition for shared neural resources between SA and WM encoding takes place. This would concur with our findings that SPL activity reflects individual differences in the degree of overlapping process between these cognitive functions. Conversely, prefrontal regions being clearly sensitive to attentional and memory demands but not showing the same response pattern points to a differential role of PFC from the posterior areas. For instance, PFC activation has been linked to a variety of control processes in the context of WM (Miller & Cohen, 2001), such as strategic or organisational processing required in complex stimulus encoding (Bor et al., 2003), and endogenously generating top-down biasing signals directing goal-related processing (Gazzaley et al., 2007; Gazzaley & Nobre, 2012; Zanto et al., 2011). Further support for the role of PFC in flexibly representing information within the context of higher-level task-related processes comes from MVPA. Stimulus category (e.g. direction vs. speed of dot motions; Riggall & Postle, 2012), and task rules (e.g. perceptual vs. category-membership judgment; Lee et al., 2013) can be both be decoded from lateral PFC, depending upon what is relevant at a given moment. Overall, lateral PFC in our work potentially suggests a functionally distinct role from the posterior DAN areas, which could subserve more abstract and higher-order control processes that are not stimulus-specific but nevertheless integral to SA and WM.

Lastly, the function of anterior insula (BA 13) has labelled as a “core” node in a separate cingulo-opercular cognitive control network that typically consists of anterior cingulate gyrus, anterior insula and the adjacent frontal operculum (Dosenbach et al., 2007). This network is dissociable from the frontoparietal network (containing lateral PFC and IPS) by intrinsic connectivity patterns (Yeo et al., 2011) and functionally is proposed to implement and update basic task sets on a slower time scale, relative to the frontoparietal network which initiates and adapts control on a moment-to-moment basis (Dosenbach et al., 2007). Sadaghiani & D’Esposito (2015) provided supporting evidence to this claim that this area offers cognitive

control over a longer timescale, showing that manipulation in the amount of alertness (i.e. the effortful and self-initiated preparedness to process and to respond) selectively activated the cingulo-opercular network, independent of manipulation in the SA demands which selectively activated the DAN. Alternatively, other investigators have observed that regions in this network only activate when participants make a discrimination response to a stimulus gradually revealed in noise (Ploran et al., 2007), or after top-down modulation from the frontoparietal network had occurred in a WM retro-cue set up (Wallis et al., 2015). According to these studies, the cingulo-opercular network plays a more downstream role, acting on evidence integrated by the other control networks, in interaction with sensory processing areas (Wallis et al., 2015).

To sum up, performing SA and WM tasks coactivated regions of several networks involved in related but separate control processes. Specifically, the rSPL was identified as a neural correlate for individual variability in shared processes between these cognitive functions, possibly reflecting the degree to which participants relied on the same attentional mechanism for both perceptual selection and WM encoding. One particular area we have refrained from discussion is the occipital cortex, which was repeatedly and robustly indicated in our analyses (with the exception of the GCSS correlation difference). Visual areas have been linked to processing low-level stimulus information as well as holding the stimulus representation (Kuo et al., 2016; Serences et al., 2009); activity within the occipital region also scale with SA and WM demands (Anderson et al., 2007). However, as the fMRI study did not have a control condition in which sensorimotor aspects of the tasks were matched, enabling subtraction of related signals from the SA and WM tasks, we could not distinguish co-activation in the visual area due to low-level sensory processing, or that due to other mechanisms.

6.3 Limitations

There are a number of limitations to the research in the thesis that should be addressed in future work. Firstly, the main challenge was trying to translate experimental tasks into individual differences measures (e.g. Barch et al., 2008; Boy & Sumner, 2014). We designed our tasks based on well-established paradigms that produce robust cognitive effects in SA and WM in the literature, such as the difference in rapid pop-out search and difficult conjunction search (Treisman & Gelade, 1980). These measures aim to precisely characterize a cognitive mechanism based on the typical or average response to a range of experimental variables. Within this context low between-subject variance is advantageous, because it leads to a smaller standard error of the mean, and thus more significant between-condition statistics (Hedge et al., 2017). In contrast, we wanted to capture individual differences – our correlational approach relied upon it. This could be potentially problematic for our purpose of selecting groups with certain performance profiles, as these measures may not be fully capable of consistently ranking individuals. Moreover, it may also partially

explain the lack of associations across our tasks, including our so-called validation tasks, as “the ability to detect relationships with other constructs will be compromised by the inability to effectively distinguish between individuals on that dimension” (Spearman, 1910). We have attempted to account for this difficulty by using mixed linear modelling which produces more reliable estimates of individual differences than aggregated statistics (e.g. means and standard deviations; Haines et al., 2020). However, if possible, future work could benefit from redesigning perceptually matching tasks according to other established tests, better designed for correlational studies.

Secondly, the sample sizes of the fMRI and cognitive training studies are on the smaller side. A growing body of research has suggested that task-based neuroimaging studies suffer from lack of statistical power due to insufficient sample sizes, which consequently inflate the estimation of the effect size. We adhered to previous recommendations of a minimum $N = 20$ per group (Thirion et al., 2007), but subsequent studies have shown that with standard general linear models the averaged replicability of patterns of clusters and peaks across different cognitive tasks were below 50%, even with sample size of 36, and remained below perfect even with a large sample (e.g. $N > 100$; Turner et al., 2018). What this means for us is that it is relatively unlikely to remedy the issue by recruiting even higher number of participants, at least within the scope of a single study, due to constraints of funding, global pandemics, time and rarity of the cognitive profiles within the population. Alternatively, more studies with moderate samples focusing on similar questions could be pulled together to draw inferences with a meta-analysis. Furthermore, results from the present work could inform a prospective power calculation, aiding the design of future projects. Likewise, the same applies to the cognitive training study. Although the power of this could be boosted more easily because of access to large samples online.

Thirdly, we defined groups with “coupled” or “dissociated” processes as those whose performance level on the SA and WM tasks were relatively matching – i.e. taken from roughly the same or different points on the respective normative distributions for each task. There are limitations associated with this approach. First, as a one-time assessment, extreme performance on a task owing at least partly to measurement noise could result in erroneous profiles, and potentially regression to the mean in subsequent testing. Although this was partially mitigated by testing individuals twice (once in the pre-screening and once in the fMRI and training study) and showing that group differences continued to exist, we did not check whether individual participants’ level of performance replicated (except for the small sample for test-retest reliability). Second, it was assumed that individuals performing well on one task and poorly on the other had less mechanistic overlap between the cognitive functions compared to individuals performing equally well or poorly on both tasks. Whilst it may be true to an extent, there may be multiple routes to the same profiles: for example, participants with “dissociated” processes could achieve high scores on both tasks

because their SA- and WM-related processes function at a high level independently. Behaviourally, we have no way to distinguish these kinds of individual from those with a truly overlapping process. Moreover, the overlap between SA and WM may not be matter of kind but a matter of degree. Namely, assigning participants to one of two groups could have simplified the story. A plausible solution that can be explored is to use two sets of tasks each tap into a single cognitive construct. The relationship (e.g. covariance) between the constructs on a latent level within each individual could then be taken as a better measure of the degree to which processes are shared.

Lastly, as discussed in the RNN chapter, there were some inconsistencies in how we abstracted the fMRI tasks in order to create inputs for training the model. Specifically, the number of stimuli simulated in the RNN was lower than that used in the actual tasks, and the stimulus features were reduced to two dimensions: the frame colour to manipulate WM load, and the orientation to change perceptual similarity among items. However, in the fMRI SA task, search target was defined by its shape instead of its orientation. We conducted our comparison analysis between RNN and brain patterns on the basis that both produced similar behavioural results, but it is less clear how the discrepancy in simulation choice affected the processes inside the neural network. On the other hand, our work with the model is expandable in many directions. For example, we can design a new component for the RNN specialised in image processing to emulate functions of the visual cortex, thus separating related mechanisms from those of our interests (i.e. attentional modulation, encoding and maintenance). We could also control model architectures to mirror the connectivity structure of the brain, for instance, long range sparse connectivity between different key hubs of the frontoparietal network and dense connections within each hub, which might allow us to study the nature of computations carried out by different brain regions.

6.4 Future directions

Research to date has primarily focused on group-level approaches to understanding the functional relationship between SA and WM. This thesis expanded this view by investigating individual variation in the shared processes, with multiple methodological tools. In short, we have attempted to identify behavioural profiles that were potentially associated with differentially coupled processes between SA and WM and examined the neural and cognitive consequences of these differences. These results provide a valuable first step towards an individual differences perspective in studying relationships between complex cognitive functions.

As the next step it would be particularly useful to confirm the relationship between behavioural profiles and degree of overlap between cognitive functions. As discussed in the limitations, our subgrouping method was based on assumption that individuals with more shared processes would also perform both tasks

similarly. This is not without its issues. Ideally, an improved approach would allow a more direct mapping between behavioural testing and functional overlaps. For example, it is plausible that instead of two tasks, two sets of tasks each assessing a single cognitive construct could be used. The relationship (e.g. covariance) between the constructs on a latent level within each individual could then be taken as a better measure of the degree to which processes are shared.

