Interactions between Visual Working Memory and Attention: Evidence from Functional Magnetic Resonance Imaging

Dissertation zur Erlangung
des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich Psychologie
und Sportwissenschaften
der Goethe-Universität
in Frankfurt am Main

von
Jutta Mayer
aus Mainz

Frankfurt am Main, 2008

(D30)

vom Fachbereich Psychologie und Sportwissenschaften der Goethe-Universität als Dissertation angenommen.

Dekan:	Prof. Dr. H. Moosbrugger
Gutachter:	Prof. Dr. R. Sireteanu Prof. Dr. Dr. D. Linden

Datum der Disputation: 03.12.2008

Danksagung

Die vorliegende Arbeit entstand im Labor für Klinische Neurophysiologie und Neuroimaging in der Klinik für Psychiatrie, Psychosomatik und Psychotherapie am Universitätsklinikum Frankfurt am Main. Viele Menschen haben direkt oder indirekt zur erfolgreichen Durchführung der Untersuchungen beigetragen. Ihnen möchte ich an dieser Stelle meinen Dank aussprechen.

Danken möchte ich zunächst Frau Prof. Ruxandra Sireteanu für ihre wertvollen Anregungen und ihre umfassende Unterstützung bei der Umsetzung meines Dissertationsvorhabens. Mein ganz besonderer Dank gilt Herrn Prof. David Linden, der den Anstoß für die vorliegende Arbeit gab und ihr Entstehen durch seine Ideen und durch fruchtbare Diskussionen in allen Phasen hilfreich begleitete. Bedanken möchte ich mich auch bei den zuständigen Direktoren, Prof. Konrad Maurer und Prof. Wolf Singer, für die Bereitstellung der notwendigen Ressourcen für das Tandem-Projekt zwischen dem Labor für Klinische Neurophysiologie und Neuroimaging und dem Max-Planck-Institut für Hirnforschung. Mein Dank gilt auch der Studienstiftung des deutschen Volkes für die finanzielle Unterstützung.

Die Ergebnisse der ersten fMRT-Studie wurden 2007 in der Zeitschrift Neuroimage veröffentlicht. Die Ergebnisse der Verhaltensstudie wurden 2007 in der Zeitschrift Advances in Cognitive Psychology veröffentlicht. Für ihre tatkräftige Unterstützung auf dem Weg dorthin möchte ich mich bei allen Koautoren, namentlich Prof. David Linden, Dr. Danko Nikolić, Robert Bittner, Dr. Christoph Bledowski und Prof. Rainer Goebel herzlich bedanken. Mein ganz besonderer Dank gilt Dr. Danko Nikolić für seine hervorragende fachliche Unterstützung ohne die, die Realisierung der Verhaltensstudie nicht möglich gewesen wäre.

Die zweite fMRT-Studie wurde unter Mitarbeit von Fabian Fußer durchgeführt, der im Rahmen seiner Doktorarbeit im Fach Medizin, einen wesentlichen Beitrag zur Erhebung und Auswertung der Pilotdaten (Verhaltensexperiment 5) und insbesondere der fMRT-Daten geleistet hat. Ihm gilt mein Dank für die engagierte und kooperative Zusammenarbeit.

Darüber hinaus möchte ich mich bei allen Kollegen und Kolleginnen des Labors für Klinische Neurophysiologie und Neuroimaging dafür bedanken, dass sie mich durch ihre Hilfsbereitschaft in vielfältiger Weise unterstützt haben. Ganz besonders möchte ich der Leiterin Dr. Corinna Haenschel für ihre wertvolle wissenschaftliche und persönliche Unterstützung danken.

Mein ganz besonderer Dank richtet sich jedoch an meine Eltern, Evi und Klaus Mayer, für ihr großes Interesse an meiner Arbeit und die Selbstverständlichkeit, mit der sie mich bei dem was ich tue unterstützen.

Table of Contents

Chapter 1 - Introduction

Visual working memory (WM) and selective attention are fundamental cognitive mechanisms both operating at the interface between perception and action. They are related because both are concerned with the control of information, and both are postulated to have limits with respect to how much information can be processed. For instance, while looking on a map and following the route between two locations, it might be extremely difficult to search for a second destination at the same time. Subjectively, giving attention to one part of the map leaves less attention available to another part. Similarly, keeping in mind the turns, significant landmarks and other information needed to reach the destination might not be perfect, and one might need to search again when losing the relevant information. The longer or the more complex the route is, the higher the risk to forget the previously stored information while accomplishing the search.

The idea that attention and memory are closely intertwined has already been noted in 1759 by the English poet Samuel Johnson (Bate et al., 1963): "*The true art of memory is the art of attention.* No man will read with much advantage, who is not able, at pleasure, to evacuate his mind, or who brings not to his author an intellect defecated and pure, neither turbid with care, nor agitated by pleasure. If the repositories of thought are already full, what can they receive? If the mind is employed on the past or future, the book will be held before the eyes in vain. What is read with delight is commonly retained, because pleasure always secures attention; but the books which are consulted by occasional necessity, and perused with impatience, seldom leave any traces on the mind."

However, throughout the modern history of cognitive psychology and cognitive neuroscience, research has focused on these processes as separate topics. Attention studies have typically ignored the importance of perceptual experience and past knowledge, whereas memory studies have not explored the role of attentional selection during the encoding and short-term storage of information. Therefore, behavioural models of how the mechanisms that select the relevant information from the incoming stimulus stream and maintain and manipulate information over short periods of time, interact with one another, have been described only recently, whereas neurophysiological models still lack. Given that attention and WM are intimately linked, specifying when and how these mechanisms interact, is crucial to increase the understanding of the cognitive and neural mechanisms underlying each of these abilities.

This dissertation aims at directly investigating interactions between visual WM and selective attention. Specifically, interactions are investigated with respect to one characteristic feature of WM and attention, namely their limitation in capacity. Although the existence of processing limitations in attention and WM has been largely described, little is known about what actually causes these limitations.

The present dissertation is embedded within the cognitive neuroscience approach, specifically devoted to understanding brain-mind relationships. Cognitive neuroscience integrates the conceptual models and methodological strategies provided by cognitive and experimental psychology with functional imaging techniques. Particularly, the development of functional magnetic resonance imaging (fMRI) in the early 1990s represents a landmark in cognitive neuroscience as this non-invasive technique allows neuroscientists to examine the neural substrates of cognitive constructs through imaging the brain in awake, behaving, human subjects as they perform a cognitive task. In this dissertation fMRI is used to study interactions between visual WM and attention in terms of BOLD (blood oxygen level dependent) activation. Two fMRI experiments in which the demands on WM and attention were manipulated orthogonally within one unitary paradigm are presented in this dissertation. The fMRI experiments were preceded by a behavioural study using the same stimuli. The aim of this study was i) to validate the engagement of the relevant attentional and WM processes by the chosen task manipulations and ii) to investigate whether and how participants can cope with the common capacity limitations of visual WM and attention These behavioural experiments will be presented in the first part of this dissertation.

In the introduction part I will first describe the psychological concepts of visual attention and visual WM in more detail. Behavioural studies investigating interactions between these mechanisms will be introduced and discussed in the context of models of capacity constraints of visual WM. Then, I will give an overview of the neural bases of visual attention and visual WM. Again, interactions will be discussed in the context of the neural capacity constraints of visual WM. The introductory chapter closes with an outline of the studies presented in this dissertation, their objectives and specific hypotheses.

1.1 The concept of visual selective attention

Capacity limitations in visual perception and selective mechanisms have been a central topic in cognitive and experimental psychology over the past 50 years. Using diverse experimental paradigms such as selective looking, dual-task, visual search, cuing, the psychological refractory period, and the attentional blink, the basic properties of attention, the mechanism that accomplishes selection, have been extensively explored (Pashler, 1998). Visual attention increases processing efficiency of relevant stimuli and reduces the interference from irrelevant distractors. The rate at which visual information can be attended is severely limited (Duncan et al., 1994), as is the number of objects that can be simultaneously attended among distractors. In a typical multiple object visual tracking task, subjects are shown a display of identical items that move within a rectangular area. At the beginning of each trial a subset of these items are marked briefly to assign them as targets. All the items then start to move independently and randomly within the display and subjects are instructed to keep track of the targets. After a period of time the items stop moving and subjects have to point out which items were the targets. Pylyshyn and Storm (1988) found that subjects were able to successfully track four to five items. Further studies have confirmed this limit of the attentional capacity (Scholl, 2001; Cavanagh and Alvarez, 2005).

Attention is highly flexible and can be deployed to locations, visual features, or objects. It can be driven exogenously, by an external stimulus event that automatically captures attention (“bottom-up”) or endogenously, by factors such as knowledge, expectation, or current goals of the subject (“top-down”). Furthermore, sustained attention can be distinguished from shifts of attention that can be accomplished with and without the concurrent performance of eye-movements (overt vs. covert attention). Finally, central capacity limitations have been distinguished from purely perceptual (for a review see Pashler, 1998; Chun and Wolfe, 2001). Taken together, attention does not denote a singular concept, but stands for a variety of psychological phenomena. Integrating those within a unitary theoretical framework is still a great challenge in the field of attention.

1.1.1 Traditional theories of selective attention

Capacity limits of information processing have been traditionally interpreted in terms of processing bottlenecks that occur if several stimuli act upon a single limited-capacity channel (Broadbent, 1957). Therefore, the system has to be proceeded by a filter mechanism that selects a portion of the incoming information for more elaborate processing. As a result stimuli are processed serially, with one stimulus after the other. Whether the selection mechanism works early or late in processing has been a long standing debate. Broadbent (1958) advocated filtering of irrelevant sensory information based on physical attributes such as location or loudness ("early selection"). The alternative, late selection view holds that selection occurs only after semantic analysis of all input has occurred (Deutsch and Deutsch, 1963; Duncan, 1980). Intermediate views include attenuation theory which proposes that rejected information is attenuated rather than completely filtered or completely identified (Treisman, 1960). An alternative account with elements of compromise between early and late selection is graded capacity sharing. Capacity or resource models (Kahneman, 1973) argue that perceptual processing takes limited mental resources. These finite resources can be allocated to different stimuli in parallel but if the capacity limits are exceeded processing of each stimulus becomes less efficient. Thus, allocation of mental resources rather than a filter mechanism determines which stimuli are processed (for a review see Pashler, 1998).

1.1.2 Visual search and attention

The visual search paradigm is known as one of the dominant methods that have been used to examine the efficiency with which observers can deploy attention to the relevant aspects of a scene (for a review see Wolfe, 1998a). In a standard visual search task, subjects look for a target item among a varying number of distractor items. The total number of items in the display is known as the set size. The target is presented for some percentage of trials, typically 50%. Subjects press one button if the target is present and another button if only distractors appear. The display remains visible until the subject responds and reaction time (RT) and accuracy are measured. Accuracy is usually high and RT is analysed as a function of set size, producing two functions - one for target present and one for target absent trials. The slopes and the intercepts of these RT x set size functions are used to quantify search efficiency and to draw inferences about the underlying search processes.

It has been shown that the efficiency of search tasks varies in a systematic way with the nature of the search stimuli. For some tasks, performance does not depend on set size. For example, in a search for a red X among green distractors, the number of green items is irrelevant. Accuracy is high and RT fast for all set sizes. The resulting RT x set size slopes are near zero ms/item indicating that the target item, when present, is detected easily without interference from the distractor items. In other words, the red item "pops out" and makes its presence known (see Figure 1.1, left panel). For other tasks, RT is roughly a linear function of set size. For example, in a search for a target defined by conjunctions of two feature dimensions such as a red X among green Xs and red Os, RTs typically increase at a rate of approximately 20 to 30 ms/item for target-present trials and 40 to 60 ms/item for target-absent trials (Figure 1.1, centre). Steeper search slopes indicate less efficient search and a greater cost for each additional distractor.

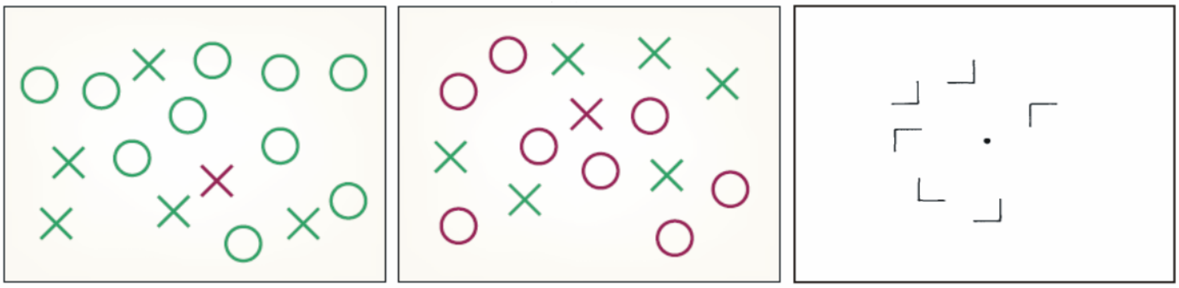


Figure 1.1 Examples of displays used in visual search studies (from Duncan and Humphreys, 1989; Robertson, 2003). Left panel. Feature search. The target is the red X among green distractors. Search in this case is highly efficient. The red target seems to pop out of the display. Centre. Conjunction search. The red X is more difficult to detect among green Xs and red Os. Search is less efficient, that is the number of distractors strongly affects RT. Right panel. Spatial-configuration search. Detecting the upright L is inefficient as it is presented among similar and heterogeneous distractors.

What determines the efficiency of visual search? Treisman's Feature Integration Theory (Treisman and Gelade, 1980; Treisman and Sato, 1990) is an early and influential account proposing that efficient and inefficient visual search differ qualitatively with regard to the underlying processes. According to this theory efficient feature search is assumed to be preattentive in nature, occurring in parallel across the visual field. In contrast, inefficient conjunction search produces non-flat search functions because it requires serial deployment of spatial attention to the individual items in turn in order to allow correct binding of the constituent features of an object.

However, zero-search slopes have been demonstrated also with conjunctions and conversely, feature searches can produce steep search functions when differences between targets and distractors are sufficiently small (Wolfe, 1998a). For instance, a search for a upright L among Ls rotated 90° clockwise or counter clockwise from the target position produces a target present slope of 38 ms/item and a target absent slope of about 71 ms/item (Duncan and Humphreys, 1989) (Figure 1.1, right panel). This indicates that feature and conjunction searches might be distinguished quantitatively rather than qualitatively and led to the idea of a continuum of search tasks from highly efficient to inefficient (Duncan and Humphreys, 1989; Wolfe, 1998a). This notion has been incorporated in parallel models that assume that in all kinds of search tasks the items are processed at once (Duncan and Humphreys, 1989). The simultaneous analysis of the whole array becomes less efficient, and thus slows down, as a finite processing capacity is approached. According to the Attentional Engagement Theory (Duncan and Humphreys, 1989), the parallel, competitive mechanism that is involved in both search for features and conjunctions, is based on mutual inhibitory interactions among units activated by the various elements in the array.

Most important for the present dissertation, irrespectively of the assumed model of visual search, the literature indicates that the degree of attentional deployment required for target-distractor discrimination can be systematically varied by manipulating search efficiency. Evidence suggests that search efficiency increases with decreasing target-distractor similarity and/ or increasing distractor homogeneity (Treisman and Gormican, 1988; Duncan and Humphreys, 1989).

1.2 The concept of WM

WM research has grown out of the research on short-term memory (STM) that has been described in detail within the famous model of human memory given by Atkinson and Shiffrin (1968). According to their account, information from the environment flows through a series of temporary sensory registers into a limited-capacity short-term store, which feeds information into and out of long-term memory (LTM) that has unlimited capacity. STM was conceptualised as a unitary system that holds a small amount of information for about 15 to 30 s after which it is lost due to decay or interference.