Going forward, one interesting question not fully addressed in this thesis is the nature of the shared neural mechanisms between SA and WM. We speculated based on the region's location within the DAN that its functional role is related to process of shifting attention, however, it is not possible to exclude other possibilities with the current design, such as other general top-down control signals sensitive to increased task demand. It will be important to use different experimental designs that dissociate these different functional roles. Alternatively, computational modelling could provide useful insight, by further distinguishing clusters of artificial neurons based on their activity profiles. For instance, we only divided our network into two functional clusters, whereas it is certain that there could be further differentiations within each cluster, as evidenced by the variation in the task selectivity metric. By using these more functionally specific clusters, we may be able to find brain regions uniquely and strongly associated with more computationally specified processes. We can then interrogate the computational process supported by these clusters, by studying the temporal dynamic of each cluster's outputs (e.g. encoding vs. maintenance) and through perturbation (e.g. lesioning).

In order to pursue this line of work, an experimental design better suited for analysing representational similarity is needed. Researchers using this method to study higher visual and categorical representations of objects in the visual processing stream have typically adopted condition-rich designs (Nili et al., 2014), whereby a large number of conditions (in this case, visual objects) are included. This results in a large similarity matrix that samples the stimulus space more thoroughly. Our study with the RNN only had six conditions in total, which resulted in a generally high degree of relatedness between the brain and RNN similarity matrices. Increasing the number of conditions in our tasks is likely to increase the sensitivity of detecting brain-RNN relationships.

Finally, we have demonstrated the potential importance of individual differences in modulating training-related benefits to WM. However, some key points will need to be clarified before drawing firm conclusions. First, the perceptual training control group in our study also showed comparable improvement to one of the SA training groups, rendering it impossible to attribute any improvements to either baseline difference or the practice regimes, or a combination of them. A better alternative, as mentioned in that chapter, is to adopt a control training program that draws less on basic perceptual processing, but rather on a conceptually distant system, such as crystallized intelligence. Related to this, it will also be crucial to ensure that the SA

task is challenging enough throughout training sessions to maximize the potential enhancement and motivation for all participants.

6.5 Conclusions

Across four empirical chapters, this thesis used four different techniques to explore the relationships between SA and WM. Large-scale online testing created normative distributions of carefully controlled tasks, enabling us to create subgroups. The task loadings within these subgroups varied significantly, indicating that different groups of participants were drawing on subtly different underlying processes to perform the tasks. Functional neuroimaging revealed a broadly distributed fronto-parietal network coactivated by both tasks. The activity of one area in particular, the rSPL, mirrored the behavioural findings by indicating differential scaling with task difficulty depending upon the participants' profile across the SA and WM tasks. An RNN was then trained to perform these two tasks, and we systematically explored the role of different clusters of neurons in performing each task. One cluster was equally associated with performance on each task, and this corresponded to the activity pattern of a broad set of fronto-parietal regions in the human fMRI data, coinciding with the result of the conjunction analysis in the fMRI chapter. Finally, we showed that training SA could transfer to WM, provided that participants were equally good at both tasks at baseline. This ambitious combination of techniques represents an initial step towards exploring the individual differences in functional, computational and causal relationships between SA and WM.

References:

References:

- 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.
- Alloway, T. P. (2007). *Automated Working: Memory Assessment: Manual*. Pearson.
- Anderson, E. J., Mannan, S. K., Husain, M., Rees, G., Sumner, P., Mort, D. J., McRobbie, D., & Kennard, C. (2007). Involvement of prefrontal cortex in visual search. *Experimental Brain Research*, *180*(2), 289–302.
- Anderson, E. J., Mannan, S. K., Rees, G., Sumner, P., & Kennard, C. (2008). A Role for Spatial and Nonspatial Working Memory Processes in Visual Search. *Experimental Psychology*, *55*(5), 301–312. <https://doi.org/10.1027/1618-3169.55.5.301>
- Anderson, S., & Kraus, N. (2013). Auditory training: evidence for neural plasticity in older adults. *Perspectives on Hearing and Hearing Disorders: Research and Diagnostics*, *17*(1), 37–57.
- Ashby, F. G., Ell, S. W., Valentin, V. V., & Casale, M. B. (2005). FROST: A distributed neurocomputational model of working memory maintenance. *Journal of Cognitive Neuroscience*, *17*(11), 1728–1743.
- Ashby, F. G., & Helie, S. (2011). A tutorial on computational cognitive neuroscience: Modeling the neurodynamics of cognition. *Journal of Mathematical Psychology*, *55*(4), 273–289.
- Astle, D. E., Harvey, H., Stokes, M., Mohseni, H., Nobre, A. C., & Scerif, G. (2014). Distinct neural mechanisms of individual and developmental differences in VSTM capacity. *Developmental Psychobiology*, *56*(4), 601–610. <https://doi.org/https://doi.org/10.1002/dev.21126>
- 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>
- Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences*, *5*(3), 119–126.
- Awh, E., Smith, E. E., & Jonides, J. (1995). *Human rehearsal processes and the frontal lobes: PET evidence*.
- Awh, E., Vogel, E. K., & Oh, S.-H. (2006). Interactions between attention and working memory. *Neuroscience*, *139*(1), 201–208.
- Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423.
- Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29.
- Baddeley, A D, & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47–89). Elsevier.
- Baddeley, Alan D, Allen, R. J., & Hitch, G. J. (2011). Binding in visual working memory: The role of the episodic buffer. *Neuropsychologia*, *49*(6), 1393–1400.
- Baddeley, Alan David, & Della Sala, S. (1996). Working memory and executive control. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *351*(1346), 1397–1404.

References:

- Badre, D. (2008). Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends in Cognitive Sciences*, *12*(5), 193–200. <https://doi.org/10.1016/j.tics.2008.02.004>
- Badre, D., & D'Esposito, M. (2007). Functional Magnetic Resonance Imaging Evidence for a Hierarchical Organization of the Prefrontal Cortex. *Journal of Cognitive Neuroscience*, *19*(12), 2082–2099. <https://doi.org/10.1162/jocn.2007.19.12.2082>
- Baldassarre, A., Lewis, C. M., Committeri, G., Snyder, A. Z., Romani, G. L., & Corbetta, M. (2012). Individual variability in functional connectivity predicts performance of a perceptual task. *Proceedings of the National Academy of Sciences*, *109*(9), 3516–3521.
- Balinsky, B. (1941). An analysis of the mental factors of various age groups from nine to sixty. *Genetic Psychology Monographs*.
- Baltes, P. B., Staudinger, U. M., & Lindenberger, U. (1999). Lifespan psychology: Theory and application to intellectual functioning. *Annual Review of Psychology*, *50*(1), 471–507.
- Barak, O. (2017). Recurrent neural networks as versatile tools of neuroscience research. *Current Opinion in Neurobiology*, *46*, 1–6.
- Barch, D. M., Carter, C. S., & Committee, C. E. (2008). Measurement issues in the use of cognitive neuroscience tasks in drug development for impaired cognition in schizophrenia: a report of the second consensus building conference of the CNTRICS initiative. *Schizophrenia Bulletin*, *34*(4), 613–618.
- Barnes, J. J., Nobre, A. C., Woolrich, M. W., Baker, K., & Astle, D. E. (2016). Training working memory in childhood enhances coupling between frontoparietal control network and task-related regions. *Journal of Neuroscience*, *36*(34), 9001–9011.
- 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>
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246. <https://doi.org/10.1037/0033-2909.107.2.238>
- Berry, A. S., Zanto, T. P., Clapp, W. C., Hardy, J. L., Delahunt, P. B., Mahncke, H. W., & Gazzaley, A. (2010). The influence of perceptual training on working memory in older adults. *PloS One*, *5*(7), e11537.
- Berryhill, M. E., Chein, J., & Olson, I. R. (2011). At the intersection of attention and memory: The mechanistic role of the posterior parietal lobe in working memory. *Neuropsychologia*, *49*(5), 1306–1315. <https://doi.org/10.1016/j.neuropsychologia.2011.02.033>
- Bettencourt, K. C., & Xu, Y. (2016). Decoding the content of visual short-term memory under distraction in occipital and parietal areas. *Nature Neuroscience*, *19*(1), 150–157.
- Beucher, S., & Meyer, F. (1993). The morphological approach to segmentation: the watershed transformation. *Mathematical Morphology in Image Processing*, *34*, 433–481.
- Bleckley, M. K., Durso, F. T., Crutchfield, J. M., Engle, R. W., & Khanna, M. M. (2003). Individual differences in working memory capacity predict visual attention allocation. *Psychonomic Bulletin & Review*, *10*(4), 884–889.
- Bollinger, J., Rubens, M. T., Masangkay, E., Kalkstein, J., & Gazzaley, A. (2011). An expectation-based memory deficit in aging. *Neuropsychologia*, *49*(6), 1466–1475.
- Bollinger, J., Rubens, M. T., Zanto, T. P., Gazzaley, A., J., B., M. T., R., T. P., Z., & A., G. (2010).

References:

- Expectation-Driven Changes in Cortical Functional Connectivity Influence Working Memory and Long-Term Memory Performance. *Journal of Neuroscience*, 30(43), 14399–14410.
<https://doi.org/10.1523/JNEUROSCI.1547-10.2010>
- Bor, D., Duncan, J., Wiseman, R. J., & Owen, A. M. (2003). Encoding strategies dissociate prefrontal activity from working memory demand. *Neuron*, 37(2), 361–367.
- Botvinick, M. M., & Plaut, D. C. (2006). Short-term memory for serial order: a recurrent neural network model. *Psychological Review*, 113(2), 201.
- Boy, F., & Sumner, P. (2014). Visibility predicts priming within but not between people: a cautionary tale for studies of cognitive individual differences. *Journal of Experimental Psychology: General*, 143(3), 1011.
- Braver, T. S., Paxton, J. L., Locke, H. S., & Barch, D. M. (2009). Flexible neural mechanisms of cognitive control within human prefrontal cortex. *Proceedings of the National Academy of Sciences*, 106(18), 7351–7356.
- Brissenden, J. A., Levin, E. J., Osher, D. E., Halko, M. A., & Somers, D. C. (2016). Functional evidence for a cerebellar node of the dorsal attention network. *Journal of Neuroscience*, 36(22), 6083–6096.
- Broadbent, D. E. (2013). *Perception and communication*. Elsevier.
- Burnham, B. R., Sabia, M., & Langan, C. (2014). Components of working memory and visual selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, 40(1), 391–403. <https://doi.org/10.1037/a0033753>
- Cantor, J., Engle, R. W., & Hamilton, G. (1991). Short-term memory, working memory, and verbal abilities: How do they relate? *Intelligence*, 15(2), 229–246.
- Carlisle, N. B., Arita, J. T., Pardo, D., & Woodman, G. F. (2011). Attentional templates in visual working memory. *Journal of Neuroscience*, 31(25), 9315–9322.
- Carnevale, F., de Lafuente, V., Romo, R., Barak, O., & Parga, N. (2015). Dynamic control of response criterion in premotor cortex during perceptual detection under temporal uncertainty. *Neuron*, 86(4), 1067–1077.
- Casey, B. J., Galvan, A., & Hare, T. A. (2005). Changes in cerebral functional organization during cognitive development. In *Current Opinion in Neurobiology* (Vol. 15, Issue 2, pp. 239–244). Elsevier Ltd. <https://doi.org/10.1016/j.conb.2005.03.012>
- Chein, J. M., & Schneider, W. (2005). Neuroimaging studies of practice-related change: fMRI and meta-analytic evidence of a domain-general control network for learning. *Cognitive Brain Research*, 25(3), 607–623.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233–255.
- Chierchia, G., Fuhrmann, D., Knoll, L. J., Pi-Sunyer, B. P., Sakhardande, A. L., & Blakemore, S.-J. (2019). The matrix reasoning item bank (MaRs-IB): novel, open-access abstract reasoning items for adolescents and adults. *Royal Society Open Science*, 6(10), 190232.
- Chun, M. M. (2011). Visual working memory as visual attention sustained internally over time. *Neuropsychologia*, 49(6), 1407–1409.
- Clapp, W. C., Rubens, M. T., & Gazzaley, A. (2010). Mechanisms of working memory disruption by external interference. *Cerebral Cortex*, 20(4), 859–872.

References:

- Clapp, W. C., Rubens, M. T., Sabharwal, J., & Gazzaley, A. (2011). Deficit in switching between functional brain networks underlies the impact of multitasking on working memory in older adults. *Proceedings of the National Academy of Sciences*, *108*(17), 7212–7217.
- Cokely, E. T., Kelley, C. M., & Gilchrist, A. L. (2006). Sources of individual differences in working memory: Contributions of strategy to capacity. *Psychonomic Bulletin & Review*, *13*(6), 991–997.
- Cole, M. W., Laurent, P., & Stocco, A. (2013). Rapid instructed task learning: a new window into the human brain's unique capacity for flexible cognitive control. *Cognitive, Affective & Behavioral Neuroscience*, *13*(1), 1–22. <https://doi.org/10.3758/s13415-012-0125-7>
- Colflesh, G. J. H., & Conway, A. R. A. (2007). Individual differences in working memory capacity and divided attention in dichotic listening. *Psychonomic Bulletin & Review*, *14*(4), 699–703.
- Conway, A. R. A., Cowan, N., Bunting, M. F., Theriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, *30*(2), 163–183.
- Conway, A. R. A., & Kane, M. J. (2001). Capacity, control and conflict: An individual differences perspective on attentional capture. *Attention, Distraction and Action: Multiple Perspectives on Attention Capture*, *133*, 349–372.
- Corbetta, M. (1998). Frontoparietal cortical networks for directing attention and the eye to visual locations: Identical, independent, or overlapping neural systems? *Proceedings of the National Academy of Sciences*, *95*(3), 831 LP – 838. <https://doi.org/10.1073/pnas.95.3.831>
- Corbetta, M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Ollinger, J. M., Drury, H. A., Linenweber, M. R., Petersen, S. E., Raichle, M. E., & Van Essen, D. C. (1998). A common network of functional areas for attention and eye movements. *Neuron*, *21*(4), 761–773.
- Corbetta, M., Kincade, J. M., & Shulman, G. L. (2002). Neural systems for visual orienting and their relationships to spatial working memory. *Journal of Cognitive Neuroscience*, *14*(3), 508–523.
- Corbetta, M., Miezin, F. M., Shulman, G. L., & Petersen, S. E. (1993). A PET study of visuospatial attention. *Journal of Neuroscience*, *13*(3), 1202–1226.
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201–215. <https://doi.org/10.1038/nrn755>
- Corbetta, M., & Shulman, G. L. (2011). Spatial neglect and attention networks. *Annual Review of Neuroscience*, *34*, 569–599. <https://doi.org/10.1146/annurev-neuro-061010-113731>
- Couperus, J. W., & Mangun, G. R. (2010). Signal enhancement and suppression during visual-spatial selective attention. *Brain Research*, *1359*, 155–177. <https://doi.org/10.1016/j.brainres.2010.08.076>
- 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.
- Cowan, N. (1998). Visual and auditory working memory capacity. *Trends in Cognitive Sciences*, *2*(3), 77.
- Cowan, N. (1999). *An embedded-processes model of working memory*.
- 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/DOI:10.1017/S0140525X01003922>
- Cowan, N. (2008). What are the differences between long-term, short-term, and working memory?

References:

- Progress in Brain Research*, 169, 323–338.
- Cowan, N., Elliott, E. M., Saults, S. J., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, 51(1), 42–100. <https://doi.org/10.1016/j.cogpsych.2004.12.001>
- Craik, F. I. M., & Salthouse, T. A. (2011). *The handbook of aging and cognition*. Psychology press.
- Culham, J. C., Cavanagh, P., & Kanwisher, N. G. (2001). Attention response functions: characterizing brain areas using fMRI activation during parametric variations of attentional load. *Neuron*, 32(4), 737–745.
- Curtis, C. E., & D’Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences*, 7(9), 415–423.
- D’Esposito, M., & Postle, B. R. (1999). The dependence of span and delayed-response performance on prefrontal cortex. *Neuropsychologia*, 37(11), 1303–1315.
- D’Esposito, M., & Postle, B. R. (2015). The Cognitive Neuroscience of Working Memory. *Annual Review of Psychology*, 66(1), 115–142. <https://doi.org/10.1146/annurev-psych-010814-015031>
- Dalmajer, E., Knapton, H., Husain, M., Holmqvist, K., & Niehorster, D. C. (2018). *Division of labour is biased towards equality amongst collaborators in a dyadic short-term memory task*.
- De Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, 47(1), 1–12.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18(1), 193–222.
- Diamantopoulou, S., Poom, L., Klaver, P., & Talsma, D. (2011). Visual working memory capacity and stimulus categories: a behavioral and electrophysiological investigation. *Experimental Brain Research*, 209(4), 501–513.
- Donald, M. (1993). *Origins of the modern mind: Three stages in the evolution of culture and cognition*. Harvard University Press.
- Dosenbach, N. U. F., Fair, D. A., Miezin, F. M., Cohen, A. L., Wenger, K. K., Dosenbach, R. A. T., Fox, M. D., Snyder, A. Z., Vincent, J. L., & Raichle, M. E. (2007). Distinct brain networks for adaptive and stable task control in humans. *Proceedings of the National Academy of Sciences*, 104(26), 11073–11078.
- Douglas, R. J., & Martin, K. A. (1991). A functional microcircuit for cat visual cortex. *The Journal of Physiology*, 440(1), 735–769.
- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, 113(4), 501.
- Duncan, J. (2001). An adaptive coding model of neural function in prefrontal cortex. *Nature Reviews Neuroscience*, 2(11), 820–829.
- Duncan, J., & Humphreys, G. (1992). *Beyond the search surface: Visual search and attentional engagement*.
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, 96(3), 433.