The term WM was used first by Miller and colleagues (1960) for describing a process that maintains behavioural plans and goals in an active state in order to be able to modify and judge current actions against them. The theoretical concept of WM was brought to the forefront in the field of memory with the model of WM first published in 1974 by Baddeley and Hitch. In contrast to the traditional storage-oriented notion of STM (Atkinson and Shiffrin, 1968), WM was considered a more processing-oriented construct supporting the performance of complex cognitive tasks, such as learning, comprehension, and reasoning. In place of a unitary short-term store, Baddeley and Hitch (1974; Baddeley, 1986) postulated a three-component model of WM. The three components comprise a control system of limited attentional capacity, termed the central executive, which is assisted by two subsidiary slave systems, specialised for the temporary storage and manipulation of visuospatial and phonologically based material (the visuospatial sketchpad and the phonological loop, respectively). The phonological loop has been further fractionated into a passive phonological store and an active rehearsal process. It represents material in a phonological code, which decays with time, whereas the rehearsal process serves to refresh the decaying representations in the phonological store. Similarly, the visuospatial sketchpad is the storage system for visual material, defined by its main function to serve as an on-line “cache” for visuospatial material and the ability to actively rehearse the contents of WM (“inner scribe”, Logie, 1995). Baddeley (2000) has recently revised this model, postulating a fourth subsystem, the episodic buffer, which forms an interface between the phonological loop, the visual sketchpad, and LTM. It is supposed to held integrated material such as scenes and events in a multimodal code (Figure 1.2).

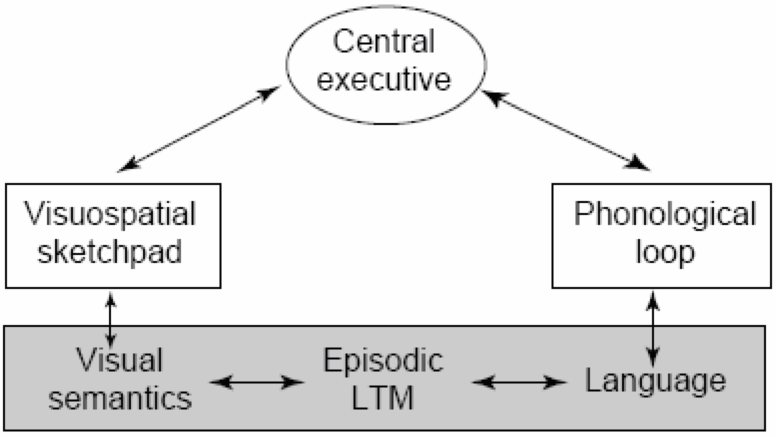


Figure 1.2 The multi-component model of WM (Baddeley, 2000).

Since its initial development in the 1970s, the multi-component WM model has initiated a great extent of research. It was a basis on which many predictions have been made and tested and has stood as the golden standard for many years. It has been especially popular in research on language processing. The visual component, however, has been proved harder to investigate. Still missing, until today, is the attempt to detail the characteristics of the visual store, paralleling the work on the articulatory loop.

1.2.1 Visual WM

The fundamental characteristics of visual WM have been investigated since the pioneering studies on short-term retention of visual stimuli conducted by Phillips in the 1970s. Phillips (1974) asked subjects to compare random patterns of black and white square matrices successively presented with variable retention intervals. With retention intervals of less than 100 ms memory performance for patterns was close to perfect but declined when increasing the duration of the retention interval at 1 s or higher. In a series of this type of experiments it was shown that visual STM is different from iconic memory (Sperling, 1960) in that it has a limited capacity, it is dependent on pattern complexity, it lasts for at least 15 s, but is lost very soon after the onset of interference, it is not maskable, and not tied to spatial position (Phillips, 1974). Also, visual STM could be distinguished from visual LTM with regard to capacity and durability (Phillips and Christie, 1977).

Around the same time research on visual WM in the context of the model proposed by Baddeley and Hitch (1974) focussed on the question of dissociable stores for verbal and visual material. Selective interference effects found in normal adults in dual-task paradigms (Baddeley, 1986; Logie, 1995), the patterns of selective impairments in brain-damaged patients (Della Sala and Logie, 1993) as well as developmental studies (Hitch, 1990) strongly supported such a separation. Later studies largely indicated that the visual sketchpad itself could not be regarded as a unique, homogeneous system, but rather as comprising different and at least partially independent subcomponents. Tresch et al. (1993), for example, demonstrated that retention of spatial patterns was impaired by a concurrent movement discrimination task but not a colour discrimination task, whereas retention of object information showed the opposite pattern of interference. A double dissociation between visual and spatial span was also revealed by Della Sala et al. (1999). For normal subjects, performance in the Corsi-block test (spatial component) was reduced by the addition of a secondary spatial tapping task, but not an irrelevant pictures task. The reverse was true for the visual pattern task (visual component). Also, neuropsychological cases could be identified, showing either disruption of spatial but not visual WM, or the opposite pattern (Della Sala and Logie, 2002). Recently, it has been demonstrated that the segregation of visual WM according to the type of information applies also to active WM processes such as the manipulation and integration of information held in WM that require executive control (Mohr and Linden, 2005).

1.2.2 Capacity constraints of visual WM

One of the hallmark characteristics of WM is its severe capacity limitation. William James (1890) already stated that, unlike the virtually unlimited amount of knowledge that can be stored in a person’s secondary memory, only a small amount of information can be kept conscious at any one time in one’s primary memory. This limitation was described in terms of the absolute amount of information that can be maintained. Undoubtedly, the best-known estimation of WM capacity has been George Miller’s proposal of a “magic number 7 plus or minus 2” (Miller, 1956). This capacity estimation was based on the observation that subjects were able to repeat verbatim about seven items in immediate-recall procedures. However, that number was meant more as a rough estimate and a rhetorical device than as a real capacity limit (Cowan, 2001, 2005). A more central focus of this seminal paper was the ability to increase the effective storage capacity through the grouping of information into higher-order chunks. The present stance is that the number seven estimates a commonly obtained, compound capacity limit when the number of chunks is unclear, rather than a pure capacity limit in which chunking has been eliminated. Evidence derived from procedures that circumvented supplementary mechanisms of memory such as grouping, rehearsal, and sensory memory, suggests a pure WM capacity in adults of three to five chunks (“the magical number 4 ± 1”, Cowan, 2001). Individual scores appear to range more widely from about two up to about six chunks.

It has been shown that this limit applies not only to the verbal but also to the visual component of WM (Sperling, 1960; Pashler, 1988; Cowan, 2001; Vogel et al., 2001). In one influential study Luck and Vogel (1997) measured the capacity of WM for simple features using a variant of the change detection paradigm developed by Phillips (1974). A sample and a test array containing a varying number of visual objects were presented in close succession and subjects indicated whether they were identical or differed in terms of a single feature such as colour or orientation. Performance, assessed as a function of set size, was nearly perfect for arrays of 1 to 3 items and declined systematically as the set size increased from 4 to 12 items. Capacity estimation indicated that observers were able to retain information about only four colours or orientations in visual WM at one time. Importantly, similar results were found when they measured WM capacity for objects defined by a conjunction of two or even four features indicating that the 4-item limit applied to integrated objects rather than individual features within objects. However, subsequent findings (Alvarez and Cavanagh, 2004) suggested that the complexity of objects or the number of their features also contribute to the capacity of WM. In this study, WM capacity for objects from different classes of stimuli (colours, polygons, Chinese characters, shaded cubes, and letters) was estimated within a change detection paradigm. It turned out that the greater the complexity of each item in a stimulus class was, the fewer objects from that class subjects were able to hold in memory, with the estimates ranging from 1.6 for cubes to 4.4 for colours. The upper bound on capacity was of approximately four or five objects. Thus, both the total information load and the number of objects imposed capacity limits on visual WM.

Also related to the discussion on the capacity constrains of visual WM is the phenomenon of change blindness whereby prominent objects in scenes can disappear, change colour, or move between one display and the next without people noticing this change (Rensink, 2002; Simons and Rensink, 2005). Therefore, it has been inferred that only little information from our visual environment is consciously perceived and stored in visual STM (Rensink, 2002). Importantly, it appears that people can monitor just between one and four items for a change (Rensink, 2000) which corresponds well to the estimates of the capacity of visual WM.

1.3 Interactions between WM and attention

1.3.1 WM and attention as separate cognitive capacities

Traditional models of human information processing characterised attention as a filtering mechanism that limits the amount of information entering a memory store (Broadbent, 1958; Atkinson and Shiffrin, 1968). In these early theories temporary memory and attention were considered distinct, associated with separate functions. Broadbent's influential Filter Theory (Broadbent, 1958) was developed as an ordered series of memory stages in which information was first held in a sensory store of unlimited capacity. From this store, some information was selected for further processing by passing a selective-attention device or filter into a limited-capacity short-term store. Here information was fully perceived and available for further processing including long-term storage of past events. The multi-store model was made more explicit by Atkinson and Shiffrin's (1968) emphasis on control processes that manage the transfer of information between sensory, short-term, and long-term stores under the subject's effortful, voluntary control. The short-term store was assumed to be capable of utilizing a range of control processes with attention controlling the transfer from sensory registers to the short-term store. Rehearsal and coding processes were thought to operate to maintain relevant information in the short-term store and to store it in the long-term store.

The operation of an attentional gate in visual WM has been supported by recent findings (Duncan and Humphreys, 1989; Bundesen, 1990). For instance, Schmidt et al. (2002) combined a spatial cuing with a visual WM task for colours. They found that colours were more likely to be remembered when they were presented at cued vs. uncued locations, and this was the case even when the cues did not predict which colour would be tested. Thus, focusing attention onto a spatial location increased the probability that information at that location was transferred into visual WM. In addition, recent models of WM that view attention and WM as serving separate functions exist as well. One example is the computational model proposed by Schneider (1999) that simulates cognitive processing in a hierarchical network of connectionist modules. Within this architecture, WM is stored in activation patterns across modules and in short-term connection changes, and supports the maintenance and the association of information. In this model, attention generally gates information through the network and monitors the activity of the resulting transmissions.

Taken together, these early and recent theories of information processing bear on the idea that attention and WM are distinct mechanisms that work at different stages of processing, with attention taking place earlier and controlling which sensory information gets encoded into visual (short-term) WM.

1.3.2 WM and attention as different aspects of the same cognitive capacity

1.3.2.1 WM and executive attentional processes

A key role of WM is to enable higher level cognitive functions that require a rapidly accessible and easily updated memory system. Executive attentional processes participate in the active manipulation and updating of contents in WM. The multi component model of WM proposed by Baddeley and Hitch (1974) incorporates such higher level control processes by postulating a central executive system. In the original model, this component was simply treated as a pool of general processing capacity that could be used to support either control or storage processes. This concept was further advanced by adopting the model of attentional control proposed by Norman and Shallice (1986) which made a distinction between automatic, habitual control and attentional, supervisory control (the supervisory activating system). Today the central, overarching executive system within the Baddeley and Hitch model is considered as reflecting a range of separable processes such as focussing, dividing, and switching attention required for the integration of information and the control of action which should be necessary, at least minimally, in all WM tasks. As noted by Baddeley (1993), from an attentional viewpoint it would have been appropriate to use the term working attention rather than working memory pointing to a substantial overlap between the concepts of attentional control and WM.

1.3.2.2 Visual WM and selective attention

It is a defining characteristic of visual WM to be limited in capacity. As mentioned above, up to four objects can be held in mind. A similar number of objects can be attentively tracked. This similarity in the capacity limitations of visual WM and selective attention has led researchers to suggest that visual WM and selective attention rely on a common capacity-limited mechanism (Awh and Jonides, 2001; Cowan, 2001; Rensink, 2002; Wheeler and Treisman, 2002; Cowan, 2005; Fougnie and Marois, 2006).

1.3.2.2.1 The embedded-processes model of WM

This idea has been made explicit within the embedded-processes model of WM proposed by Cowan (1988, 1993, 2005). Cowan offers the view that the contents of WM are best understood as activated representations from within LTM that are currently within the focus of attention (Figure 1.3). According to this model a stimulus that is presented to the subject first enters a sensory store that preserves its physical properties for a period of up to several hundred milliseconds. During this time, information in LTM has started to become activated producing stimulus coding and STM storage of the activated set of codes from LTM. Thus, LTM represents the source of activated memory. However, activated memory also contributes to the formation of long-term memories. Most stimulus situations in life include novel combinations of familiar features. New links are formed between the elements that are concurrently activated in memory and may then become part of LTM. Activated codes corresponding to stimuli to which the subject has habituated remain outside awareness, i.e. they do not enter the focus of attention. The focus of attention is thought to be controlled by a combination of automatic orienting responses to changes in the environment and voluntary effort arising from central executive processes. Most importantly, whereas activated memory is suspected of having limits caused by interference from incoming similar items and/ or from memory decay over time, it is the focus of attention that is limited by its capacity (rather than time) to about three to five separate chunks of information at any given time (Cowan, 1998, 2001). By this view, attention is the limited-capacity process that constrains WM capacity.

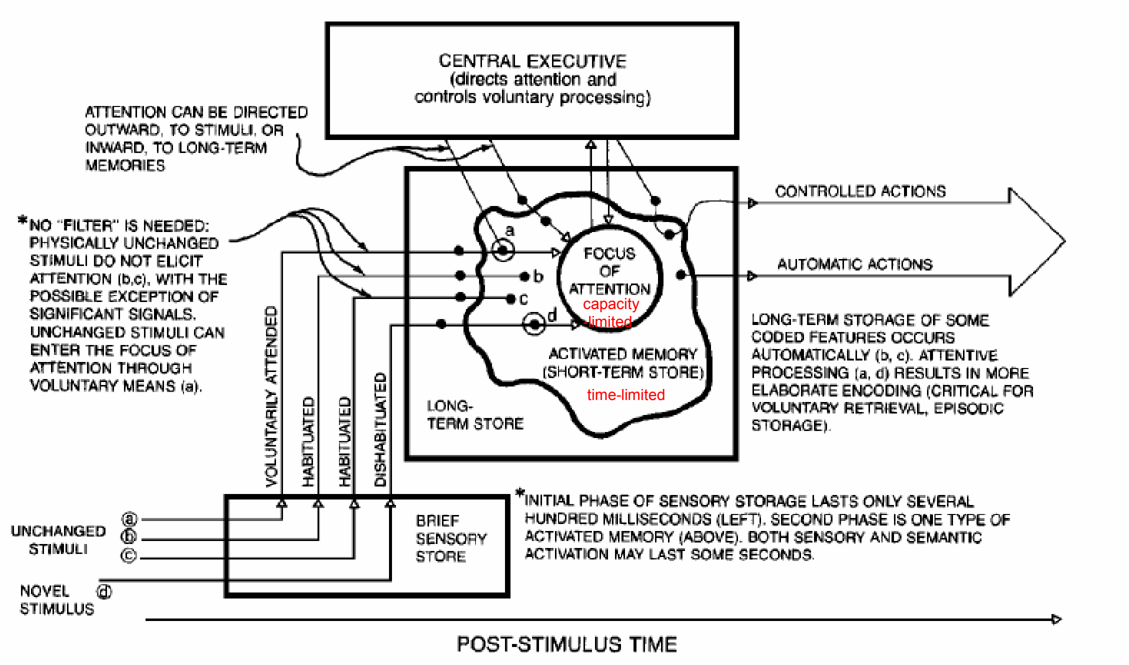


Figure 1.3 The embedded-processes model of WM (modified from Cowan, 1988).

1.3.2.2.2 Binding in WM

A different account of how selective attention contributes to the limited capacity of visual WM has been offered by Wheeler and Treisman (2002). They hypothesised that attention is required to maintain bound information in WM in a similar vein as it is required for creating bindings between object features in visual perception (Treisman and Gelade, 1980; Treisman and Gormican, 1988). This hypothesis was tested in a series of experiments using a change detection paradigm similar to that used in the study by Luck and Vogel (1997). However, they failed to replicate Luck and Vogel's finding that subjects could memorise two values within the same dimension (bicoloured squares) as easily as they could memorise objects with a single value on that dimension (single-coloured squares). Therefore, the authors reasoned that memory capacity within a feature dimension is limited by the number of feature values rather than by the number of objects. Furthermore, Wheeler and Treisman (2002) tested WM for binding between different feature dimensions such as colour and location and found that performance was significantly worse compared to the memory for only one or either one of both features. In addition, memory for binding was selectively impaired by the perceptual condition at test, with significant worse performance when presenting a whole display test vs. a single probe. Based on these results they proposed separate mechanisms that limit visual STM for features on the one hand and for binding between these features on the other. According to their model, feature values from different dimensions are stored in parallel in separate dimension-specific caches each with its own capacity limit. Within a dimension the features compete for limited capacity representation which is typically about three to four items, but between dimensions there is little or no competition. Maintaining binding information costs only little in terms of feature capacity; however, it depends on another more general attention resource and is more vulnerable to interference. In conclusion, the capacity limit of visual WM is viewed as a product of the interaction between limited-capacity attentional processes needed to integrate information from different dimensions and the independent capacity of distinct feature stores.