References:

- Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, 23(10), 475–483.
- Eckstein, M. P. (2011). Visual search: A retrospective. *Journal of Vision*, 11(5), 14.
- Edin, F., Klingberg, T., Johansson, P., McNab, F., Tegnér, J., & Compte, A. (2009). Mechanism for top-down control of working memory capacity. *Proceedings of the National Academy of Sciences*, 106(16), 6802–6807.
- Ehrlich, D. B., Stone, J. T., Brandfonbrener, D., Atanasov, A., & Murray, J. D. (2021). PsychRNN: An Accessible and Flexible Python Package for Training Recurrent Neural Network Models on Cognitive Tasks. *Eneuro*, 8(1).
- Eimer, M. (2014). The neural basis of attentional control in visual search. In *Trends in Cognitive Sciences* (Vol. 18, Issue 10, pp. 526–535). Elsevier Ltd. <https://doi.org/10.1016/j.tics.2014.05.005>
- Emrich, S. M., Al-Aidroos, N., Pratt, J., & Ferber, S. (2009). Visual search elicits the electrophysiological marker of visual working memory. *PloS One*, 4(11), e8042.
- Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). *Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex*.
- 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.
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 16(1), 143–149.
- Eriksson, J., Vogel, E. K., Lansner, A., Bergström, F., & Nyberg, L. (2015). Neurocognitive Architecture of Working Memory. *Neuron*, 88(1), 33–46. <https://doi.org/10.1016/j.neuron.2015.09.020>
- Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., & Snyder, M. (2019). fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature Methods*, 16(1), 111–116.
- Estes, W. K., & Taylor, H. A. (1964). A detection method and probabilistic models for assessing information processing from brief visual displays. *Proceedings of the National Academy of Sciences of the United States of America*, 52(2), 446.
- Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review*, 9(1), 59–79.
- Fedorenko, E., Hsieh, P.-J., Nieto-Castañón, A., Whitfield-Gabrieli, S., & Kanwisher, N. (2010). New method for fMRI investigations of language: defining ROIs functionally in individual subjects. *Journal of Neurophysiology*, 104(2), 1177–1194.
- Feredoes, E., Heinen, K., Weiskopf, N., Ruff, C., & Driver, J. (2011). Causal evidence for frontal involvement in memory target maintenance by posterior brain areas during distracter interference of visual working memory. *Proceedings of the National Academy of Sciences*, 108(42), 17510–17515.
- Fletcher, P. C., & Henson, R. N. A. (2001). Frontal lobes and human memory: insights from functional neuroimaging. *Brain*, 124(5), 849–881.
- Foster, D. H., & Ward, P. A. (1991). Asymmetries in oriented-line detection indicate two orthogonal filters in early vision. *Proceedings of the Royal Society of London. Series B: Biological Sciences*,

References:

- 243(1306), 75–81.
- Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences*, *102*(27), 9673–9678.
- Frank, M. J., Loughry, B., & O'Reilly, R. C. (2001). Interactions between frontal cortex and basal ganglia in working memory: a computational model. *Cognitive, Affective, & Behavioral Neuroscience*, *1*(2), 137–160.
- Frost, M. A., & Goebel, R. (2012). Measuring structural–functional correspondence: spatial variability of specialised brain regions after macro-anatomical alignment. *Neuroimage*, *59*(2), 1369–1381.
- Fukuda, K., Vogel, E., Mayr, U., & Awh, E. (2010). Quantity, not quality: The relationship between fluid intelligence and working memory capacity. *Psychonomic Bulletin & Review*, *17*(5), 673–679.
- Fusser, F., Linden, D. E. J., Rahm, B., Hampel, H., Haenschel, C., & Mayer, J. S. (2011). Common capacity-limited neural mechanisms of selective attention and spatial working memory encoding. *European Journal of Neuroscience*, *34*(5), 827–838. <https://doi.org/https://doi.org/10.1111/j.1460-9568.2011.07794.x>
- Gajewski, P. D., Hanisch, E., Falkenstein, M., Thönes, S., & Wascher, E. (2018). What Does the n-Back Task Measure as We Get Older? Relations Between Working-Memory Measures and Other Cognitive Functions Across the Lifespan . In *Frontiers in Psychology* (Vol. 9, p. 2208). <https://www.frontiersin.org/article/10.3389/fpsyg.2018.02208>
- Garrett, H. E. (1946). A developmental theory of intelligence. *American Psychologist*, *1*(372–378).
- Gaspar, J. M., Christie, G. J., Prime, D. J., Jolicœur, P., & McDonald, J. J. (2016). Inability to suppress salient distractors predicts low visual working memory capacity. *Proceedings of the National Academy of Sciences*, *113*(13), 3693–3698. <https://doi.org/10.1073/pnas.1523471113>
- Gazzaley, A., Clapp, W., Kelley, J., McEvoy, K., Knight, R. T., & D'Esposito, M. (2008). Age-related top-down suppression deficit in the early stages of cortical visual memory processing. *Proceedings of the National Academy of Sciences*, *105*(35), 13122–13126.
- Gazzaley, A., Cooney, J. W., McEvoy, K., Knight, R. T., & D'Esposito, M. (2005). Top-down enhancement and suppression of the magnitude and speed of neural activity. *Journal of Cognitive Neuroscience*, *17*(3), 507–517.
- Gazzaley, A., Cooney, J. W., Rissman, J., & D'Esposito, M. (2005a). Top-down suppression deficit underlies working memory impairment in normal aging. *Nature Neuroscience*, *8*(10), 1298–1300. <https://doi.org/10.1038/nn1543>
- Gazzaley, A., Cooney, J. W., Rissman, J., & D'Esposito, M. (2005b). Top-down suppression deficit underlies working memory impairment in normal aging. *Nature Neuroscience*, *8*(10), 1298–1300. <https://doi.org/10.1038/nn1543>
- Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: bridging selective attention and working memory. *Trends in Cognitive Sciences*, *16*(2), 129–135. <https://doi.org/10.1016/j.tics.2011.11.014>
- Gazzaley, A., Rissman, J., Cooney, J., Rutman, A., Seibert, T., Clapp, W., & D'Esposito, M. (2007). Functional interactions between prefrontal and visual association cortex contribute to top-down modulation of visual processing. *Cerebral Cortex*, *17*(suppl_1), i125–i135.
- Gazzaley, A., Rissman, J., D'Esposito, M., & D'Esposito, M. (2004). Functional connectivity during

References:

- working memory maintenance. *Cognitive, Affective, & Behavioral Neuroscience*, 4(4), 580–599. <https://doi.org/10.3758/CABN.4.4.580>
- Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., & Ghosh, S. S. (2011). Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. *Frontiers in Neuroinformatics*, 5, 13.
- Greenwood, P. M., & Parasuraman, R. (2016). The mechanisms of far transfer from cognitive training: Review and hypothesis. *Neuropsychology*, 30(6), 742.
- Greig, T. C., Zito, W., Wexler, B. E., Fiszdon, J., & Bell, M. D. (2007). Improved cognitive function in schizophrenia after one year of cognitive training and vocational services. *Schizophrenia Research*, 96(1–3), 156–161.
- Griffin, I. C., & Nobre, A. C. (2003). Orienting Attention to Locations in Internal Representations. *Journal of Cognitive Neuroscience*, 15(8), 1176–1194. <https://doi.org/10.1162/089892903322598139>
- Grinband, J., Wager, T. D., Lindquist, M., Ferrera, V. P., & Hirsch, J. (2008). Detection of time-varying signals in event-related fMRI designs. *Neuroimage*, 43(3), 509–520.
- Guye, S., De Simoni, C., & von Bastian, C. C. (2017). Do individual differences predict change in cognitive training performance? A latent growth curve modeling approach. *Journal of Cognitive Enhancement*, 1(4), 374–393.
- Hagler Jr, D. J., & Sereno, M. I. (2006). Spatial maps in frontal and prefrontal cortex. *Neuroimage*, 29(2), 567–577.
- Haines, N., Kvam, P., Irving, L., Smith, C., Beauchaine, T., Pitt, M., Ahn, W.-Y., & Turner, B. (2020). *Learning from the reliability paradox: How theoretically informed generative models can advance the social, behavioral, and brain sciences*. <https://doi.org/10.31234/osf.io/xr7y3>
- Han, S.-H., & Kim, M.-S. (2004). Visual search does not remain efficient when executive working memory is working. *Psychological Science*, 15(9), 623–628.
- Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, 458(7238), 632–635.
- Haut, K. M., Lim, K. O., & MacDonald, A. (2010). Prefrontal Cortical Changes Following Cognitive Training in Patients with Chronic Schizophrenia: Effects of Practice, Generalization, and Specificity. *Neuropsychopharmacology*, 35(9), 1850–1859. <https://doi.org/10.1038/npp.2010.52>
- Hedge, C., Powell, G., & Sumner, P. (2017). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, 50(3), 1166–1186.
- Heitz, R. P., & Engle, R. W. (2007). Focusing the spotlight: individual differences in visual attention control. *Journal of Experimental Psychology: General*, 136(2), 217.
- Hirschfeld, G., & Von Brachel, R. (2014). Improving Multiple-Group confirmatory factor analysis in R—A tutorial in measurement invariance with continuous and ordinal indicators. *Practical Assessment, Research, and Evaluation*, 19(1), 7.
- Holmes, J., Woolgar, F., Hampshire, A., & Gathercole, S. E. (2019). Are working memory training effects paradigm-specific? *Frontiers in Psychology*, 10, 1103.
- Houtkamp, R., & Roelfsema, P. R. (2009). Matching of visual input to only one item at any one time. *Psychological Research PRPF*, 73(3), 317–326.