1.3.2.2.3 Attention-based rehearsal in WM

Evidence for the proposal that visual WM and attention rely on a common limited-capacity process also comes from studies demonstrating that visual WM and attention can interfere with each other. In a study conducted by Smyth and Scholey (1994) subjects were asked to remember the temporal order of a set of locations within a pre-defined spatial array while performing various secondary tasks during the retention period (e.g., touching visual targets, repeating heard words, listening to tones from spatially separated locations, pointing to these tones, pointing to visual targets, and categorising spatial targets as being from the left or right). Serial spatial WM was selectively impaired by those tasks that required shifts of spatial attention away from the memorised locations (touching visual targets, listening to tones from spatially separated locations, pointing to these tones, pointing to visual targets, and categorising spatial targets as being from the left or right). The authors concluded that covert shifts of spatial attention could aid in the active maintenance information in spatial WM, in much the same way that covert articulation serves to refresh the decaying representations in the phonological loop (Baddeley and Hitch, 1974). Likewise, Awh and colleagues (1998) investigated how shifts of spatial attention affect information previously stored in spatial WM. Participants performed a colour discrimination task during the retention interval of a memory task for a single location. Maintenance of information in spatial WM was incompatible with the secondary discrimination task when this task required participants to shift their attention to non-memorised locations. In contrast, when colour could be discriminated without a shift of attention, WM was not impaired. The authors proposed that mechanisms of spatial attention were recruited in the service of a rehearsal-like function in order to maintain information active in spatial WM. Importantly, the observed interference effects suggest that shifts of attention to locations represented in WM were a necessary part of accurate WM storage pointing to a functional overlap in the mechanisms of spatial WM and spatial selective attention.

Whether a similar relationship between attention and WM is at work in the object domain is still an open question. Some evidence comes from studies demonstrating attentional capture by objects held in WM. For example, Downing (2000) asked participants to encode a sample object into WM (a face or a line drawing) which was followed by the presentation of two objects during a delay period, one matching the sample and the other novel. When a secondary probe stimulus appeared at the location at which the originally encoded object was presented immediately before, reaction times to the probe stimulus were faster than when the probe appeared at the location of the new item. Thus, maintaining an object in WM increased the probability that attention was drawn towards the object held in memory rather than the novel one. However, clear evidence that the storage of an object in WM necessarily leads to attentional capture by subsequent presentations of that object has not yet been provided (Downing and Dodds, 2004). So far, a true functional role of selective attention in WM maintenance could be demonstrated only in the spatial domain.

1.3.2.2.4 WM and visual search

Interference between visual WM and attention has also been demonstrated in dual-task studies that tested the influence of concurrent WM loads on performance in visual search tasks. In two independent studies (Oh and Kim, 2004; Woodman and Luck, 2004) subjects were asked to perform a visual search task during the delay of a WM task for spatial locations. In both cases, spatial WM load impaired search efficiency, and the search process impaired spatial WM accuracy in the dual task condition as compared with when the two tasks were tested in isolation. These results support the assumption that selective attention and spatial WM storage rely on a common limited-capacity process. In contrast, interactions between visual search and the storage of information in object WM have not been consistently found. Performing a visual search task while maintaining colours in visual WM did not result in impaired search efficiency and impaired memory accuracy in the study by Oh and Kim (2004). Similar results were reported by Woodman et al. (2001) when combining a visual search task with WM tasks for either colours or objects. However, interference between object WM and visual search has been demonstrated when the target for the search was not consistent throughout the experimental session but changed on a trial-by-trial basis (Awh et al., 2006).

In a different line of research, Lavie and collegues have shown that visual selective attention is sensitive to interference from WM requirements in conditions of high memory load (de Fockert et al., 2001; Lavie et al., 2004; Lavie and de Fockert, 2005). For instance, Lavie and de Fockert (2005) demonstrated that WM load specifically affects attentional capture by a salient but irrelevant colour singleton in visual search. That is, when subjects needed to maintain information in verbal WM while performing the visual search, attentional capture by the distractor increased. These results converged with previous findings of a study that combined neuroimaging and behavioural experiments (de Fockert et al., 2004). In this study subjects memorised a digit order while performing a selective attention task that required them to classify written famous names and to ignore irrelevant distractor faces. Greater interference on RTs were observed from incongruent distractors (e.g., Bill Clinton’s face with Mick Jagger’s name) versus neutral or congruent distractors under high vs. low WM load. Moreover, the neuroimaging results showed increased face-related activity in the visual cortex related to the presence vs. absence of distractor faces under conditions of high vs. low WM load. These results provide support for the hypothesis that WM serves to maintain the distinction between relevant and irrelevant stimuli indicating a role of WM in the control of selective attention (Chun and Turk-Brown, 2007).

1.4 The neural bases of visual selective attention and visual WM

Insights into the neural mechanisms subserving visual selective attention and visual WM have been derived from extensive research starting with single-cell recordings and lesion studies in monkeys, and the investigation of patients suffering from attentional and memory deficits due to brain damage. In addition, studies using functional imaging techniques such as electroencephalography (EEG), positron emission tomography (PET), and fMRI have brought important contributions to the fields of attention and WM. Following the tradition of investigating selective attention and WM in isolation, findings on their neural substrates will be presented separately in the first part. So far, little is known from targeted comparisons and results from these studies will be described in the second part. The section will be closed with recent findings on neural capacity constraints of visual WM, which provides the framework within interactions between visual WM and attention are addressed in this dissertation.

1.4.1 Neural correlates of visual selective attention

The neural basis of visual attention has been investigated with regard to two fundamental aspects of this mechanism: i) the effects of attentional modulation on neural processing in the visual cortex and ii) the top-down control of these modulations by source areas in the parietal and frontal cortex.

1.4.1.1 The effects of attention on visual processing

Evidence from single-unit studies in monkeys indicates that attention increases activity in areas of the brain that are specialised for the processing of stimuli at attended locations or of attributes of attended stimuli, such as colour, motion, texture, or shape (for a review see Kanwisher and Wojciulik, 2000; Kastner and Ungerleider, 2000; Pessoa et al., 2003). In a typical experiment, an attended condition, in which the monkey focuses attention on a visual stimulus that is placed within the receptive field (RF) of a cell, is compared with an unattended condition, in which the same visual stimulation conditions are present but the monkey focuses attention on a stimulus outside the RF. The paradigmatic finding is that when attention is directed to a single stimulus in the RF, there is an increase in the firing rate of neurons that respond to the attended stimulus (Motter, 1993). These attentional effects are retinotopically organised, that is the cortical topography of purely attention-driven activity precisely matches the topography of activity evoked by visual targets. In a similar fashion, increases in fMRI signals in humans have been reported for a stimulus at an attended relative to an unattended location as well as in favour of an attended stimulus attribute. For instance, attention to shape and colour leads to response enhancement in regions of the posterior portion of the fusiform gyrus, including area V4. The physiology literature has reported attention-related modulations in many extrastriate cortical areas, including V2, V4, temporal-occipital area (TEO), and middle temporal area (MT). Relatively few reports suggest that attentional modulation occurs in the primary visual cortex (V1). In sharp contrast, fMRI studies have demonstrated robust effects of attention also on V1 responses, probably reflecting longer-latency feedback processes from other areas (Martinez et al., 1999).

Visual attention does not only enhance the neural representation of the attended stimulus but can also inhibit the representation of the unattended stimulus (Kanwisher and Wojciulik, 2000; Kastner and Ungerleider, 2000; Pessoa et al., 2003). In the absence of attention the neuronal response to a single effective stimulus in V4 is reduced when an additional, ineffective stimulus is present in the same RF (Reynolds et al., 1999). However, attention can resolve the competition among multiple stimuli by counteracting the suppressive influences of nearby stimuli (stimuli falling within the same RF) in higher-level visual areas, thereby enhancing information processing at the attended location (“Biased Competition Model”, Desimone and Duncan, 1995; Desimone, 1998).

A further effect of visual attention on visual processing has been termed the baseline shift. That is, in expectation of a visual stimulus but before it is presented, the spontaneous firing rates for neurons or populations of neurons in the retinotopically appropriate region within the visual cortex are increased by a constant amount independent of the strength of the stimulus (Luck et al., 1997; Kastner et al., 1999). This increase of baseline activity during the expectation period was followed by a further increase of activity evoked by the onset of the stimulus presentations. Such a shift in baseline activity in visual cortex in the absence of visual stimulation presumably increases sensitivity to a stimulus at a given location or to the stimulus feature, thereby providing a competitive advantage.

1.4.1.2 The control of visual attention

The frontal and parietal lobes have been implicated in the direction of visual attention, on the basis of patient studies demonstrating that damage in these regions leads to attentional deficits such as neglect (Posner and Petersen, 1990), and single-unit studies which show that neurons in these areas produce stronger responses to attended than unattended stimuli (Goldberg et al., 2002). Furthermore, there exists an anatomical substrate for top-down influences, inasmuch as tract-tracing studies in monkeys have demonstrated direct feedback projections to extrastriate visual areas V4 and TEO from parietal cortex and to inferior temporal (IT) cortex from prefrontal cortex, as well as indirect feedback projections to areas V4 and TEO from prefrontal cortex via parietal cortex (Pessoa et al., 2003).

Results from fMRI studies in humans further support the idea that areas in the frontal and parietal cortex are involved in the generation and control of attentional top-down signals. In a typical visuospatial attention task, subjects are asked to attend to a central cue and, based on the nature of the cue, covertly (without making eye movements) direct their attention to a peripheral visual stimulus for target detection or discrimination (e.g., Corbetta et al., 1993; Nobre et al., 1997). A cortical activation pattern comprising the superior parietal lobule (SPL), the intraparietal sulcus (IPS), the frontal eye field (FEF), and the supplementary eye field (SEF) has been consistently found to be activated. In addition, but less consistently, activations in the inferior parietal lobule (IPL), the lateral prefrontal cortex in the region of the middle frontal gyrus (MFG) and superior frontal gyrus (SFG), the superior temporal gyrus (STG) and the anterior cingulate cortex (ACC) have also been observed (for a review see Kanwisher and Wojciulik, 2000; Kastner and Ungerleider, 2000; Pessoa et al., 2003).

The specific functions of frontal and parietal regions in visual attention could be further disentangled in a study conducted by Culham and colleagues (2001). They used a parametric manipulation of an attententive tracking task that required subjects to track one to five balls within a display of nine randomly moving balls. Visual, frontal, and parietal regions were activated in the attention-demanding task. A subset of these regions in frontal and parietal cortex showed a monotonic increase in activation from attention load 1 to 5 suggesting that these areas were directly involved in attentional processing. However, in other areas (e.g., FEF, parietal area 7) they found an increase in activation compared to a passive baseline condition with no additional increase when more items needed to be tracked indicating that these regions were involved in task-specific functions that supported overall performance. Such functions might include the preparation and suppression of eye movements (Corbetta et al., 1998).

Neuroimaging studies using the visual search paradigm to study the neural correlates of selective attention have reported activation in similar regions in the frontal (in particular FEF), posterior parietal (IPS, SPL) and occipital cortex (Corbetta et al., 1995; Corbetta and Shulman, 1998; Leonards et al., 2000; Donner et al., 2002; Nobre et al., 2003). Additional sites of activation include the cingulate gyrus, the superior colliculus, and the cerebellum (Gitelman et al., 2002). The fronto-parietal activation pattern revealed in these visual search studies appeared to be lateralised to the right hemisphere (Chelazzi, 1999). Although patient studies suggest a right parietal dominance in visuospatial attention as well, this has not been unequivocally supported by fMRI studies on spatial attention.

One major distinction has been made between endogenous and exogenous attention, two functions that are supposed to be subserved by partially segregated but interacting networks. According to the model proposed by Corbetta and Shulman (2002), a dorsal fronto-parietal system (bilateral SPL, IPS, and FEF) is involved in the generation of attentional sets associated with goal-directed stimulus-response selection. A second, ventral system supports the detection of behaviourally relevant stimuli and works as an alerting mechanism for the first system when these stimuli are detected outside the focus of processing. Some evidence suggests that the ventral system that includes the temporoparietal junction and the middle and inferior frontal gyri is lateralised to the right hemisphere (Corbetta and Shulman, 2002), whereas other studies did not support such hemispheric specialisation (Linden et al, 1999; Bledowski et al., 2004).

Recent evidence suggests that top-down control of attention to visual features draws on cortical regions that essentially overlap with those revealed by divers spatial attention tasks pointing to a general network related to the control of visual attention (Giesbrecht et al., 2003; Yantis and Serences, 2003; Serences et al., 2004). Finally, fMRI has been used to study the dynamics of attentional control. For instance, in a study by Yantis et al. (2002) subjects were asked to detect a digit that appeared in a stream of letters presented in rapid succession on the right or the left side of the display. Two different types of targets instructed them either to maintain attention on the same side or to switch attention on the other side. A rapid, transient increase in activation was observed in the SPL when a change in the locus of attention was required whereas IPS elicited sustained activation. These findings suggest a unified functional system of attentional control that initiates and maintains the desired attentive state (Yantis and Serences, 2003).

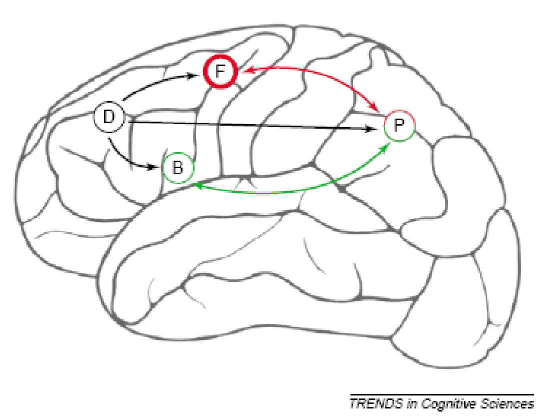
1.4.2 Neural correlates of visual WM

The question of where in the brain information is stored in WM has been extensively investigated since the pioneering studies in monkeys performing delayed response tasks. Two key findings from these early experiments suggested a crucial role for the prefrontal cortex (PFC). First, experimental lesions of the dorsolateral prefrontal cortex (DLPFC), especially within and around the principal sulcus (BA 46), greatly impaired WM performance (Goldman and Rosvold, 1970; Bauer and Fuster, 1976; Funahashi et al., 1993). Second, single-cell recordings from the DLPFC showed stimulus-specific sustained activity throughout the retention interval (Fuster and Alexander, 1971). This sustained activity has been taken as the neural correlate of maintenance processes that take place during the delay. Integrating the neurophysiological evidence and the findings from anatomical connectivity studies in non human primates with the idea of domain-specific storage buffers (Baddeley, 1986), Goldman-Rakic (1987) proposed her influential model of PFC function. Following the separation of the posterior visual areas into a dorsal pathway involved in the processing of spatial and a ventral pathway involved in the processing of object information (Ungerleider and Mishkin, 1982), she claimed that the DLPFC would be engaged in the “on-line” maintenance of spatial memoranda, while the ventrolateral prefrontal cortex (VLPFC) would support the storage of object information. Results from monkey electrophysiology (Wilson et al., 1993), experimental psychology (Tresch et al., 1993; Della Sala et al., 1999; Mohr and Linden, 2005), human neurophysiology (Owen et al., 1997; Postle et al., 1997) and neuroimaging (Smith et al., 1995; Courtney et al., 1996; Munk et al., 2002; Mohr et al., 2006, for a review see Courtney, 2004) confirmed the validity of this domain-specific organisation of PFC, at least within the posterior portions of the PFC [e.g., BAs 8 for spatial and 45/ 47 (in humans) and 45/ 12 (in monkeys) for nonspatial information].