References:

- Hughes, R. W., Hurlstone, M. J., Marsh, J. E., Vachon, F., & Jones, D. M. (2013). Cognitive control of auditory distraction: impact of task difficulty, foreknowledge, and working memory capacity supports duplex-mechanism account. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(2), 539.
- Hutchinson, J. B., Uncapher, M. R., & Wagner, A. D. (2009). Posterior parietal cortex and episodic retrieval: convergent and divergent effects of attention and memory. *Learning & Memory*, *16*(6), 343–356.
- Ikkai, A., & Curtis, C. E. (2011). Common neural mechanisms supporting spatial working memory, attention and motor intention. *Neuropsychologia*, *49*(6), 1428–1434.
- Jenness, J. L., Rosen, M. L., Sambrook, K. A., Dennison, M. J., Lambert, H. K., Sheridan, M. A., & McLaughlin, K. A. (2018). Violence exposure and neural systems underlying working memory for emotional stimuli in youth. *Development and Psychopathology*, *30*(4), 1517–1528. <https://doi.org/DOI: 10.1017/S0954579417001638>
- Jerde, T. A., Merriam, E. P., Riggall, A. C., Hedges, J. H., & Curtis, C. E. (2012). Prioritized maps of space in human frontoparietal cortex. *Journal of Neuroscience*, *32*(48), 17382–17390.
- Jha, A. P., & McCarthy, G. (2000). The influence of memory load upon delay-interval activity in a working-memory task: an event-related functional MRI study. *Journal of Cognitive Neuroscience*, *12*(Supplement 2), 90–105.
- Jonides, J., Lacey, S. C., & Nee, D. E. (2005). Processes of working memory in mind and brain. *Current Directions in Psychological Science*, *14*(1), 2–5.
- Jost, K., Bryck, R. L., Vogel, E. K., & Mayr, U. (2011). Are old adults just like low working memory young adults? Filtering efficiency and age differences in visual working memory. *Cerebral Cortex*, *21*(5), 1147–1154.
- Julian, J. B., Fedorenko, E., Webster, J., & Kanwisher, N. (2012). An algorithmic method for functionally defining regions of interest in the ventral visual pathway. *Neuroimage*, *60*(4), 2357–2364.
- Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, *130*(2), 169.
- Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, *133*(2), 189.
- Kane, M. J., Poole, B. J., Tuholski, S. W., & Engle, R. W. (2006). Working memory capacity and the top-down control of visual search: Exploring the boundaries of " executive attention ". *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(4), 749.
- Kang, M.-S., & Woodman, G. F. (2014). The neurophysiological index of visual working memory maintenance is not due to load dependent eye movements. *Neuropsychologia*, *56*, 63–72.
- Kanwisher, N., & Wojciulik, E. (2000). Visual attention: insights from brain imaging. *Nature Reviews Neuroscience*, *1*(2), 91–100.
- Karbach, J., Könen, T., & Spengler, M. (2017). Who benefits the most? Individual differences in the transfer of executive control training across the lifespan. *Journal of Cognitive Enhancement*, *1*(4), 394–405.
- Karbach, J., & Verhaeghen, P. (2014). Making working memory work: a meta-analysis of executive-

References:

- control and working memory training in older adults. *Psychological Science*, 25(11), 2027–2037.
- Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron*, 22(4), 751–761.
- Kelly, A. M. C., & Garavan, H. (2005). Human functional neuroimaging of brain changes associated with practice. *Cerebral Cortex*, 15(8), 1089–1102.
- Kessels, R. P. C., van Zandvoort, M. J. E., Postma, A., Kappelle, L. J., & de Haan, E. H. F. (2000). The Corsi Block-Tapping Task: Standardization and Normative Data. *Applied Neuropsychology*, 7(4), 252–258. https://doi.org/10.1207/s15324826an0704_8
- Keye, D., Wilhelm, O., Oberauer, K., & van Ravenzwaaij, D. (2009). Individual differences in conflict-monitoring: testing means and covariance hypothesis about the Simon and the Eriksen Flanker task. *Psychological Research PRPF*, 73(6), 762–776. <https://doi.org/10.1007/s00426-008-0188-9>
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *ArXiv Preprint ArXiv:1412.6980*.
- Kiyonaga, A., & Egner, T. (2013). Working memory as internal attention: Toward an integrative account of internal and external selection processes. *Psychonomic Bulletin & Review*, 20(2), 228–242.
- Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlström, K., Gillberg, C. G., Forsberg, H., & Westerberg, H. (2005). Computerized training of working memory in children with ADHD—a randomized, controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, 44(2), 177–186.
- Koechlin, E., & Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences*, 11(6), 229–235.
- Konoike, N., Kotozaki, Y., Jeong, H., Miyazaki, A., Sakaki, K., Shinada, T., Sugiura, M., Kawashima, R., & Nakamura, K. (2015). Temporal and motor representation of rhythm in fronto-parietal cortical areas: an fMRI study. *PloS One*, 10(6), e0130120.
- Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008). Representational similarity analysis—connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, 2, 4.
- Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1), 79–86.
- Kundu, B., Sutterer, D. W., Emrich, S. M., & Postle, B. R. (2013). Strengthened effective connectivity underlies transfer of working memory training to tests of short-term memory and attention. *Journal of Neuroscience*, 33(20), 8705–8715.
- Kuo, B.-C., Nobre, A. C., Scerif, G., & Astle, D. E. (2016). Top-Down Activation of Spatiotopic Sensory Codes in Perceptual and Working Memory Search. *Journal of Cognitive Neuroscience*, 28(7), 996–1009. https://doi.org/10.1162/jocn_a_00952
- Kuo, B.-C., Rao, A., Lepsien, J., Nobre, A. C., B.-C., K., A., R., J., L., & A. C., N. (2009). Searching for Targets within the Spatial Layout of Visual Short-Term Memory. *Journal of Neuroscience*, 29(25), 8032–8038. <https://doi.org/10.1523/JNEUROSCI.0952-09.2009>
- Kuo, B.-C., Yeh, Y. Y., Chen, A. J. W., & D’Esposito, M. (2011). Functional connectivity during top-down modulation of visual short-term memory representations. *Neuropsychologia*, 49(6), 1589–1596. <https://doi.org/10.1016/j.neuropsychologia.2010.12.043>

References:

- 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.
- Labar, K. S., Gitelman, D. R., Parrish, T. B., & Mesulam, M. M. (1999). Neuroanatomic overlap of working memory and spatial attention networks: A functional MRI comparison within subjects. *NeuroImage*, *10*(6), 695–704. <https://doi.org/10.1006/nimg.1999.0503>
- Landau, S. M., Schumacher, E. H., Garavan, H., Druzgal, T. J., & D'Esposito, M. (2004). A functional MRI study of the influence of practice on component processes of working memory. *NeuroImage*, *22*(1), 211–221. <https://doi.org/10.1016/j.neuroimage.2004.01.003>
- Landman, R., Spekreijse, H., & Lamme, V. A. F. 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](https://doi.org/10.1016/s0042-6989(02)00402-9)
- Lange, K., Kühn, S., & Filevich, E. (2015). "Just Another Tool for Online Studies"(JATOS): An easy solution for setup and management of web servers supporting online studies. *PloS One*, *10*(6), e0130834.
- Lavie, N., & De Fockert, J. (2005). The role of working memory in attentional capture. *Psychonomic Bulletin & Review*, *12*(4), 669–674.
- Lee, E.-Y., Cowan, N., Vogel, E. K., Rolan, T., Valle-Inclan, 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.
- Lee, S.-H., Kravitz, D. J., & Baker, C. I. (2013). Goal-dependent dissociation of visual and prefrontal cortices during working memory. *Nature Neuroscience*, *16*(8), 997–999.
- Lepsien, J., Griffin, I. C., Devlin, J. T., & Nobre, A. C. (2005). Directing spatial attention in mental representations: Interactions between attentional orienting and working-memory load. *NeuroImage*, *26*(3), 733–743. <https://doi.org/10.1016/j.neuroimage.2005.02.026>
- Lepsien, J., & Nobre, A. C. (2006). Attentional Modulation of Object Representations in Working Memory. *Cerebral Cortex*, *17*(9), 2072–2083. <https://doi.org/10.1093/cercor/bhl116>
- Lepsien, J., & Nobre, A. C. (2007). Attentional Modulation of Object Representations in Working Memory. *Cerebral Cortex*, *17*(9), 2072–2083. <https://doi.org/10.1093/cercor/bhl116>
- Lewis, C. M., Baldassarre, A., Committeri, G., Romani, G. L., & Corbetta, M. (2009). Learning sculpts the spontaneous activity of the resting human brain. *Proceedings of the National Academy of Sciences*, *106*(41), 17558–17563.
- 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>
- Li, S.-C., Lindenberger, U., Hommel, B., Aschersleben, G., Prinz, W., & Baltes, P. B. (2004). Transformations in the couplings among intellectual abilities and constituent cognitive processes across the life span. *Psychological Science*, *15*(3), 155–163.
- López-Luengo, B., & Vázquez, C. (2003). Effects of attention process training on cognitive functioning of schizophrenic patients. *Psychiatry Research*, *119*(1–2), 41–53.
- Lövdén, M., Brehmer, Y., Li, S.-C., & Lindenberger, U. (2012). Training-induced compensation versus magnification of individual differences in memory performance. *Frontiers in Human Neuroscience*, *6*, 141.