However, an increasing number of evidence suggests that this is not the only organisation principle of PFC. It has been shown that a dorsal-ventral gradient exists also according to the types of processing (Petrides, 1994; Wager and Smith, 2003). The process-specific account states that the VLPFC (BAs 12/ 47 and 45) is involved in the maintenance of information, whereas the DLPFC (BAs 9 and 46) is recruited preferentially to support control functions such as monitoring (Petrides, 2000) or the manipulation of items held in WM (D'Esposito et al., 1999; Owen et al., 1999; Smith and Jonides, 1999).

Domain- and process-specific accounts of the functional subdivisions in frontal cortex are not necessarily mutually exclusive. Models have been developed that successfully integrate these different accounts. For example, Curtis and D’Esposito (2003) suggest that the rehearsal of different types of information occurs in specific frontal premotor areas [dorsal premotor cortex for spatial and ventral premotor cortex (Broca’s area) for verbal rehearsal]. In contrast, the DLPFC/ middle frontal gyrus (BAs 46/ 9) is thought to influence all types of rehearsal in a domain-independent manner by selecting and managing the information to be rehearsed (see Figure 1.4).

Importantly, the lateral PFC is not unique in its responsiveness to retained stimuli. Depending on the type of stimulus, cells with sustained responses have been found in the parietal and the inferior temporal cortex during WM tasks for spatial and object information, respectively (Miller and Desimone, 1994; Chafee and Goldman-Rakic, 1998). As in monkeys, several studies in humans have shown that regions outside the PFC also exhibit sustained delay activity, with object-selective activation in the inferior temporal cortex, and spatial-selective activation in the dorsal parietal cortex (e.g., Munk et al., 2002; Wager and Smith, 2003; Ranganath, 2006). In addition, right posterior brain lesions have been associated with deficits in spatial WM (Jonides et al., 2005). These findings have led to the suggestion that information is stored in visual WM through persistent activity in posterior brain regions, the same regions that are also involved in the perceptual processing of that information (Slotnick, 2004; Jonides et al., 2005; Pasternak and Greenlee, 2005; Postle, 2006; Ranganath, 2006). In other words, the short-term storage of information seems to recapitulate perception (Jonides et al., 2005). These models place less emphasis on a storage role for the PFC and instead (or additionally) emphasise its role in providing extra-mnemonic top-down control over the posterior regions were the information is actually stored (Smith and Jonides, 1999; Curtis and D'Esposito, 2003) (see Figure 1.4).



TRENDS in Cognitive Sciences

Figure 1.4 Simplified model of the neural substrate of spatial WM (red) and verbal WM (green) (Curtis and D'Esposito, 2003). Top-down signals from DLPFC (D) control both the rehearsal of information in domain-specific regions in the frontal cortex (F = FEF, B = Broca's area) and the storage of information in the posterior parietal cortex (P). FEF might reflect spatial rehearsal that could involve the reactivation of saccade goals that would shift gaze to the target location if the eye movement were allowed. Verbal rehearsal involving Broca's area might be mediated through sub-vocalizations of to-be-remembered items.

One milestone in the study of the neural correlates of WM was the development of event-related fMRI techniques in the 1990s as it became possible to isolate the component processes involved in delayed-recognition tasks. Traditionally, WM tasks distinguish the encoding phase, associated with the transfer of information generated from perceptual input into durable storage (Jolicœur and Dell'Acqua, 1998; Ranganath et al., 2004), from the delay period, during which the information is actively maintained (Courtney et al., 1997; Munk et al., 2002), and the retrieval phase, where a test item has to be compared to the stored information (Pessoa et al., 2002; Bledowski et al., 2006). The majority of WM studies (as reviewed above) have focused on the delay period whereas the encoding and retrieval periods were considered in more detail only recently. During the encoding phase activation has been reported in occipito-temporal and occipito-parietal regions associated with perceptual processes and in fronto-parietal regions that are also recruited during the delay phase (Munk et al., 2002; Pessoa et al., 2002). Some evidence suggests that activation in these regions differs according to the type of information (Ranganath et al., 2004; Mohr et al., 2006) and is modulated by WM load (Linden et al., 2003; Leung et al., 2004). One challenge in the study of the neural correlates of WM encoding processes is to distinguish activity related to mnemonic processing from activity reflecting the percteptual processing of the stimulus. One study addressing this issue directly, investigated whether activity in the fusiform face area (FFA) and the hippocampal place area (PPA), regions that are known to exhibit material-specific responses during the perception of faces or scenes, would be modulated by the demands to encode and maintain faces and scenes (Ranganath et al., 2004). Subjects were presented with a set of faces and scenes and needed to encode and maintain only one type of stimulus. Thus, the task relevance of faces and scenes was varied whereas the perceptual content of information was constant across trials. The results showed that the FFA response during the encoding and maintenance period was greater when faces were task-relevant than when scenes were task-relevant. Conversely, PPA activity was greater during the encoding and maintenance period when scenes were task-relevant than when faces were task-relevant. Thus, independent of perceptual stimulation, FFA and PPA activity was enhanced by the demand to actively encode each region’s preferred stimulus type. These results support the hypothesis that WM encoding is implemented through modulation of regions that have evolved for perceptual processing.

Another advance has been the development of parametric designs that systematically vary the demands on WM. Parametric designs are of advantage because all processes except the process of interest are held constant across conditions. In this way parametric designs avoid the pitfalls that arise when comparing activation in a WM task with a non-WM control task. A region, whose activity increases systematically when the number of items to be maintained increases is thus a candidate substrate for memory storage. Effects of WM load have been successfully demonstrated in fronto-parietal regions in a number of studies using various stimulus types such as faces, objects, and positions (e.g., (Leung et al., 2002; Druzgal and D'Esposito, 2003; Linden et al., 2003; Leung et al., 2004). However, other studies have failed to find load-sensitivity in PFC (Postle and D'Esposito, 1999; Jha and McCarthy, 2000). These inconsistent findings from studies on WM load suggest that the PFC might be involved in control functions necessary during WM maintenance rather the storage per se.

1.4.3 Interactions between visual WM and attention

1.4.3.1 Common neural correlates of WM and attention - Evidence from targeted comparisons

Overlap of the cerebral networks of WM and attention has been recently demonstrated in targeted comparisons (LaBar et al., 1999; Pollmann and von Cramon, 2000; Corbetta et al., 2002). For instance, comparing the brain regions engaged in a verbal WM and a covert spatial attention task within the same set of subjects, LaBar et al. (1999) found common activation in fronto-parietal regions including regions along the ventral and dorsal precentral sulcus (PrcS), the supplementary motor area (SMA), and the IPS. Additional sites of overlap included the thalamus, the temporal cortex, the insula and cerebellum. Corbetta et al. (2002) manipulated the allocation of attention to a peripheral location and its maintenance over a 7-s delay interval within the same task and revealed sustained activation in identical regions within the IPS and the FEF. In another study, Pollmann and von Cramon (2000) combined a delayed discrimination task for geometrical objects with a visual search task. Subjects were asked to memorise an object, which had to be matched, after a variable delay, to a target object that was placed in an 11-item array. Visual search difficulty was manipulated by presenting the target and distractor objects within different frames. In the difficult search condition the targets appeared within closed and the distractors within open frames. In the easy search condition it was the other way around. The results revealed a high degree of overlap in the brain areas that showed delay activity as well as activity related to visuospatial orienting including the FEF, the precentral gyri, the pre-supplementary motor area (pre-SMA) and the IPS. Furthermore, it has been shown that spatial attention to representations held in WM is subserved by fronto-parietal brain regions similar to those recruited for spatial orienting in the perceptual domain (Nobre et al., 2004; Lepsien et al., 2005).

1.4.3.2 The cortical substrate of spatial rehearsal effects

A number of studies have shown that spatial rehearsal in WM, i.e. the allocation of attention to memory locations, modulates early sensory processing in visual areas that represent the memorised locations. The time course and the neural locus of these rehearsal effects are similar to those that are observed after manipulations of selective attention (Awh and Jonides, 2001). For instance, Postle et al. (2004) conducted an event-related fMRI study that measured posterior visual activations while subjects performed a spatial WM task. The task required subjects to memorise the location of a bar presented in the left or right visual field over a delay period of 7.5 s that was either filled with a flickering checkerborad or unfilled. Delay-epoch activity in filled trials was stronger in the hemisphere contralateral to the visual field in which the bar had been presented. This attention-based rehearsal effect was found in the extrastriate but not the striate visual cortex. In addition, delay activity in these regions, as well as in the parietal cortex, was also lateralised in unfilled trials, suggesting that attention-based rehearsal produces a baseline shift in areas representing the to-be-remembered location in space. Using a similar delayed-recognition task Jha (2002) studied the time course of visual modulations with EEG. Sensory-evoked event-related potentials (ERPs) were recorded to task-irrelevant probes that were presented early and late during the delay period. The results revealed phasic modulations of the P1 and N1 components, with higher amplitude responses for probes occurring at memorised locations in comparison to probes presented at other locations. This pattern was observed for early and late delay probes consistent with a model of spatial WM in which perceptual level selective attention is utilised throughout the entire period of active maintenance to keep relevant spatial information in mind.

1.4.3.3 Neural capacity constraints of visual WM

The neural capacity constraints of visual WM have been studied using parametric designs that manipulate WM load beyond the capacity limit as indicated by behavioural estimates (Linden et al., 2003; Todd and Marois, 2004; Xu and Chun, 2006). These studies suggest that the capacity limit of visual WM is reflected in the posterior parietal cortex by a load-dependent increase in BOLD activation that reaches a plateau when the capacity limit is approached.

In the study conducted by Linden and colleagues (2003) WM capacity for nonnatural objects was tested. Subjects were presented with up to four complex objects and asked to memorise them over a 12s-delay period. Distributed fronto-parietal regions showed consistently higher activation when multiple objects needed to be remembered as compared to only one object. This effect was present at encoding and continued through the entire delay and retrieval period. Most importantly, whereas activity in the prefrontal and medial frontal cortex monotonically increased in response to WM load, activity in posterior regions including the FEF and IPS peaked when subjects had to maintain only two or three objects and decreased in the highest load condition. This inverted U-shaped response function correlated negatively with the number of items subjects were able to store. Therefore, the authors suggested that the cognitive operations mediated by the IPS and FEF failed to support visual WM when the capacity limit was approached. As these regions are critically involved in visual attention (see above), the authors proposed that the limitation of visual WM is caused by limited attentional resources in posterior regions.

Studies using the change detection paradigm have also localised the capacity limit of visual WM in the posterior cortex. In two complementary studies (Todd and Marois, 2004; Vogel and Machizawa, 2004) subjects were presented with arrays consisting of a variable number of coloured items. After a short retention interval a second array appeared and subjects were asked to detect if there was a change in object identity (colour) and/ or location. The fMRI study (Todd and Marois, 2004) revealed that activity of the posterior parietal/ superior occipital cortex increased from set size 1 to set size 4, levelling off with higher set sizes and this activity strongly correlated with the number of objects stored in VSTM. This activation was observed during both the encoding and the maintenance periods of the task. Consistently, the electrophysiological study (Vogel and Machizawa, 2004) revealed large negative slow waves at posterior parietal and lateral occipital electrode sites that persisted throughout the duration of the memory retention interval with the amplitudes increasing with set size and reaching a limit with arrays of four items. Moreover, the increase in ERP amplitude correlated with the individual differences in VSTM capacity.

Xu and Chun (2006) further dissociated the roles of parietal and occipital cortices for visual WM capacity. In this study subjects were required to remember one to six simple or complex shapes within a change detection paradigm. Estimates of visual WM capacity indicated a maximum of about four objects for simple shapes and only two objects for complex shapes. FMRI activations in the superior IPS and lateral occipital complex (LOC) tracked these capacity estimates, increasing with WM load for simple shapes, but not for complex shape features. For complex shapes a plateau was reached with set size 2. In contrast, regardless of object complexity, activity in the inferior IPS increased with increased WM load reaching a plateau with set size 4. These neural response patterns were observed during both the encoding and maintenance periods. The findings suggest that the capacity of visual WM is determined both by object complexity (represented in superior IPS and LOC) and by a fixed number of objects (represented in inferior IPS). Most importantly, the authors proposed that it is an attention mechanism that selects and determines the maximum number of objects. Thus, these findings again raise the possibility that mechanisms of visual selective attention subserved by the parietal cortex represent one factor that critically constrains the capacity of visual WM.

1.5 Objectives and hypotheses of this dissertation

Visual WM and selective attention have been central topics in cognitive psychology over the last 50 years. Both mechanisms have been extensively studied, however largely in isolation and interactions between the two have rarely been addressed. The general aim of this dissertation was to investigate interactions between these two cognitive systems in terms of behavioural performance and neural activation.

Traditional models of human information processing considered temporary memory and attention distinct, associated with separate functions. Attention and WM were thought to operate at different stages of processing, with attention taking place earlier and controlling which sensory information gets encoded into WM (Broadbent, 1958; Atkinson and Shiffrin, 1968). In this case, visual WM and attention might be represented by different neural substrates. However, recent models of WM suggest that selective attention and WM may rely on a common capacity-limited cognitive mechanism (Cowan, 1988; Baddeley, 1993). Specifically, selective attention has been implicated as a limiting factor for the storage capacity of visual WM (Cowan, 1998, 2001; Wheeler and Treisman, 2002). This view predicts that visual WM and attention share common neural resources. Thus, the main question addressed by this dissertation was the following: Are visual WM and attention mediated by different or the by same cognitive and neural substrates? Following the above mentioned studies, fMRI was specifically used to test the hypothesis that the capacity limitation of visual WM is due to limited-capacity neural resources shared with the process of visual selective attention.

Frontal and parietal brain regions are the primary areas involved both in WM and visual attention (Pessoa and Ungerleider, 2004). Overlap of the cerebral networks of WM and attention has been demonstrated in targeted comparisons (LaBar et al., 1999; Pollmann and von Cramon, 2000; Corbetta et al., 2002). However, neuroanatomic overlap alone cannot be interpreted as direct evidence for common neural and cognitive mechanisms because a small-scale regional specialisation may exist below the resolution of functional imaging (Nieder, 2004). Furthermore, neurons within the same anatomical region may carry out task-specific adaptive functions (Rao et al., 1997), evoking the impression that different cognitive functions, e.g., WM and selective attention, are mediated by the same cortical region. Finally, overlap between the neural substrates that support WM and attention does not necessarily entail a functional relationship between the two cognitive domains. For example, one cannot exclude that shifts of visuospatial attention associated with activation of a given brain region are epiphenomenal to the core processes that encode and maintain information in visual WM (Awh et al., 2006). By demonstrating that memory performance declines when shifts of attention are prevented, it becomes possible to infer a true functional role of attention in visual WM (Smyth and Scholey, 1994; Awh et al., 1998; Oh and Kim, 2004; Woodman and Luck, 2004).

The conceptual link between visual WM and attention addressed in this dissertation stems from one characteristic feature of visual WM and attention, namely their limitation in capacity. It has recently been demonstrated that the capacity limit of visual WM is reflected in the posterior parietal cortex by a load-dependent increase in BOLD activation that reaches a plateau when the capacity limit is approached (Linden et al., 2003; Todd and Marois, 2004; Xu and Chun, 2006). That is, a limit in cognitive processing is correlated with a limit in neural activation, namely a plateau in BOLD activity that cannot be exceeded with increasing demands. One can thus reason that if visual WM and attention share common limited-capacity cognitive and neural resources, these resources will become exhausted in conditions that make high demand on both processes, thus resulting in interference. The present dissertation was therefore motivated by the need to orthogonally manipulate the demand on WM and attention within one single task and to identify brain areas which show an interference effect. Such interference would indicate a limitation of the neural resources available for WM encoding and attentional processing.