References:

- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281.
- 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>
- Lundqvist, A., Grundström, K., Samuelsson, K., & Rönnerberg, J. (2010). Computerized training of working memory in a group of patients suffering from acquired brain injury. *Brain Injury*, *24*(10), 1173–1183.
- Luria, R., Balaban, H., Awh, E., & Vogel, E. K. (2016). The contralateral delay activity as a neural measure of visual working memory. *Neuroscience & Biobehavioral Reviews*, *62*, 100–108.
- Luria, R., & Vogel, E. K. (2011). Visual Search Demands Dictate Reliance on Working Memory Storage. *Journal of Neuroscience*, *31*(16), 6199–6207. <https://doi.org/10.1523/JNEUROSCI.6453-10.2011>
- Lustig, C., May, C. P., & Hasher, L. (2001). Working memory span and the role of proactive interference. *Journal of Experimental Psychology: General*, *130*(2), 199.
- Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, *17*(3), 347.
- Majerus, S., Attout, L., D'Argembeau, A., Degueldre, C., Fias, W., Maquet, P., Martinez Perez, T., Stawarczyk, D., Salmon, E., & Van der Linden, M. (2012). Attention supports verbal short-term memory via competition between dorsal and ventral attention networks. *Cerebral Cortex*, *22*(5), 1086–1097.
- Majerus, S., Péters, F., Bouffier, M., Cowan, N., & Phillips, C. (2018). The dorsal attention network reflects both encoding load and top-down control during working memory. *Journal of Cognitive Neuroscience*, *30*(2), 144–159.
- Makino, Y., Yokosawa, K., Takeda, Y., & Kumada, T. (2004). Visual search and memory search engage extensive overlapping cerebral cortices: An fMRI study. *NeuroImage*, *23*(2), 525–533. <https://doi.org/10.1016/j.neuroimage.2004.06.026>
- Mante, V., Sussillo, D., Shenoy, K. V., & Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature*, *503*(7474), 78–84.
- 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 & Adolescent Psychiatry*, *44*(4), 377–384.
- Masse, N. Y., Rosen, M. C., & Freedman, D. J. (2020). Reevaluating the role of persistent neural activity in short-term memory. *Trends in Cognitive Sciences*, *24*(3), 242–258.
- Masse, N. Y., Yang, G. R., Song, H. F., Wang, X.-J., & Freedman, D. J. (2019). Circuit mechanisms for the maintenance and manipulation of information in working memory. *Nature Neuroscience*, *22*(7), 1159–1167.
- Mayer, J. S., Bittner, R. A., Nikolić, D., Bledowski, C., Goebel, R., & Linden, D. E. J. (2007). Common neural substrates for visual working memory and attention. *NeuroImage*, *36*(2), 441–453. <https://doi.org/10.1016/j.neuroimage.2007.03.007>
- McNab, F., & Klingberg, T. (2008). Prefrontal cortex and basal ganglia control access to working memory. *Nature Neuroscience*, *11*(1), 103–107.

References:

- Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology, 49*(2), 270.
- Miller, B. T., Vytlačil, J., Fegen, D., Pradhan, S., & D'Esposito, M. (2011). The prefrontal cortex modulates category selectivity in human extrastriate cortex. *Journal of Cognitive Neuroscience, 23*(1), 1–10.
- Miller, E. K., & Cohen, J. D. (2001). An Integrative Theory of Prefrontal Cortex Function. *Annual Review of Neuroscience, 24*(1), 167–202. <https://doi.org/10.1146/annurev.neuro.24.1.167>
- Minear, M., Brasher, F., Guerrero, C. B., Brasher, M., Moore, A., & Sukeena, J. (2016). A simultaneous examination of two forms of working memory training: Evidence for near transfer only. *Memory & Cognition, 44*(7), 1014–1037.
- Mishra, J., de Villers-Sidani, E., Merzenich, M., & Gazzaley, A. (2014). Adaptive Training Diminishes Distractibility in Aging across Species. *Neuron, 84*(5), 1091–1103. <https://doi.org/10.1016/j.neuron.2014.10.034>
- Mishra, J., Rolle, C., & Gazzaley, A. (2015). Neural plasticity underlying visual perceptual learning in aging. *Brain Research, 1612*, 140–151.
- Miyake, A., & Shah, P. (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. Cambridge University Press.
- Moreau, D., Kirk, I. J., & Waldie, K. E. (2016). Seven pervasive statistical flaws in cognitive training interventions. *Frontiers in Human Neuroscience, 10*, 153.
- 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.
- Murray, A. M., Nobre, A. C., Clark, I. A., Cravo, A. M., & Stokes, M. G. (2013). Attention restores discrete items to visual short-term memory. *Psychological Science, 24*(4), 550–556.
- Nee, D. E., & Jonides, J. (2009). Common and distinct neural correlates of perceptual and memorial selection. *Neuroimage, 45*(3), 963–975.
- Neisser, U. (1964). Visual search. *Scientific American, 210*(6), 94–103.
- Nieder, A. (2004). The number domain—can we count on parietal cortex? *Neuron, 44*(3), 407–409.
- Nieto-Castañón, A., & Fedorenko, E. (2012). Subject-specific functional localizers increase sensitivity and functional resolution of multi-subject analyses. *NeuroImage, 63*(3), 1646–1669. <https://doi.org/10.1016/j.neuroimage.2012.06.065>
- Nili, H., Wingfield, C., Walther, A., Su, L., Marslen-Wilson, W., & Kriegeskorte, N. (2014). A Toolbox for Representational Similarity Analysis. *PLOS Computational Biology, 10*(4), e1003553. <https://doi.org/10.1371/journal.pcbi.1003553>
- Nobre, A. C., Coull, J. T., Walsh, V., & Frith, C. D. (2003). Brain activations during visual search: contributions of search efficiency versus feature binding. *Neuroimage, 18*(1), 91–103.
- Norman, D. A., & Shallice, T. (1986). *Attention to Action BT - Consciousness and Self-Regulation: Advances in Research and Theory Volume 4* (R. J. Davidson, G. E. Schwartz, & D. Shapiro (eds.); pp. 1–18). Springer US. https://doi.org/10.1007/978-1-4757-0629-1_1
- Oberauer, K. (2001). Removing irrelevant information from working memory: a cognitive aging study

References:

- with the modified Sternberg task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(4), 948.
- 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.
- Oberauer, K. (2005). Control of the contents of working memory--a comparison of two paradigms and two age groups. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(4), 714.
- Oberauer, K. (2013). The focus of attention in working memory—from metaphors to mechanisms. *Frontiers in Human Neuroscience*, 7, 673.
- Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (2012). Modeling working memory: An interference model of complex span. *Psychonomic Bulletin & Review*, 19(5), 779–819.
- Oh, S.-H., & Kim, M.-S. (2004). The role of spatial working memory in visual search efficiency. *Psychonomic Bulletin & Review*, 11(2), 275–281.
- Pessoa, L., Kastner, S., & Ungerleider, L. G. (2003). Neuroimaging Studies of Attention: From Modulation of Sensory Processing to Top-Down Control. *The Journal of Neuroscience*, 23(10), 3990 LP – 3998. <https://doi.org/10.1523/JNEUROSCI.23-10-03990.2003>
- Peterson, D. J., & Berryhill, M. E. (2013). The Gestalt principle of similarity benefits visual working memory. *Psychonomic Bulletin & Review*, 20(6), 1282–1289.
- Peterson, M. S., Kramer, A. F., Wang, R. F., Irwin, D. E., & McCarley, J. S. (2001). Visual search has memory. *Psychological Science*, 12(4), 287–292.
- Ploran, E. J., Nelson, S. M., Velanova, K., Donaldson, D. I., Petersen, S. E., & Wheeler, M. E. (2007). Evidence accumulation and the moment of recognition: dissociating perceptual recognition processes using fMRI. *Journal of Neuroscience*, 27(44), 11912–11924.
- Pollmann, S., & von Cramon, D. Y. (2000). Object working memory and visuospatial processing: functional neuroanatomy analyzed by event-related fMRI. *Experimental Brain Research*, 133(1), 12–22.
- Poole, B. J., & Kane, M. J. (2009). Working-memory capacity predicts the executive control of visual search among distractors: The influences of sustained and selective attention. *Quarterly Journal of Experimental Psychology*, 62(7), 1430–1454.
- Posner, M. I., & DiGirolamo, G. J. (1998). *18 Executive Attention: Conflict, Target Detection, and Cognitive Control*.
- Postle, B. R., Awh, E., Jonides, J., Smith, E. E., & D'Esposito, M. (2004). The where and how of attention-based rehearsal in spatial working memory. *Cognitive Brain Research*, 20(2), 194–205. <https://doi.org/10.1016/j.cogbrainres.2004.02.008>
- Pratt, N., Willoughby, A., & Swick, D. (2011). Effects of working memory load on visual selective attention: behavioral and electrophysiological evidence. *Frontiers in Human Neuroscience*, 5, 57.
- Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. In *Developmental Review* (Vol. 41, pp. 71–90). Mosby Inc. <https://doi.org/10.1016/j.dr.2016.06.004>
- Qi, S., Ding, C., & Li, H. (2014). Neural correlates of inefficient filtering of emotionally neutral distractors from working memory in trait anxiety. *Cognitive, Affective, & Behavioral Neuroscience*,

References:

- 14(1), 253–265.
- Rainer, G., Asaad, W. F., & Miller, E. K. (1998). Selective representation of relevant information by neurons in the primate prefrontal cortex. *Nature*, 393(6685), 577–579.
- Rajan, K., & Abbott, L. F. (2006). Eigenvalue Spectra of Random Matrices for Neural Networks. *Phys. Rev. Lett.*, 97(18), 188104. <https://doi.org/10.1103/PhysRevLett.97.188104>
- Rao, S. C., Rainer, G., & Miller, E. K. (1997). Integration of what and where in the primate prefrontal cortex. *Science*, 276(5313), 821–824.
- Raposo, D., Kaufman, M. T., & Churchland, A. K. (2014). A category-free neural population supports evolving demands during decision-making. *Nature Neuroscience*, 17(12), 1784–1792. <https://doi.org/10.1038/nn.3865>
- Rappport, M. D., Chung, K.-M., Shore, G., & Isaacs, P. (2001). A conceptual model of child psychopathology: Implications for understanding attention deficit hyperactivity disorder and treatment efficacy. *Journal of Clinical Child & Adolescent Psychology*, 30(1), 48–58.
- Redick, T. S., & Engle, R. W. (2006). Working memory capacity and attention network test performance. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 20(5), 713–721.
- Redick, T. S., & Lindsey, D. R. B. (2013). Complex span and n-back measures of working memory: A meta-analysis. *Psychonomic Bulletin & Review*, 20(6), 1102–1113.
- Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., Kane, M. J., & 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.
- Reerko, 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>
- Reynolds, J. H., & Chelazzi, L. (2004). Attentional modulation of visual processing. *Annu. Rev. Neurosci.*, 27, 611–647.
- Richter, F. R., Cooper, R. A., Bays, P. M., & Simons, J. S. (2016). Distinct neural mechanisms underlie the success, precision, and vividness of episodic memory. *Elife*, 5, e18260.
- Riggall, A. C., & Postle, B. R. (2012). The relationship between working memory storage and elevated activity as measured with functional magnetic resonance imaging. *Journal of Neuroscience*, 32(38), 12990–12998.
- Romo, R., Brody, C. D., Hernández, A., & Lemus, L. (1999). Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature*, 399(6735), 470–473.
- Roome, H. E., Towse, J. N., & Jarrold, C. (2014). How do selective attentional processes contribute to maintenance and recall in children’s working memory capacity? *Frontiers in Human Neuroscience*, 8, 1011.
- Roth, J. K., Serences, J. T., & Courtney, S. M. (2005). Neural System for Controlling the Contents of Object Working Memory in Humans. *Cerebral Cortex*, 16(11), 1595–1603. <https://doi.org/10.1093/cercor/bhj096>
- Rottschy, C., Langner, R., Dogan, I., Reetz, K., Laird, A. R., Schulz, J. B., Fox, P. T., & Eickhoff, S. B. (2012). Modelling neural correlates of working memory: A coordinate-based meta-analysis.

References:

- NeuroImage*, 60(1), 830–846. <https://doi.org/10.1016/j.neuroimage.2011.11.050>
- Ruff, C. C., Bestmann, S., Blankenburg, F., Bjoertomt, O., Josephs, O., Weiskopf, N., Deichmann, R., & Driver, J. (2008). Distinct causal influences of parietal versus frontal areas on human visual cortex: evidence from concurrent TMS–fMRI. *Cerebral Cortex*, 18(4), 817–827.
- Ruff, C. C., Driver, J., & Bestmann, S. (2009). Combining TMS and fMRI: from ‘virtual lesions’ to functional-network accounts of cognition. *Cortex*, 45(9), 1043–1049.
- Rutman, A. M., Clapp, W. C., Chadick, J. Z., & Gazzaley, A. (2010). Early top-down control of visual processing predicts working memory performance. *Journal of Cognitive Neuroscience*, 22(6), 1224–1234. <https://doi.org/10.1162/jocn.2009.21257>
- Sadaghiani, S., & D’Esposito, M. (2015). Functional Characterization of the Cingulo-Opercular Network in the Maintenance of Tonic Alertness. *Cerebral Cortex*, 25(9), 2763–2773. <https://doi.org/10.1093/cercor/bhu072>
- Salthouse, T. A., Hancock, H. E., Meinz, E. J., & Hambrick, D. Z. (1996). Interrelations of age, visual acuity, and cognitive functioning. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 51(6), P317-330.
- Sauseng, P., & Klimesch, W. (2008). What does phase information of oscillatory brain activity tell us about cognitive processes? *Neuroscience & Biobehavioral Reviews*, 32(5), 1001–1013.
- Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X.-N., Holmes, A. J., Eickhoff, S. B., & Yeo, B. T. T. (2018). Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cerebral Cortex*, 28(9), 3095–3114.
- Schielzeth, H., Dingemanse, N. J., Nakagawa, S., Westneat, D. F., Algue, H., Teplitsky, C., Réale, D., Dochtermann, N. A., Garamszegi, L. Z., & Araya-Ajoy, Y. G. (2020). Robustness of linear mixed-effects models to violations of distributional assumptions. *Methods in Ecology and Evolution*, 11(9), 1141–1152. <https://doi.org/https://doi.org/10.1111/2041-210X.13434>
- Schmiedek, F., Lövdén, M., & Lindenberger, U. (2014). A task is a task is a task: putting complex span, n-back, and other working memory indicators in psychometric context. *Frontiers in Psychology*, 5, 1475. <https://doi.org/10.3389/fpsyg.2014.01475>
- Schurgin, M. W., Wixted, J. T., & Brady, T. F. (2020). Psychophysical scaling reveals a unified theory of visual memory strength. *Nature Human Behaviour*, 4(11), 1156–1172. <https://doi.org/10.1038/s41562-020-00938-0>
- Schwark, J., Sandry, J., & Dolgov, I. (2013). Evidence for a Positive Relationship between Working-Memory Capacity and Detection of Low-Prevalence Targets in Visual Search. *Perception*, 42(1), 112–114. <https://doi.org/10.1068/p7386>
- Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science*, 20(2), 207–214.
- Serences, J. T., Schwarzbach, J., Courtney, S. M., Golay, X., & Yantis, S. (2004). Control of object-based attention in human cortex. *Cerebral Cortex*, 14(12), 1346–1357.
- Sestieri, C., Shulman, G. L., & Corbetta, M. (2012). *Orienting to the environment: Separate contributions of dorsal and ventral frontoparietal attention networks.*
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, 84(2), 127.

References:

- Shipstead, Z., Lindsey, D. R. B., Marshall, R. L., & Engle, R. W. (2014). The mechanisms of working memory capacity: Primary memory, secondary memory, and attention control. *Journal of Memory and Language*, *72*(1), 116–141. <https://doi.org/10.1016/j.jml.2014.01.004>
- Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin*, *138*(4), 628.
- Silk, T. J., Bellgrove, M. A., Wrafter, P., Mattingley, J. B., & Cunnington, R. (2010). Spatial working memory and spatial attention rely on common neural processes in the intraparietal sulcus. *Neuroimage*, *53*(2), 718–724.
- Silver, M. A., & Kastner, S. (2009). Topographic maps in human frontal and parietal cortex. *Trends in Cognitive Sciences*, *13*(11), 488–495.
- Sligte, I. G., Scholte, H. S., & Lamme, V. A. F. (2008). Are There Multiple Visual Short-Term Memory Stores? *PLOS ONE*, *3*(2), e1699. <https://doi.org/10.1371/journal.pone.0001699>
- Smith, E. E., Jonides, J., Koeppe, R. A., Awh, E., Schumacher, E. H., & Minoshima, S. (1995). Spatial versus object working memory: PET investigations. *Journal of Cognitive Neuroscience*, *7*(3), 337–356.
- Smoleń, T., Jastrzebski, J., Estrada, E., & Chuderski, A. (2018). Most evidence for the compensation account of cognitive training is unreliable. *Memory & Cognition*, *46*(8), 1315–1330.
- Sobel, K. V., Gerrie, M. P., Poole, B. J., & Kane, M. J. (2007). Individual differences in working memory capacity and visual search: The roles of top-down and bottom-up processing. *Psychonomic Bulletin & Review*, *14*(5), 840–845.
- Song, H. F., Yang, G. R., & Wang, X.-J. (2016). Training Excitatory-Inhibitory Recurrent Neural Networks for Cognitive Tasks: A Simple and Flexible Framework. *PLOS Computational Biology*, *12*(2), e1004792. <https://doi.org/10.1371/journal.pcbi.1004792>
- Sörqvist, P. (2010). High working memory capacity attenuates the deviation effect but not the changing-state effect: Further support for the duplex-mechanism account of auditory distraction. *Memory & Cognition*, *38*(5), 651–658.
- Spearman, C. (1910). Correlation calculated from faulty data. *British Journal of Psychology*, *3*(3), 271.
- Spinnler, H., Sala, S. Della, Bandera, R., & Baddeley, A. (1988). Dementia, ageing, and the structure of human memory. *Cognitive Neuropsychology*, *5*(2), 193–211.
- Sreenivasan, K. K., Curtis, C. E., & D’Esposito, M. (2014). Revisiting the role of persistent neural activity during working memory. In *Trends in Cognitive Sciences* (Vol. 18, Issue 2, pp. 82–89). Elsevier Ltd. <https://doi.org/10.1016/j.tics.2013.12.001>
- Sreenivasan, K. K., & Jha, A. P. (2007). Selective attention supports working memory maintenance by modulating perceptual processing of distractors. *Journal of Cognitive Neuroscience*, *19*(1), 32–41.
- Steiger, J. H. (1990). Structural Model Evaluation and Modification: An Interval Estimation Approach. *Multivariate Behavioral Research*, *25*(2), 173–180. https://doi.org/10.1207/s15327906mbr2502_4
- Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders’ method. *Acta Psychologica*, *30*, 276–315.
- Stokes, M. G., Kusunoki, M., Sigala, N., Nili, H., Gaffan, D., & Duncan, J. (2013). Dynamic Coding for Cognitive Control in Prefrontal Cortex. *Neuron*, *78*(2), 364–375. <https://doi.org/10.1016/j.neuron.2013.01.039>