Two separate fMRI experiments were conducted that combined visual search and delayed visual discrimination for either objects (experiment 1) or locations (experiment 2). In experiment 1 participants performed easy or difficult visual search in order to encode one or three complex objects into visual WM. In experiment 2 they performed easy or difficult visual search in order to encode one, three, or five locations into visuospatial WM. Attentional demand was manipulated by implementing two search conditions in which target items had either unique features (i.e., colour) and were highly discriminable from the distractors ("pop-out condition" [PO] = low attentional demand) or shared the features with the distractors and were difficult to discriminate ("non pop-out condition" [NPO] = high attentional demand) (Treisman and Gormican, 1988; Duncan and Humphreys, 1989) (see Figure 3.1).

Neural capacity constraints for visual WM have been observed both during the encoding and maintenance of visual information (Linden et al., 2003; Todd and Marois, 2004; Vogel and Machizawa, 2004; Xu and Chun, 2006). In this disseration I focussed on the encoding phase during which the search array was presented. Physical properties of the stimulus display were identical across conditions, which ruled out differences in brain activation owed to differences in sensory stimulation. This design provided the possibility of investigating common and selective activation for visual WM and attention and most importantly to test for interactions between both processes in terms of BOLD activity.

Prior to the fMRI studies of this dissertation a behavioural study was conducted consisting of a series of five experiments that used the same stimuli as in the fMRI experiments. These experiments served two purposes. First, the behavioural experiments were used to validate the engagement of the relevant attentional and WM processes by the chosen task manipulations and to optimise the design of the fMRI experiments. Second, these experiments addressed the question of whether and how subjects can cope with the common capacity limitations of visual WM and attention.

1.5.1 Hypotheses

1.5.1.1 Behavioural study - Attentional demand influences strategies for encoding into visual working memory (Chapter 3)

In many real-life situations the demands on visual attention and WM occur simultaneously. Given the evidence that visual attention and visual WM share common resources, and thus interfere when engaged simultaneously (Awh et al., 1998; Jolicœur & Dell'Acqua, 1998, 1999; Smyth and Scholey, 1994; Oh and Kim, 2004; Woodman and Luck, 2004; Barrouillet et al., 2007), the question arises how these limitations can be overcome. To answer this question in a laboratory setting, the behavioural study tested whether and how participants can encode complex objects into WM while engaging selective attention for a visual search task. In this study, the stimuli, procedure, and design were the same as in the fMRI study, except for the following two differences. First, WM load varied from 1 to 5. Second, the stimulus array was not shown for a fixed amount of time but instead remained visible until the participant pressed a response key. Thus, the most important dependent variable was the presentation time of the stimulus array that participants needed to achieve good WM performance. It was investigated how this time changed as a function of WM load and attentional demand. On the one hand, this time was used to determine the duration of the encoding period implemented in the fMRI study. On the other hand, this time allowed isolating the processes that enabled participants to deal with concurrent demands on visual attention and encoding into visual WM.

1.5.1.2 fMRI study - Common neural substrates for encoding into visual WM and selective attention (Chapters 4 and 5)

In the fMRI part of this dissertation the hypothesis was tested that the capacity limitation of visual WM is due to common limited-capacity neural resources shared by visual WM and selective attention. Two separate fMRI experiments were conducted that combined visual search and delayed visual discrimination for either objects (experiment 1) or locations (experiment 2). The paradigm allowed differentiating between three patterns of activation that were associated with different contributions to the cognitive task components.

1. An exclusive main effect for difficulty of either encoding into WM or attentional selection was expected in areas that preferentially subserve that particular task component.
2. Overlap areas that mediate both processes should show main effects for both task manipulations with an additive increase in BOLD activation as a consequence of an increase in the demands on WM encoding and visual search difficulty.
3. One can expect to reveal areas showing an interaction effect between attentional demand and WM load. Activation in these regions should demonstrate a less than additive increase in BOLD activation with increasing demands on WM and visual search. Thus, activation should reach a plateau as WM and attentional demands increase with the difference in the BOLD response between high WM load and low WM load levelling off in the non pop-out condition. Such interference would indicate a limitation of the neural resources available for WM encoding and attentional processing and offer direct evidence for common neural resources shared by the processes of encoding into visual WM and visual selective attention.

Following the idea of distinct cognitive and neural mechanisms for the WM storage of object and spatial information (e.g, Della Sala et al., 1999 , Munk et al., 2002; Mohr et al., 2006), two fMRI experiments were conducted that required subjects to encode either objects or locations into WM. Thus, the question whether effects of interference between visual attention and WM encoding depend on the particular information being encoded or generalise across different classes of stimuli could be addressed. In the former case I expected to find effects of interference in distinct ventral and dorsal fronto-parietal regions for the encoding of object and spatial information, respectively. In the latter case an interaction effect between WM load and attentional demand should be observed in similar fronto-parietal regions in experiments 1 and 2.

The empirical part of this dissertation starts with the behavioural study (chapter 3). In this chapter the hypotheses addressed by each of the five experiments will be outlined in more detail. The fMRI part of the dissertation consists of two experiments that combined visual search with either object WM (chapter 4) or spatial WM (chapter 5). Yet, before the studies of this dissertation are presented and discussed, in chapter 2 I will give a brief introduction into the methods applied.

Chapter 2 - Methods

2.1 Psychophysics: Assessing stimulus-response relationships

In its original view psychophysics refers to the quantitative study of the relationship between the stimulus intensity as specified in physical terms and the intensity of sensations and perceptions evoked by these stimuli (Fechner, 1860). Based on the assumption that the human perceptual system is a measuring instrument yielding results (experiences, judgments, responses) that can be systematically analysed, psychophysics introduced the objective measurement of mental operations into the field of psychology. Departing from the observation that all mental operations take time (“mental chronometry”, Donders, 1969), early experimental and cognitive psychologists adapted the psychophysical approach by using reaction time (RT) and response accuracy as most accessible indicators for higher cognitive processes. With the re-emergence of cognitive psychology in the middle of the 20th century, speed and accuracy data provided the empirical background for the development of sophisticated models of cognitive functions within the framework of information processing (Neisser, 1974). In essence, RT data offered useful information for the isolation of different component operations constituting specified cognitive functions (Sternberg, 1969), such as attention, memory, problem solving, decision making or learning.

2.1.1 The additive-factor method

One of the principal methods that have been used to decompose mental processes into their constituent stages is the additive-factor method introduced by Sternberg (1969). The additive-factor method is applied to RT data from factorial experiments in which the effects of two or more experimental variables are studied. RT is treated as a composite measure that reflects the entire process and that can be divided into serial distinct stages or processes. The following predictions are made: First, the effects of a variable that affects overall RT by varying the time to complete one stage will add to effects of variables that affect other stages. In other words, the effect of varying one factor is unchanged by varying another factor. Second, two variables that influence at least one stage in common will lead to an interaction effect. With this framework, it becomes possible to determine at which stage a new factor has its influence.

In the classical Sternberg task (Sternberg, 1966) subjects were asked to determine whether or not a probe digit had been present in a just previously presented series of digits. The factors stimulus quality (intact vs. degraded), number of items (1 to 6), response type (positive vs. negative), and frequency of response type were manipulated and Sternberg found that they had an additive effect on RT. For instance, the increase in RTs resulting from degradation was the same regardless of the set size, the response type, and the frequency of response. In addition, RTs increased linearly from set size 1 to set size 6. These additive effects were interpreted in terms of a series of distinct stages including a sensory, a serial-comparison, a binary decision, and a response selection stage. An example of an interaction effect has been described when using the additive-factor method to examine the effects of experimentally induced sleep deprivation on cognitive processing within a digit-naming task (see Sternberg, 2001). In this study, the effect of the factor stimulus quality (intact vs. degraded stimuli) on RTs considerably increased under sleep deprivation. This interaction indicated that both factors influenced the stage of stimulus identification.

2.2 Functional magnetic resonance imaging

2.2.1 Basic principles: The interaction of physics and physiology

FMRI takes advantage of the coupling between neural activity and haemodynamics in the brain, i.e. the local control of blood flow and oxygenation. Changes in the level of cortical blood oxygenation influence the signal intensities in magnetic resonance images. Thus, fMRI utilises the technique of magnetic resonance imaging (MRI) in order to visualise brain signal dynamics that represent indirect measurements of neural activation.

The fundamental signal for MRI comes from hydrogen atoms (protons), which are abundant in the water molecules of the brain (Heeger and Ress, 2002). The protons rotate (they have a spin) and thus possess a magnetic moment. In field-free space, these magnetic dipoles are oriented randomly. In the presence of an external magnetic field, like for instance the magnetic field of the MR scanner, the protons align parallel or antiparallel to this field. The two orientations are not equally frequent, which results in a net magnetisation parallel to the magnetic field, also called longitudinal magnetisation. At the same time, the protons rotate around the axis of the applied field which is termed precession. The speed of this movement or precession frequency is dependent on the strength of the external magnetic field. If a radio-frequency pulse is applied at the same frequency (Lamor frequency), energy is absorbed by the protons and the equilibrium between parallel and anti-parallel protons changes. The protons alter their alignment from the direction of the main magnetic field (“longitudinal magnetisation”) to the direction opposite the main magnetic field (“transverse magnetisation”). As the protons try to realign with the main magnetic field, they emit the absorbed energy in the form of a radio frequency magnetic field until they return to their equilibrium state (relaxation). This field can be detected by a receiver coil and represents the signal measured with MRI (for further details see Cohen, 1996). The measured radio-frequency signal decays exponentially over time (measured as relaxation time) depending on the different chemical and physical properties of the local tissue surrounding the protons. There are two types of relaxation times important for MRI. One is called the longitudinal relaxation time (T1), representing the time the spin system needs to recover to its thermal equilibrium, and the second one is called the transverse relaxation time (T2), representing the time needed by the excited spins to develop a phase incoherence before relaxing back to the equilibrium state. The transverse relaxation depends on both mutual interactions between spins and random inhomogeneities in the magnetic field. The combined relaxation time is called T2*. By applying adequate pulse sequences, the relaxation time properties of the probe tissue can be determined. As fat and water have different T1 and T2* parameters, a contrast between them can be observed in the reconstructed image of the probe tissue.

Important for functional MRI is the fact that contrast can also be obtained between oxygenated haemoglobin (HbO_2) and deoxygenated haemoglobin (dHb), because HbO_2 and dHb slightly differ in their magnetic susceptibility: dHb is paramagnetic and introduces an inhomogeneity into the magnetic field of the surrounding tissue.

HbO_2 is diamagnetic and has little effect. This image contrast has been termed blood oxygen level dependent (BOLD) effect (Ogawa et al., 1990). In the brain, the ratio between HbO_2 and dHb in a blood vessel, in particular in the venoles, depends on the energy consumption of adjacent neurons. When a neural event occurs, increased energy consumption of the neuron increases oxygen extraction from the blood vessel. This leads to a decrease of diamagnetic HbO_2 compared to paramagnetic dHb. This is reflected in the fMRI as a decrease of T2*-weighted signal in the first few seconds after the onset of the neural activity and often referred to as initial dip (Buxton, 2001). Following the initial dip, the increased energy consumption of the neuron results in a compensatory increase of regional cerebral blood flow. This effect is called the haemodynamic response. As a consequence, the relative concentration of HbO_2 and dHb in the blood is altered in the favour of HbO_2. The relative decrease in dHb associated with neural activity leads to an increase in the local homogeneity of the magnetic field which in turn results in an increase of the T2*-weighted signal (for a review see Howseman and Bowtell, 1999; Heeger and Ress, 2002). The increased T2*-weighted fMRI signal due to the BOLD effect has been widely used as a measure of the local neural activity. It has been demonstrated that the BOLD signal is best predicted by local field potentials (LFPs) of an ensemble of neurons which reflect the synaptic input into a neural population rather than the spiking activity (Logothetis et al., 2001). However, the detailed cellular and molecular mechanisms that underlie the coupling of neural activity and the haemodynamic changes still need to be determined.

2.2.2 Event-related fMRI

The echo planar imaging sequence (Mansfield, 1977) which is sensitive to changes of the oxygenation level in the tissue is considered to be the method of choice in fMRI experiments (Kwong, 1995). Due to its short scanning time, a single slice can be measured in less than 100 ms, it becomes possible to characterise and detect transient haemodynamic responses to brief stimuli or tasks (event-related fMRI, efMRI) (e.g., Buckner, 1998). In contrast to blocked design procedures, in which the signal is temporally integrated across a series of trials, event-related designs allow to isolate individual trial events or subcomponents of trial events. Using efMRI design and analysis methods it has become possible to separate the different component processes involved in a WM task (e.g., encoding, maintenance, and retrieval) (Courtney et al., 1997; Munk et al., 2002). In standard blocked trial procedures these processes would be blurred together and the differential contributions across regions difficult to appreciate. However, efMRI designs are also of advantage because they allow to randomise the order of trials of different conditions, thus eliminating potential confounds, such as habituation, anticipations, subject’s cognitive set or other strategy effects. In addition, response amplitudes are assessed for each individual trial which allows the correlation between behavioural performance and fMRI responses.

2.2.2.1 The nature of the event-related BOLD signal

2.2.2.1.1 Reliability of the signal

The reliability of the BOLD signal has been indicated by a number of studies showing that within a given subject and cerebral region the BOLD response is consistent from one set measurements to the next (e.g., Miezin et al., 2000). A typical haemodynamic respose function exhibits a small initial dip in the first 2 s, a rise and a peak at 4 to 6 s after stimulus onset, a decay to baseline, and a undershoot. The return to baseline takes about 16-20 s. The general shape of the BOLD response appears similar across early sensory regions. However, considerable variation in the timing and shape of responses has been observed across the brain, particularly across higher cortical regions (Buckner, 1998). In addition, the BOLD response appears to vary considerably across people (Aguirre et al., 1998).

2.2.2.1.2 Sensitivity of the signal

The BOLD signal is seen as reflecting the result of a transient increase in neuronal activity which occurs in response to even very brief periods of stimulation (Heeger and Ress, 2002). An extreme example of this sensitivity has been reported by Savoy et al. (1995) demonstrating that visual stimulation in the range of tens of milliseconds is sufficient to elicit a detectable signal change. Initially, transient signal changes were found in the sensory and motor cortex. However, similar increases have also been shown in cognitive task paradigms, even though the signal changes can be considerably smaller in magnitude (Buckner et al., 1996).

2.2.2.1.3 Linearity of the signal

The central assumption guiding inferences that are made from efMRI data about neural activity is that the fMRI response is the output of a linear time-invariant system (“linear transform model”, Boynton et al., 1996). According to this model the haemodynamic response does not change with time and summates over time and sequential events in a roughly linear fashion. Thus, the shape of the BOLD response to a given period of stimulation is predictable and relatively stable across events, even when there is an overlap in the responses to successive events. Different trial types can then be randomly intermixed and statistical methods can be used to estimate the separate contributions of these different types to the variability in the measured fMRI signal.

The linearity of the signal has been demonstrated for simple visual stimulation even at very short intertrial intervals of up to 2 s in primary sensory cortex (Boynton et al., 1996; Dale and Buckner, 1997). For instance, the fMRI signals evoked by a single event (e.g., 1 s flickering checkerboard) exhibit the characteristic impulse-response function, with the haemodynamic responses to subsequently presented trials additively superimposing. Moreover, the estimated response to each successive trial closely matches that of the first trial (Dale and Buckner, 1997). In addition, it has been shown that the reponse to a 12-s stimulus can be predicted by summing the response to a 6-s stimulus with a copy of the same response delayed by 6 s (Boynton et al., 1996).