References:

- Stokes, M., Thompson, R., Cusack, R., & Duncan, J. (2009). Top-down activation of shape-specific population codes in visual cortex during mental imagery. *Journal of Neuroscience*, *29*(5), 1565–1572.
- Strenziok, M., Parasuraman, R., Clarke, E., Cisler, D. S., Thompson, J. C., & Greenwood, P. M. (2014). Neurocognitive enhancement in older adults: comparison of three cognitive training tasks to test a hypothesis of training transfer in brain connectivity. *Neuroimage*, *85*, 1027–1039.
- Szczepanski, S. M., Pinsk, M. A., Douglas, M. M., Kastner, S., & Saalmann, Y. B. (2013). Functional and structural architecture of the human dorsal frontoparietal attention network. *Proceedings of the National Academy of Sciences*, *110*(39), 15806–15811.
- Tamber-Rosenau, B. J., Esterman, M., Chiu, Y.-C., & Yantis, S. (2011). Cortical mechanisms of cognitive control for shifting attention in vision and working memory. *Journal of Cognitive Neuroscience*, *23*(10), 2905–2919. <https://doi.org/10.1162/jocn.2011.21608>
- Tamez, E., Myerson, J., & Hale, S. (2012). Contributions of associative learning to age and individual differences in fluid intelligence. *Intelligence*, *40*(5), 518–529.
- Tamm, L., Menon, V., & Reiss, A. L. (2002). Maturation of Brain Function Associated with Response Inhibition. *Journal of the American Academy of Child and Adolescent Psychiatry*, *41*(10), 1231–1238. <https://doi.org/10.1097/00004583-200210000-00013>
- Thirion, B., Pinel, P., Mériaux, S., Roche, A., Dehaene, S., & Poline, J.-B. (2007). Analysis of a large fMRI cohort: Statistical and methodological issues for group analyses. *Neuroimage*, *35*(1), 105–120.
- Thorell, L. B., Lindqvist, S., Bergman Nutley, S., Bohlin, G., & Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental Science*, *12*(1), 106–113.
- Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature*, *428*(6984), 751–754.
- Tomasi, D., Chang, L., Caparelli, E. C., & Ernst, T. (2007). Different activation patterns for working memory load and visual attention load. *Brain Research*, *1132*, 158–165.
- Treisman, A. M. (1964). Selective attention in man. *British Medical Bulletin*, *20*(1), 12–16.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136.
- Treisman, A. M., & Souther, J. (1985). Search asymmetry: a diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General*, *114*(3), 285.
- Turner, B. O., Paul, E. J., Miller, M. B., & Barbey, A. K. (2018). Small sample sizes reduce the replicability of task-based fMRI studies. *Communications Biology*, *1*(1), 62. <https://doi.org/10.1038/s42003-018-0073-z>
- Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*(2), 127–154.
- Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval. *Cognitive Psychology*, *71*, 1–26.
- Van der Molen, Mj., Van Luit, J. E. H., Van der Molen, M. W., Klugkist, I., & Jongmans, M. J. (2010). Effectiveness of a computerised working memory training in adolescents with mild to borderline intellectual disabilities. *Journal of Intellectual Disability Research*, *54*(5), 433–447.

References:

- Van Der Velde, F., & De Kamps, M. (2001). From knowing what to knowing where: Modeling object-based attention with feedback disinhibition of activation. *Journal of Cognitive Neuroscience*, *13*(4), 479–491.
- Van Vreeswijk, C., & Sompolinsky, H. (1996). Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science*, *274*(5293), 1724–1726.
- Van Wert, M., Nova, N., Horowitz, T., & Wolfe, J. (2008). What does performance on one visual search task tell you about performance on another? *Journal of Vision*, *8*(6), 312. <https://doi.org/10.1167/8.6.312>
- Vandierendonck, A. (2014). Symbiosis of executive and selective attention in working memory. *Frontiers in Human Neuroscience*, *8*, 588.
- Vincent, J. L., Kahn, I., Snyder, A. Z., Raichle, M. E., & Buckner, R. L. (2008). Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *Journal of Neurophysiology*, *100*(6), 3328–3342.
- Vincent, J. L., Snyder, A. Z., Fox, M. D., Shannon, B. J., Andrews, J. R., Raichle, M. E., & Buckner, R. L. (2006). Coherent spontaneous activity identifies a hippocampal-parietal memory network. *Journal of Neurophysiology*, *96*(6), 3517–3531.
- Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*(6984), 748–751. <https://doi.org/10.1038/nature02447>
- Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, *438*(7067), 500–503. <https://doi.org/10.1038/nature04171>
- Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory. *Cognitive, Affective, & Behavioral Neuroscience*, *3*(4), 255–274.
- Wallis, G., Stokes, M., Cousijn, H., Woolrich, M., & Nobre, A. C. (2015). Frontoparietal and Cingulo-occipital Networks Play Dissociable Roles in Control of Working Memory. *Journal of Cognitive Neuroscience*, *27*(10), 2019–2034. https://doi.org/10.1162/jocn_a_00838
- Wass, S. V., Scerif, G., & Johnson, M. H. (2012). Training attentional control and working memory—Is younger, better? *Developmental Review*, *32*(4), 360–387.
- Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *Journal of Strength and Conditioning Research*, *19*(1), 231–240. <https://doi.org/10.1519/15184.1>
- Westerberg, H., Jacobaeus, H., Hirvikoski, T., Clevberger, P., Östensson, M.-L., Bartfai, A., & Klingberg, T. (2007). Computerized working memory training after stroke—a pilot study. *Brain Injury*, *21*(1), 21–29.
- 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>
- Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. *Biological Psychiatry*, *57*(11), 1336–1346.
- Williams, D. E., Reingold, E. M., Moscovitch, M., & Behrmann, M. (1997). Patterns of eye movements during parallel and serial visual search tasks. *Canadian Journal of Experimental Psychology/Revue*

References:

- Canadienne de Psychologie Expérimentale*, 51(2), 151.
- Williams, L. H., & Drew, T. (2018). Working memory capacity predicts search accuracy for novel as well as repeated targets. *Visual Cognition*, 26(6), 463–474. <https://doi.org/10.1080/13506285.2018.1490370>
- Wolfe, J. M. (1994). Guided search 2.0 a revised model of visual search. *Psychonomic Bulletin & Review*, 1(2), 202–238.
- Wolfe, J. M. (2015). *Visual search*.
- Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: an alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 15(3), 419.
- Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, 5(6), 495–501.
- Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, 1(3), 58. <https://doi.org/10.1038/s41562-017-0058>
- Woodman, G. F., & Arita, J. T. (2011). Direct electrophysiological measurement of attentional templates in visual working memory. *Psychological Science*, 22(2), 212–215.
- Woodman, G. F., Vogel, E. K., & Luck, S. J. (2001). Visual search remains efficient when visual working memory is full. *Psychological Science*, 12(3), 219–224.
- Woolgar, A., Afshar, S., Williams, M. A., & Rich, A. N. (2015). Flexible coding of task rules in frontoparietal cortex: an adaptive system for flexible cognitive control. *Journal of Cognitive Neuroscience*, 27(10), 1895–1911.
- Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, 440(7080), 91–95. <https://doi.org/10.1038/nature04262>
- Yang, G. R., Joglekar, M. R., Song, H. F., Newsome, W. T., & Wang, X.-J. (2019). Task representations in neural networks trained to perform many cognitive tasks. *Nature Neuroscience*, 22(2), 297–306. <https://doi.org/10.1038/s41593-018-0310-2>
- Yeo, B. T. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., & Polimeni, J. R. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*.
- Zanto, T. P., Rubens, M. T., Thangavel, A., & Gazzaley, A. (2011). Causal role of the prefrontal cortex in top-down modulation of visual processing and working memory. *Nature Neuroscience*, 14(5), 656–661. <https://doi.org/10.1038/nn.2773>
- Zhang, M., Rennie, J. P., Hawkins, E., Bathelt, J., & Astle, D. E. (2020). Mapping differential responses to cognitive training using machine learning. *Developmental Science*, 23(4), e12868.
- Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, 453(7192), 233–235. <https://doi.org/10.1038/nature06860>
- Zinke, K., Zeintl, M., Rose, N. S., Putzmann, J., Pydde, A., & Kliegel, M. (2014). Working memory training and transfer in older adults: effects of age, baseline performance, and training gains. *Developmental Psychology*, 50(1), 304.