However, situations in which stimulus events occur extremely rapidly (shorter than 2 s apart) can show marked departures from linearity (e.g., Rees et al., 1997). This nonlinearity is typically a saturation whereby the response to a sequence of events is smaller than would be predicted by the summation of responses to each event alone. Also, nonlinearites in the amplitude of the BOLD response have been found as a function of stimulus duration (e.g. for visual stimuli less than 4 s duration) and stimulus contrast (e.g. for stimuli less than 40% contrasts) (Boynton et al., 1996;

Vazquez and Noll, 1998). The sources of these saturation effects are still incompletely understood and might occur at the haemodynamic (e.g., a non linear relationship between the extraction of oxygen and stimulus parameters) and/ or neuronal level (e.g, a non linear relationship between the neuronal activity and stimuls parameters). Therefore, to further evaluate the linear transform model of the fMRI signal direct comparisons between fMRI and neuronal signals are needed. A recent, seminal study by Logothetis et al. (2001) that correlated the BOLD response with electrophysiological measurements in the monkey has emphasised the closeness of the BOLD signal to LFPs rather than spiking activity and demonstrated a linear relationship between the magnitude of the LFPs and the BOLD signal. At some recording sites and for a specific stimulus contrast, the linear transform model predicted the measured fMRI responses well, explaining more than 90% of the variance in the fMRI signals. However, substantial non-linearities were also observed depending on the stimulus contrast. A 12% stimulus contrast evoked about half the maximum fMRI response but much less than half the maximum LFP indicating a monotic but non linear relationship between the fMRI signals, neuronal responses, and stimulus intensity. In addition, in another study recording electrical and haemodynamic responses to low and high contrast visual stimuli in the cat visual cortex, the haemodynamic reponses were positively correlated with stimulus intensity, spiking activity, and most strongly with the power of LFP oscillations in the gamma frequency range (Niessing et al., 2005). Because these oscillations increase with the synchrony of synaptic events, a close link between haemodynamic responses and neuronal sychnchronisation has been suggested.

2.2.3 Analysis of efMRI data

Changes in the obtained fMRI signal across time assumingly rely on the BOLD effect that reflects neural activity. However, other factors such as head motion, changes in overall blood flow, or changes in the static magnetic field can influence the signal as well. If a certain area shows fMRI responses, this might be ideally caused by the experimental stimulation, but could also have different reasons, e.g. unwanted cognitive activities by the subject. Thus, fMRI experiments entail a probabilistic approach that is repeated experimental stimulation and statistical analysis.

A powerful technique for analysing efMRI data is to explicitly model the predicted time course of the voxel-wise fMRI signal within the General Linear Model (GLM). The application of GLM and respective statistical analyses in fMRI was introduced by Friston et al. (1995). Providing a single framework for many statistical tests and models, the GLM gives great flexibility in analysing multifactorial designs.

In the GLM the observed time series Y_{ij} at each voxel j = 1,…, J and time point (e.g., scan) i = 1,…, I, can be predicted by a linear combination of explanatory functions (“predictors”) x_{ik} for the different trial events that are thought to contribute to the variability in the BOLD signal plus an error term (Kiebel and Holmes, 2003):

$$Y_{ij} = x_{i1}\beta_{1j} + x_{i2}\beta_{2j} + \ldots.\ x_{iK}\beta_{iK} + e_{ij} \qquad e_{ij} \sim N(0, \sigma_j^2), \qquad (1)$$

where the errors (e_{ij}) are independent and identically distributed normally. The explanatory functions x_{ik} are defined according to the layout of the design and a model of the haemodynamic response. Assuming that the BOLD signal is the output of a linear system, then the predictors can be expressed as the convolution of the neural activity with a haemodynamic response function (HRF) (Henson, 2003). The use of the GLM in its simplest form implies the assumption that the fMRI signal changes immediately after stimulation in a rectangular pulse manner. Other functions have been shown to provide a reasonably good fit to the haemodynamic response such as the gamma function (Boynton et al., 1996) and the canonical HRF (Friston et al., 1998) which was used in the present study. This function is characterised by two gamma functions, one modelling the peak and one modelling the undershoot of the haemodynamic response.

The β_{kj} (k = 1,…, K) are K unknown parameters for each voxel j describing the amplitude of the expected time course. These parameters can be estimated such that the predicted time course is as close as possible to the measured time course.

Equation (1) can be expressed in the usual matrix form for the GLM:

$Y = X\beta + e,$ (2)

where Y is the data matrix, which has one column per voxel and one row per scan. The matrix X which is composed of the coefficients x_{ik} is called the design matrix or the model. The design matrix has one column for every modelled effect and one row for each scan. β is a vector of K weights or parameters for voxel j. e is the matrix of the error terms.

Within multiple regression analyses a least-mean-squares fit of this model to the fMRI time series data produces estimates for the $ß_{kj}$ weights of the predictors and can be tested against the residual errors using F statistics. The resulting multiple correlation coefficient gives the linear correlation between the predicted time course as determined by the full model and the actually measured time course of the signal. The β value of a certain predictor however represents the partial correlation of this predictor with the actually measured neuronal activation and can hence be interpreted in comparison and in contrast to another predictor or set of predictors of the factorial design. The significance of these contrasts can be indexed by t-tests. In addition, the obtained ß weights can serve as input for a second-level ANOVA analysis. Thus, the ß values of subjects can be treated explicitly as realisations of one or several within-subjects factors which allows to directly test for main effects and, most importantly in the context of the present work, for interactions between different factors.

2.2.4 Random effects analysis

Neuroimaging data from multiple subjects can be analysed using fixed-effects or random-effects analysis (Penny and Holmes, 2003). In fixed-effects analysis only within-subject variance (e.g. the variability between scans) is taken into account, and thus the reported results are only valid for the group included in the analysis. To make inferences about the population from which subjects are drawn, random-effects analysis is required considering both within-subject and between-subject variance. This means that the subject variable is treated as a random selection from the population as it is a standard procedure in psychophysics. In fMRI including the present experiments this can be realised by using a hierarchical analysis modelling within-subject variance at the first level and then taking the parameter estimates from each subject to a second level of inference in which their distributions are compared directly (Josephs and Henson, 1999). In order to estimate general population effects a minimum sample size of 12 subjects is recommended (Desmond and Glover, 2002).

2.2.5 Statistical thresholding using the false discovery rate

Finding objective and effective thresholds for voxelwise statistics derived from neuroimaging data has been a long-standing problem. The alpha level, which gives the probability of classifying a voxel as active when in fact the null hypothesis of no difference between specified experimental conditions is true, is conventionally set to $p < .05$. However, when performing multiple tests, in an fMRI experiment about 100,000 voxels are measured separately, the number of false positives becomes very high (with a given alpha level of $p < .05$ about 5000 voxels would be declared active when they are really inactive). The Bonferroni correction is one of the most common methods for controlling the false-positive rate by adjusting the threshold depending on the number N of independent tests: $p_{corr} = p / N$ (Bortz, 1993). However, in fMRI experiments the time series in neighbouring voxels are not independent but correlate to some degree. Therefore, the Bonferroni correction is a conservative method that has strong control of the alpha error with the disadvantage of increasing the beta error when applied to the entire data set. The Bonferroni correction has a tendency to wipe out both false and true positives. An alternative approach to correct for multiple comparisons is the false discovery rate (FDR) (Genovese et al., 2002). This approach has been recently adapted to fMRI data and is used in the present work. The FDR threshold controls the expected proportion of false positives (incorrect rejections of the null hypothesis) among those tests for which the null hypothesis is rejected (voxels that are declared as active). Thus, the FDR method adapts to the amount of activity in the data. The method is very strict if there is not much activity in the data, but assumes less conservative thresholds if larger regions of the brain show task-related effects. Therefore, the FDR is more sensitive than the Bonferroni correction which controls the chance of obtaining any false positives. In practice, the researcher chooses in advance a tolerable rate of false discoveries and the specific thresholds are then determined from the data.

Chapter 3 - Behavioural study: Attentional demand influences strategies for encoding into visual working memory

The goal of the behavioural study was two folded. First, the behavioural experiments were used to validate the engagement of the relevant attentional and WM processes by the chosen task manipulations and to optimise the design of the fMRI experiments. These aspects will be discussed in chapters 4 and 5. Second, the behavioural study aimed at investigating whether and how participants can cope with the common capacity limitations of visual WM and attention. The main question was the following: If visual attention and visual WM share common resources and thus, interfere when engaged simultaneously, how can these limitations be overcome?

To investigate this question, a task was designed that combined the classical features of visual search experiments, which have been widely used in the study of selective attention (Treisman and Gelade, 1980; Wolfe, 1998b), with those of visual WM studies (e.g., Wheeler and Treisman, 2002; Oh and Kim, 2004; Olsson and Poom, 2005) (see Figure 3.1). In each trial, participants were presented with an array of nine objects and had to memorise only some of them (targets), while the others could be ignored (distractors). Determination of the target locations was based on an L-shaped item located in the centre of the object, but only the outer shape of the object and its orientation had to be remembered. Thus, the present procedure allowed manipulating independently the demands on encoding into visual WM and the demands on attention for visual search of target locations. Attentional demand was manipulated by implementing two stimulation conditions in which the L-shaped items had either unique features (i.e., colour) and were highly discriminable from the distractors (resulting in perceptual “pop-out” [PO]) or shared the features with the distractors and were thus difficult to discriminate (“non pop-out” [NPO]) (Treisman and Gormican, 1988; Duncan and Humphreys, 1989). Only in the latter case it was expected that the determination of the target locations would require the attention-demanding serial search, which is commonly indicated by a linear increase in search times as a function of the number of distractor items in the array (Treisman and Gelade, 1980; Treisman and Sato, 1990). To manipulate the load of WM encoding, the number of target items was varied in each array, which ranged from one to five.

In the classical visual search paradigm, the display remains visible until the participant responds and response accuracy is usually high. Therefore, RT is the most important measure in this paradigm as it indicates the amount of time required to determine the presence or absence of a target presented among distractors (Treisman and Gelade, 1980; Duncan and Humphreys, 1989; Treisman and Sato, 1990; Wolfe, 1998a). This setup was highly instrumental for the development of one of the most successful theories in psychology - the feature integration theory (Treisman and Gelade, 1980; Treisman and Sato, 1990). Here, the same concepts were used to study the processes underlying the encoding of information into visual WM. Thus, the most important dependent variable was the presentation time of the stimulus array that participants needed to achieve good WM performance, and which they self-paced by a key press. It was investigated how this time changed as a function of memory load and of attentional demand.

A similar dependent variable has been used in a recent study that investigated the role of visual WM for the formation of visual LTM (Nikolić and Singer, 2007). The authors first estimated the WM capacity for the locations of the target stimuli that either did or did not pop out from the distractors, and then requested the participants to memorise accurately a number of target locations that grossly exceeded the capacity of WM. The participants self-paced the memorisation process and the obtained encoding times were measured reliably ($r > 0.90$) and increased linearly as a function of target set size. Importantly, the changes in the slopes of these linear functions could be predicted accurately from the changes in the estimated WM capacities for the same stimuli. The authors concluded that the capacity of WM determined the speed with which visual LTM was created. This provided the missing evidence that visual WM played a pivotal role in the storage of information in visual LTM. Nikolić and Singer reported that the self-paced measure of the encoding times was reliable given that an immediate performance feedback was supplied at each trial, which in turn enabled the participants to learn quickly, on a trial-and-error basis, the minimum amount of effort (time) that was needed to achieve the required level of performance (95% correct in that study). In contrast, if such feedback was not provided, participants tended to shorten the encoding time and hence, trade the accuracy for speed.

The study of Nikolić and Singer (2007) investigated the WM capacity for the locations of the target stimuli only, thus without any additional contents presented on the display. In that study, WM could be loaded with very short stimulus presentations of about 1 s. The present study addressed the WM for relatively complex objects that were presented at the target locations. Thus, participants needed not only to select the target locations but also to extract and memorise the various shapes that were presented at these locations. This required much longer presentation time than 1 s, as the information could not be loaded “directly” but successful encoding required the participants to engage into a more elaborated processing. The main goal of the present study was to investigate the nature of these processing steps, and to this end, two types of strategies were considered.

In a “search-and-encode strategy” participants would encode each shape as soon as they selected a relevant location, interleaving thus the search process with the WM encoding. If this was the case, presentation time should be simply divided between the two task components, and the presentation time that participants need in the non pop-out condition should be the sum of the presentation time in the pop-out condition and the time needed to select the relevant locations in the non pop-out condition. Thus, as empirical support for the search-and-encode strategy, I looked for the evidence that the times for encoding and determination of target locations are additive.

The other considered strategy was postulated to involve two separate steps of encoding (“two-step encoding strategy”). In the first step participants would select and memorise only the locations of all target items and only then would encode the associated shapes at a later step. The additional process of memorising the target locations would require additional processing time. Therefore, for that case, a super-additive combination of the times for encoding and determination of target locations was predicted in the non pop-out condition. The time needed to memorise the locations was measured directly and it was investigated whether this time corresponded to the additional time required to encode the target shapes in the non pop-out condition.

Importantly, the two-step encoding strategy but not the search-and-encode strategy implies interference between WM encoding and attention. A search-and-encode strategy should be possible if the two components need to be executed sequentially but do not interfere with each other, i.e., the search for a new target does not erase the contents stored previously in WM. As the existing evidence suggests that this is not the case (Smyth and Scholey, 1994; Awh et al., 1998; Jolicœur and Dell'Acqua, 1998, 1999; Oh and Kim, 2004; Woodman and Luck, 2004; Barrouillet et al., 2007), the two-step encoding strategy was considered as a possible tactic for overcoming this interference. Therefore, if empirical evidence favours one of the two strategies, the result provides also indirect information on whether, in this task, visual WM encoding and attention interfere.

Synopsis of experiments

Five experiments were conducted in which the study phase consisted always of identical stimuli, the tasks differing only in the instructions and in the test displays. Participants were debriefed at the end of each experiment and were asked about their subjective experience and strategies. In the main experiment (Experiment 1), participants encoded complex target shapes into WM, while determining their locations in a low or high attention-demanding visual search task (i.e., presence or lack of perceptual pop-out). WM performance was comparable across search conditions. Presentation time increased with increased WM load and, most importantly, with the lack of pop-out. Further experiments (Experiments 2 to 5) investigated the reason for the increase in the presentation time by contrasting the two, above described, strategies.

Experiment 2 and 3 tested the hypotheses of additivity vs. super-additivity of the times needed to encode and determinate the target locations. In Experiment 2, the time needed for simple visual search was measured. These times could not explain the increased presentation time produced by the lack of pop-out in Experiment 1. Therefore, Experiment 3 tested whether the slower processing in the non pop-out condition in Experiment 1 could be explained by repeated searches, owing to a putative lack of memory for visited target locations (Irwin, 1992; Peterson et al., 2001) and to the need to search the entire array. The need to search repeatedly was reduced by informing the participants at each trial about the upcoming number of targets. The time saved by this manipulation again could not explain the costs on presentation time produced by the lack of pop-out in Experiment 1. Therefore, the results from Experiments 1 to 3 indicated consistently super-additivity of the times for encoding and determination of the target locations, favouring the two-step encoding strategy.

In the remaining two experiments (Experiments 4 and 5) the two-step strategy was tested further. The times that participants needed to memorise the locations of the target items only were measured and it was investigated whether these times could explain quantitatively the difference between the pop-out and non pop-out conditions in Experiments 1 and 3. Indeed, in Experiments 4 and 5, the times needed to memorise the target locations accounted well for the presentation time offsets between pop-out and non pop-out conditions in Experiments 1 and 3, respectively. These results again favoured the two-step strategy.

3.1 Experiment 1 - Visual search and encoding of *objects* into WM

Experiment 1 addressed the question whether and how participants can encode complex objects into WM, while engaging selective attention for a visual search task. Participants memorised the shapes of only those objects whose centre items matched the target items, and were instructed to ignore all the other objects. Determination of the target locations was easy in the pop-out condition and required attention demanding serial search in the non pop-out condition. Only WM for the shapes was tested and there were no explicit requirements to use any particular strategy in this task. Thus, it was investigated whether participants could advance the WM performance in the non pop-out condition to the level of the performance in the pop-out condition, and if so, at what cost on presentation time.

3.1.1 Method

3.1.1.1 Participants

Thirty-six students and employees of the University of Frankfurt M. (15 males, 21 females) volunteered in this study. The mean age of the participants was 26.1 years (range: 19-33). In this and in all other experiments all participants reported normal or corrected-to-normal visual acuity, normal colour vision, and no history of neurological or psychiatric illness.

3.1.1.2 Apparatus and stimuli

The stimuli were presented through a PC on a 17” colour monitor using ERTS (Experimental Run-Time System, Berisoft, Frankfurt, Germany). A chinrest was used to minimise head motion and to ensure that the observer’s eyes were positioned in a constant distance of 42 cm from the screen. Response keys were located on the computer keyboard. The experiments were performed in a dimmed room.

The display in the study phase consisted of nine different grey geometric shapes (each spanning approximately 1.1° × 1.1° of visual angle), arranged in a 3 × 3 matrix, and presented in the centre of the screen and on a black background. The shapes were selected at random without replacement from a set of 12 shapes and each was oriented randomly in one of the four possible directions, having to discriminate, in total, between 48 different objects. A small L-shaped item (0.3° × 0.3°) was placed in the centre of each shape. The Ls appeared in one of four different orientations (0°, 90°, 180°, and 270°, clockwise) and were coloured either blue or red (see Figure 3.1). Participants needed to memorise only the shapes associated with an L oriented 90° (target items). The shapes associated with Ls of other orientations could be ignored (distractor items). The number of target items within each display varied randomly between one and five. In the pop-out condition target L’s always appeared in blue and distractors in red. Distractor L’s were always oriented 270°. In the non pop-out condition each target and distractor was assigned randomly either blue or red colour. In this condition, the distractor items could take any of the remaining three orientations (0°, 180°, and 270°). In the test phase participants were presented with a single shape in the centre of the screen and without the centre item. The luminance of the shapes, the blue, and the red centre items was 12.3, 6.01, and 9.87 cd/m², respectively. The background luminance was 0.01 cd/m². During the delay period a white central fixation cross was presented on a blank screen (0.2° × 0.2°, 60.06 cd/m²).

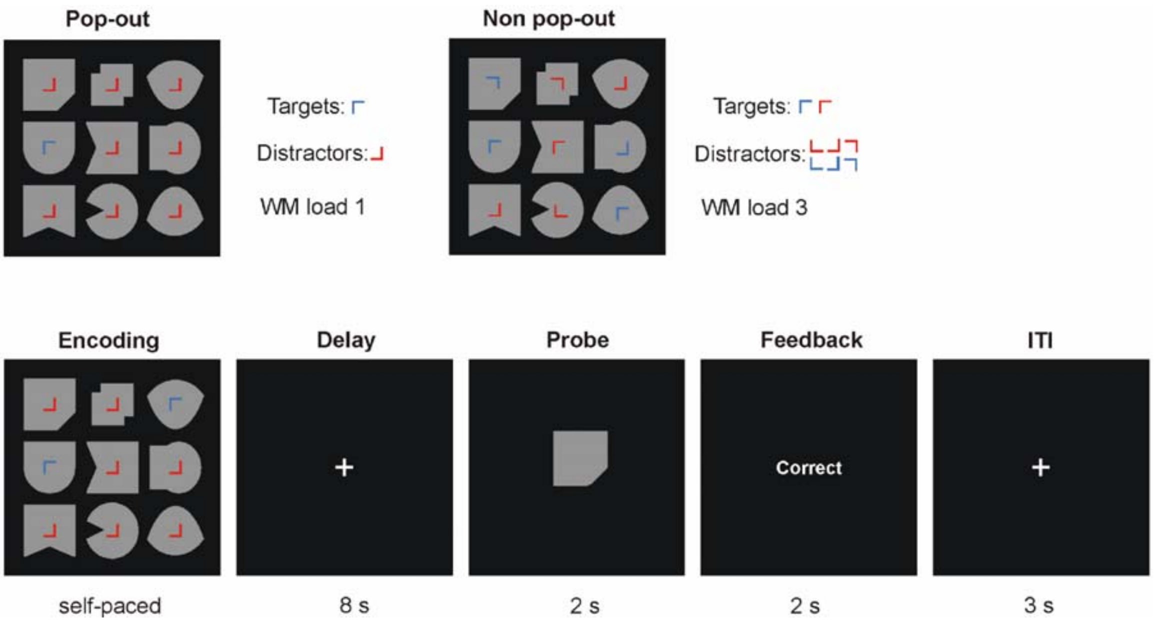


Figure 3.1 The stimuli and procedure used in Experiment 1. Participants determined the locations of the target items and memorised the shapes surrounding them. Targets and distractors were distinguished by the items presented in the centre of each object. Attentional demand for determination of the target locations was manipulated by the presence and absence of perceptual pop-out. In the pop-out condition blue target items were presented among red distractors. In the non pop-out condition colours were assigned randomly to the target items. WM load was manipulated by changing the number of targets which varied between 1 and 5. The presentation time that was needed to achieve high WM performance was self-paced. After an interval of 8 s, participants had to judge whether the test shape matched one of the target shapes. ITI: Inter-trial interval.

3.1.1.3 Design and procedure

A 2 × 5 within-subjects factorial design was used, with two levels of attentional demand for target-distractor discrimination (pop-out and non pop-out) and five levels of WM load, determined by the number of targets (one to five targets). Each of the 10 experimental conditions was presented equally often (12 trials per condition). Pop-out and non pop-out conditions were presented in separate blocks of 10 trials, with six blocks for each condition. This amounted to a total of 120 experimental trials per participant. The trials were fully randomised within blocks and pseudo-randomised across blocks and across participants. Before starting a new block, participants were always given an instruction about the targets they needed to search for. At the beginning of the experiment participants performed two practice blocks of 10 trials, one for each of the two attentional conditions.

Each trial began with the presentation of the nine-item array, which remained visible until the participant pressed the response key. Participants had to determine the target locations and to memorise the shapes associated with the targets. The time they needed to achieve high memory performance, indicated by a key-press, was used as a dependent variable (presentation time). Participants were instructed to emphasise accuracy over speed in order to ensure that response accuracy was high and comparable across different attentional-demand conditions. After the display disappeared they fixated a cross during a delay period of 8 s, which was followed by the presentation of a single test shape. Participants were then required to indicate whether the test shape matched in the form and orientation one of the target shapes presented previously by pressing the “Y” or “N” key for match and non-match, respectively. Half of the trials were matches. In 50% of the non-matches the probe stimuli differed with respect to the shape, in the other 50 % with respect to the orientation. The non-matches probe stimuli were selected from the set of all possible shapes that were not used as a target in a given trial. After each response feedback was given (“Wrong”, “Correct” or “No Response”), which was followed by an inter-trial interval of 3 s. Analyses of presentation times included only correct trials. See Figure 3.1 for an illustration of the sequence of events at each trial. The experimental procedure lasted approximately 60 min for each participant. After the experiment, participants were asked, within a semi-structured interview, to freely recall the strategies they used to accomplish the task. The following questions were included: What strategies did you use for searching the targets in the PO and NPO conditions? What strategies did you use for encoding the objects in the PO and NPO conditions? What strategies did you use for memorising the objects in the PO and NPO conditions during the delay period?

3.1.2 Results and discussion

3.1.2.1 Accuracy at test

Overall, response accuracy for the WM task was high (on average 85% correct) and decreased with the number of shapes that needed to be encoded (from 93% correct, with WM load 1, to 75% correct, with WM load 5, in the pop-out condition and from 93% correct, with WM load 1, to 78% correct, with WM load 5, in the non pop-out condition) (Figure 3.2, upper panel). These changes were significant, as tested by the main effect of number of targets in a 2 × 5 repeated measures ANOVA [$F(4, 140) = 30.4$, $p < .001$, $\eta^2 = .47$]. Neither attentional demand nor the interaction between the two factors reached significance [$F(1, 35) = 0.55$, $p = .46$ and $F(4, 140) = 1.85$, $p = .14$, respectively]. Given that response accuracy was high and comparable across the different levels of attentional demand, it was concluded that the differences in the individually chosen presentation times indicated the differences in the processes required for successful WM encoding (see presentation time section). According to Luck and Vogel (1997), the load-dependent decrease in accuracy is likely to reflect the limited ability of maintaining information in visual WM rather than the limitations of the encoding process. Thus, this drop in performance should not have affected the processes of encoding information into WM, which was the main focus of the present analyses.

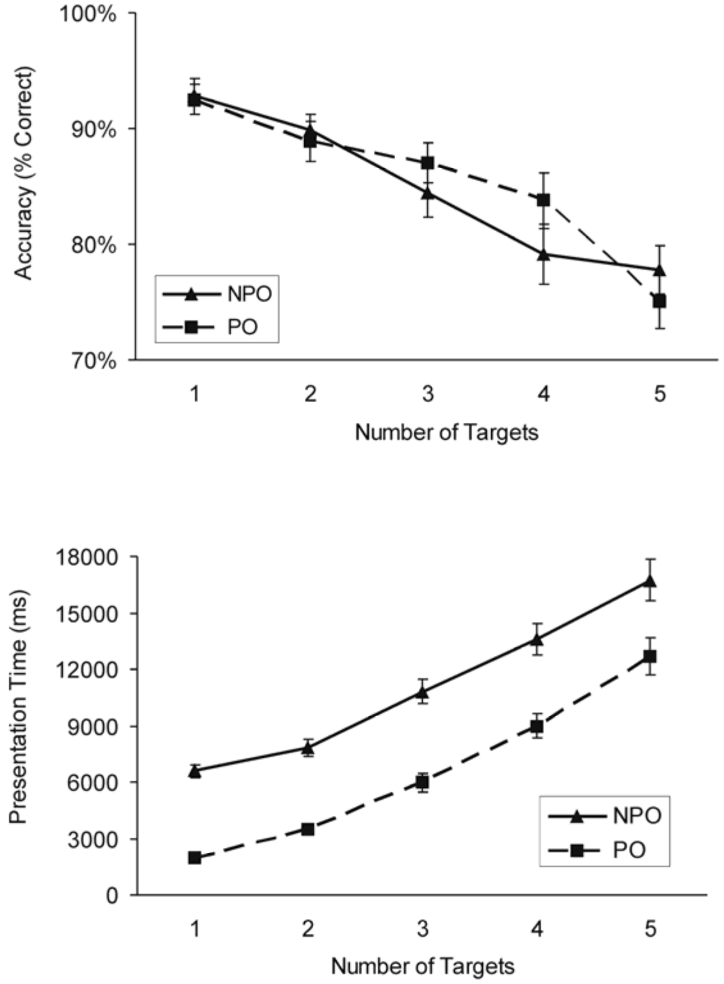


Figure 3.2 Results from Experiment 1. Mean response accuracy at test and mean presentation time as a function of number of targets and attentional demand (PO: pop-out; NPO: non pop-out). Vertical bars: the standard error of the mean.

3.1.2.2 Presentation time

Participants were slower without than with perceptual pop-out and the presentation time increased with the number of targets that needed to be encoded (Figure 3.2, lower panel). Repeated measures ANOVA, conducted with the same 2 x 5 design as for test performance, revealed significant main effects of attentional demand [$F(1, 35) = 288.41$, $p < .001$, $\eta^2 = .892$] and number of targets [$F(4, 140) = 116.62$, $p < .001$, $\eta^2 = .769$]. The increase in presentation time as a function of number of targets could be explained very well by a linear approximation and this was the case for both attentional-demand conditions (linear fits: $R^2 = .977$ for pop-out and

R^2 = .983 for non pop-out). Quadratic models explained only 2.3% (pop-out) and 1.3% (non pop-out) of additional variance. Therefore, the subsequent analyses of these data were made on the basis of linear approximation. On average, participants needed 2706 ms for encoding into WM each additional target shape in the presence, and 2606 ms in the absence, of perceptual pop-out. The relatively slow rates of these linear functions indicated that the process of encoding complex shapes into WM was difficult and capacity-demanding already in the pop-out condition.

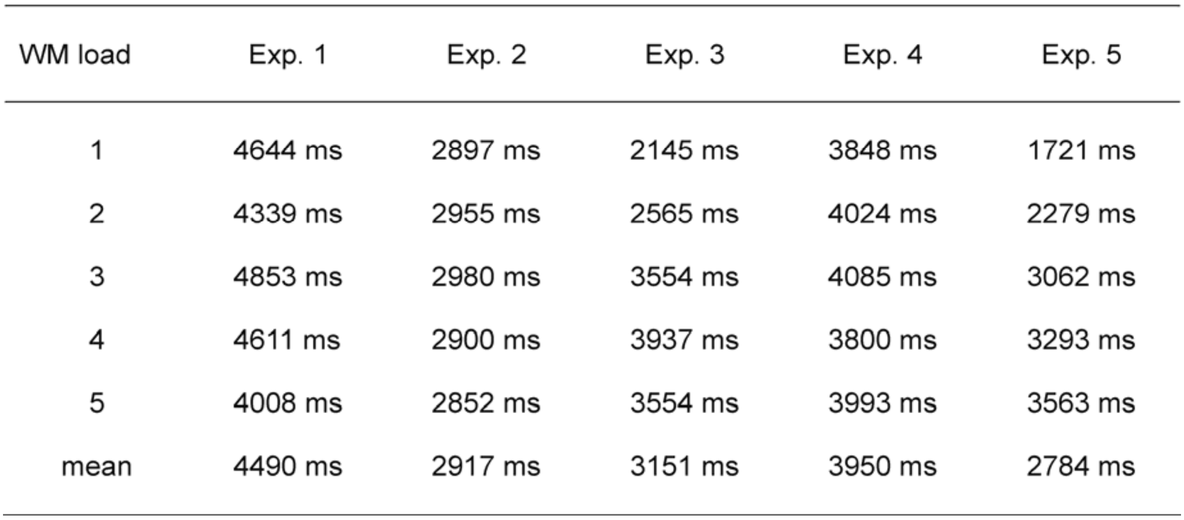
Importantly, the interaction between attentional demand and number of targets was not significant [*F*(4, 140) = 1.19, *p* = .32] indicating that the slopes relating the average presentation time to the number of targets were practically identical in the two attentional-demand conditions. The offset between the two slopes, i.e. the difference between non pop-out and pop-out conditions, ranged between 4008 ms and 4853 ms with an average of 4490 ms (see Table 3.1). Thus, the manipulation of attentional demand added considerable processing time but this time was constant across the number of targets. This result indicates that the manipulation of attentional mechanisms produced an effect on presentation time that was independent of the effect produced by the manipulation of WM load. Therefore, the results from Experiment 1 suggest that participants can achieve high memory performance despite the lack of pop-out but that this comes at the price of longer presentation time.

3.1.2.3 Reported encoding strategies

The majority of participants (32 of 36) reported that in the non pop-out condition they needed to use a “two-step” encoding strategy: In the first step they detected and memorised the locations of all the target items, encoding the associated shapes only in the second step. Three participants reported using a “search-and-encode” strategy in the non pop-out condition, encoding each target shape immediately after detecting a target item and making only one sweep through the array. One participant did not report any specific strategy. There were no significant differences in response accuracy and presentation times between participants subscribing to different encoding strategies [*F*(1, 33) = 0.25, *p* = .88

for presentations time; *F*(1, 33) = 0.06, *p* = .82 for accuracy]. However, due to vastly unequal numbers of participants in the two groups (32 vs. 3) this result should be taken with caution.

Table 3.1 Offsets in presentation time (Experiments 1, 3, 4 and 5) and counting time (Experiment 2) between non pop-out and pop-out conditions across WM load 1 to 5



WM load	Exp. 1	Exp. 2	Exp. 3	Exp. 4	Exp. 5
1	4644 ms	2897 ms	2145 ms	3848 ms	1721 ms
2	4339 ms	2955 ms	2565 ms	4024 ms	2279 ms
3	4853 ms	2980 ms	3554 ms	4085 ms	3062 ms
4	4611 ms	2900 ms	3937 ms	3800 ms	3293 ms
5	4008 ms	2852 ms	3554 ms	3993 ms	3563 ms
mean	4490 ms	2917 ms	3151 ms	3950 ms	2784 ms

3.2 Experiment 2 - Visual search only

In this experiment it was investigated whether the offset in presentation time between the two attentional-demand conditions, observed in Experiment 1, could be explained by visual search for target locations. To estimate the time to select target locations in this task, the same stimuli were presented as in Experiment 1 but participants were asked only to count the number of the target items in the array. This task required engagement of attention for determination of the targets, but not the processing of the background shapes, nor did it pose any demands on WM for shapes. Participants were again instructed to emphasise accuracy over speed in order to ensure that the criteria for determination of the target locations were similar to those in Experiment 1. If the offset in presentation time between pop-out and non pop-out conditions in Experiment 1 was due to the attention-demanding visual search, one should find a similar offset between pop-out and non pop-out conditions in the counting times.

3.2.1 Method

3.2.1.1 Participants

Fourteen students and employees of the University of Frankfurt M. (6 males, 8 females) participated in this study. Their mean age was 26.7 years (range: 19-44). Five participants had also taken part in Experiment 1.

3.2.1.2 Apparatus, stimuli, procedure, and design

Participants were required to count the target items in the same stimulus array as used in Experiment 1. After completing the count, participants indicated the search time by pressing the "return" button on the computer keyboard. After this button-press a question mark appeared in the centre of the screen prompting the participants to enter the number of the counted targets. Participants were instructed to emphasise accuracy over speed during the counting process and were informed that the time needed to enter the counted number of targets was irrelevant. After each response, the question mark disappeared and feedback ("Wrong", "Correct" or "No Response") was provided and followed by an inter-trial interval of 3 s. Only correct trials were included in the analyses of counting times. The experimental procedure lasted approximately 30 min for each participant.

A 2 × 5 within-subjects factorial design was used with two levels of attentional demand for target-distractor discrimination (pop-out and non pop-out) and five different counts (one to five targets).

3.2.2 Results and discussion

3.2.2.1 Accuracy at test

Overall, response accuracy was high (on average 97% correct). A repeated measures ANOVA revealed only a significant main effect of attentional demand [$F(1, 13) = 32.4$, $p < .001$, $\eta^2 = .71$] and neither the number of targets nor the interaction between the two factors reached significance [$F(4, 52) = 1.63$, $p = .21$, and $F(4, 52) = 0.70$, $p = .54$, respectively]. Participants counted target items less accurately in the non pop-out (on average 94.4% correct) than in the pop-out conditions (on average 98.9% correct) (Figure 3.3, upper panel). In the non pop-out condition the errors were more often underestimates (about 86%) than in the pop-out condition (about 66%), indicating that the increase in the similarity between targets and distractors increased the probability that a target item will be missed. The accuracy in the non-pop out conditions of the present task was higher than in a control version of the same task in which participants were asked to emphasise speed over accuracy [90.6% correct vs. 94.4% correct, $t(22) = 2.16$, $p < .05$; other results not shown for the control experiment]. Therefore, the results from the present task, in which accuracy was emphasised, indicated that participants followed this instruction. Thus, any increase in counting times in the non pop-out compared to the pop-out condition should be attributed to slower perceptual processing and should not be influenced by changes in speed-accuracy tradeoff across different perceptual conditions.

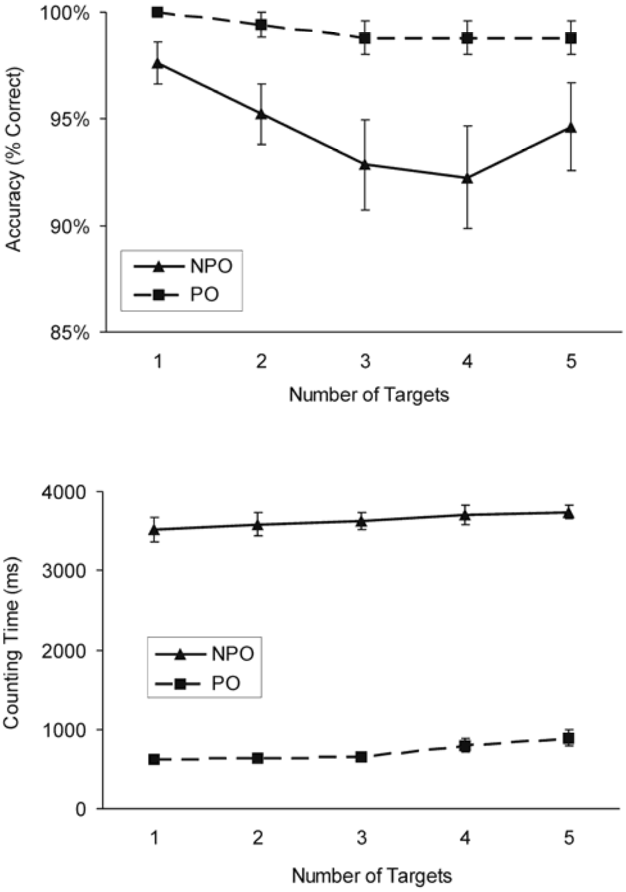


Figure 3.3 Results from Experiment 2. Mean response accuracy at test and mean counting time as a function of number of targets and attentional demand (PO: pop-out; NPO: non pop-out). Vertical bars: the standard error of the mean.

3.2.2.2 Counting time

Participants were slower in the non pop-out compared to the pop-out condition and counting times increased linearly with the number of targets (linear fits: $R^2 = .865$ for pop-out and $R^2 = .991$ for non pop-out) (Figure 3.3, lower panel). Participants needed on average 72 ms for counting each additional target item in the presence and 57 ms in the absence of perceptual pop-out. A repeated measures ANOVA revealed significant main effects of attentional demand [$F(1, 13) = 1292$, $p < .001$, $\eta^2 = 0.99$] and number of targets [$F(4, 52) = 8.4$, $p < .001$, $\eta^2 = .39$] but the interaction between attentional demand and number of targets was not significant [$F(4, 52) = 0.40$, $p = .81$]. Accordingly, the offset produced by the non pop-out condition compared to the pop-out condition was almost constant across the number of targets and was on average 2917 ms (range: 2852 ms to 2980 ms) (Table 3.1).

The similarity of the two slopes relating the counting time to the number of targets indicates that these slopes reflect mostly the time needed to perform counting operations, such as the verbal act of increasing the counter by one upon the selection of the target, and thus, that these operations are not directly related to visual search. Visual search processes should be reflected solely in the described offset in the counting times because participants needed to search always the entire arrays, regardless of the number of targets. In order to estimate the rate of this search, it was necessary to take into account the constant processing time that was not related to the sequential component of the search process (i.e., the intercept). This time could not be measured from the present data directly but, one can assume that this time largely corresponded to the counting times in the non pop-out condition. Thus, the search rate in the non pop-out condition was estimated simply by taking the mean offset of counting times between the two attentional-demand conditions and dividing this number by the number of elements in the array (nine). This resulted in an amount of 324 ms for scanning each of the nine locations. Although this time is higher than the search rates reported in standard inefficient visual search tasks (Treisman and Gormican, 1988; Duncan and Humphreys, 1989; Wolfe, 1998b), it is consistent with reports that search time increases with the complexity of the items (Alvarez and Cavanagh, 2004). The slower speed of the search in the present task than in standard visual search tasks cannot be simply explained by the need to select and count multiple targets because such tasks do not produce similar increases in response times (Horowitz and Wolfe, 2001). One can also exclude that the prolonged search time was a result of the instruction to emphasise accuracy because, in one control experiment (not reported here), 10 participants were instructed to count the target items as fast as possible and obtained only slightly faster search times (280 ms for scanning each of the nine locations). Another reason that visual search was so slow in the present experiment might be that attention tends to be locked to perceptual objects. When attention is voluntarily placed to one feature of an object it automatically spreads to other features of the same object (Duncan, 1984; Vecera and Farah, 1994; Scholl, 2001). Thus, when attention was placed on the features defining the targets in the present task, the attentional spotlight may have tended to spread over the other features of the objects, making it more difficult to scan multiple items simultaneously and/or judge whether this item was a target.

The important finding for the present study is that the offsets in the counting time between pop-out and non pop-out conditions (on average about 2.9 s) were smaller in the present experiment than the offsets in the presentation time in Experiment 1 (on average about 4.5 s) (see Table 3.1). These differences were statistically significant [$F(1, 48) = 13.42$, $p < .01$, $\eta^2 = .22$]. It was also tested whether this comparison might have been confounded by a perceptual learning effect that could have taken place for the five participants who took part also in Experiment 1. A comparison between the nine new and five old participants revealed no significant effect of the factor task exposure (new vs. old participants), and neither were significant the interactions of this factor with the factors attentional demand or WM load (repeated measure ANOVA; all *F*-values < 0.66, all *p*-values >.57). Therefore, it was concluded that serial search accounted for only about two thirds of the processing costs that arose due to the lack of pop-out in Experiment 1. These findings suggest a super-additive increase in the times for encoding and determination of target locations in the non pop-out condition, which is consistent with the idea of interference between attention and visual WM encoding. However, it was first investigated in Experiment 3 whether the remaining one third (or about 1.6 s) of the offset between pop-out and non pop-out conditions could be explained by repeated serial searches.

3.2.2.3 Reported search strategies

In the non pop-out condition all participants reported scanning the array serially, mostly from the upper left corner towards the lower right, and making one single sweep through the array. In the pop-out condition participants reported detecting the target items at a glance.

3.3 Experiment 3 - The role of repeated searches

The aim of this experiment was to assess whether repeated searches could explain the difference between the presentation time of the non pop-out and pop-out conditions of Experiment 1. Several studies have demonstrated that the temporary storage of previously searched target locations decays over time (Phillips, 1974; Irwin, 1992) and that participants sometimes need to repeat the search at target locations that they have already visited previously (Peterson et al., 2001). Repeated searches might have occurred in the non pop-out condition of Experiment 1 because i) multiple targets were presented, ii) participants had to perform a difficult additional task of encoding information into WM, and iii) participants needed to scan always the entire array, even when there was only one target. This was because they did not know how many targets would be presented at a given trial.

To assess the degree to which the lack of knowledge about the number of targets contributed to the time offsets between pop-out and non pop-out conditions and thus, to assess the extent of possible repeated searches, participants were informed in Experiment 3 about the upcoming number of targets prior to each experimental trial. This manipulation was expected to reduce the presentation time in the non pop-out conditions especially with small number of targets (1 or 2 targets). The main question then was whether this reduction would explain all of the difference between the search time, as determined from Experiment 2, and the presentation time in the non pop-out condition of Experiment 1. In this case one could conclude that repeated searches explained the non pop-out offset in

Experiment 1. This finding would support a search-and-encode strategy. Conversely, if an offset between pop-out and non pop-out conditions remained even in Experiment 3, where the number of targets was known beforehand, this would suggest that a particular cognitive process supporting WM encoding placed a particular demand on presentation time. It was suggested that it is the process of memorising all target locations.

A second question addressed by Experiment 3 concerned the role of verbal coding. The phonological store is highly efficient for serial recall and thus, participants tend to recode visually presented items into a verbal code (Baddeley, 2000). Indeed, in Experiment 1, the majority of participants reported creating their own verbal labels for the complex shapes. As the aim of the present study was to investigate visual attention and WM, it was necessary to assess the role of verbal encoding during the encoding of the shapes into WM. To this end, an articulatory suppression task was implemented that is known to reduce, albeit not completely eliminate, subvocal rehearsal and the phonological encoding of visually presented material (e.g., Murray, 1968; Besner et al., 1981; Baddeley, 2000, 2003). If presentation time and accuracy did not substantially differ between Experiment 1 without articulatory suppression and Experiment 3 with articulatory suppression, one could conclude that the encoding and storage of complex shapes depends to a high degree on visual processing of information.

3.3.1 Method

3.3.1.1 Participants

Sixteen students and employees of the University of Frankfurt M. (7 males, 9 females) participated in this experiment. The mean age of the participants was 24.6 years (range: 18-44). Six participants took part also in Experiment 2 and only one of them took part in Experiment 1.

3.3.1.2 Apparatus, stimuli, design, and procedure

The apparatus, stimuli, design, and procedure were the same as in Experiment 1, except for the following two differences. First, at the beginning of each trial a digit was presented at the centre of the screen, for 2 s. This digit indicated the number of target items that would be presented in the upcoming stimulus array. Second, the articulatory suppression task required participants to repeat aloud a syllable (“la”) throughout the duration of the trial.

3.3.2 Results and discussion

3.3.2.1 Accuracy at test

A repeated measures ANOVA revealed a significant main effect of number of targets [$F(4, 60) = 13.38$, $p < .001$, $\eta^2 = .47$] but no effect of attentional demand, [$F(1, 15) = 2.76$, $p = .12$]. The interaction between the two factors also reached significance [$F(4, 60) = 3.30$, $p < .05$, $\eta^2 = .18$] but the averages did not show any consistent relationships between the variables (Figure 3.4, upper panel) and explained only 18.1% of the variance in the dependent factor. Therefore, this interaction was not used for further interpretation of the results.

These results are highly consistent with those observed in Experiment 1, showing that response accuracy decreases with the number of targets to be remembered but does not depend on the attentional demand condition. Also, participants were about equally accurate as in Experiment 1 (on average 82% correct, range: 71% to 95%, in Experiment 3; on average 85% correct, range: 75% to 93%, in Experiment 1) and there were no significant differences between these two experiments [$F(1, 50) = 1.51$, $p = .14$, for pop-out; $F(1, 50) = 0.74$, $p = .46$, for non pop-out]. These results indicate that articulatory suppression did not affect participants’ ability to memorise the shapes. This finding suggests that in the present task it was not necessary to recode the visual information into a verbal form in order to achieve good memory performance. This conclusion was also supported by the presentation time data (see presentation time section).

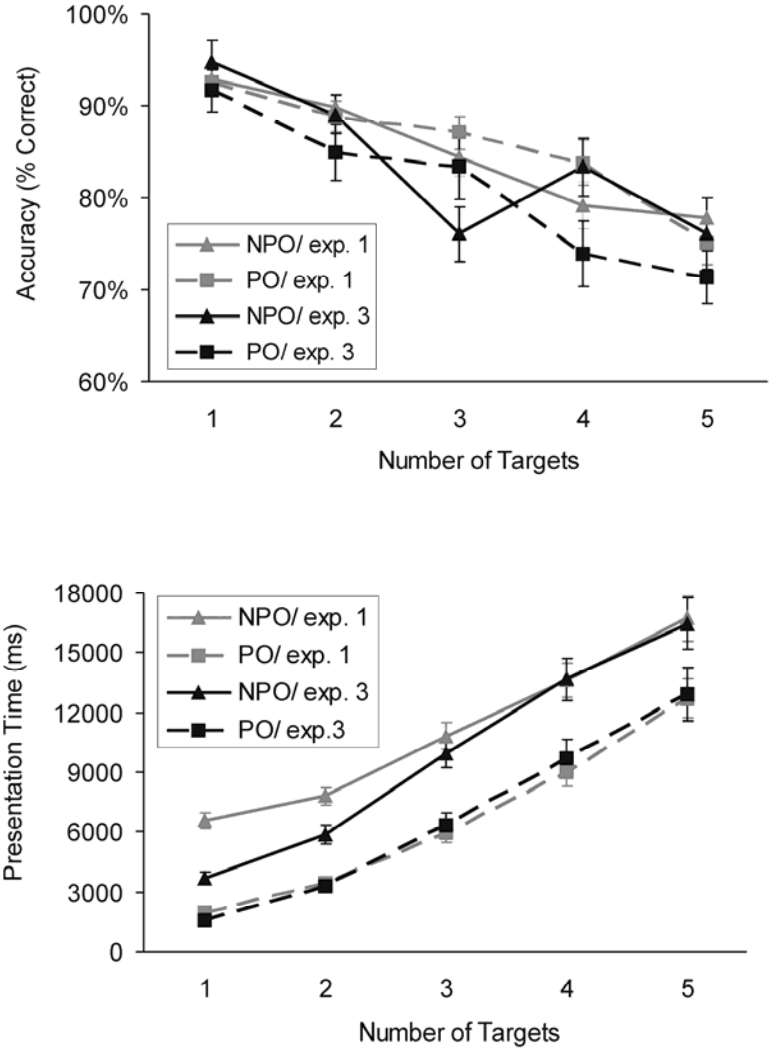


Figure 3.4 Results from Experiment 3 compared to the results from Experiment 1. Mean response accuracy at test and mean presentation time as a function of number of targets and attentional demand. (PO: pop-out; NPO: non pop-out). Vertical bars: the standard error of the mean

3.3.2.2 Presentation time

Similarly to Experiment 1, participants were slower in the non pop-out than in the pop-out condition [*F*(1, 15) = 127.9, *p* < .001, *η*² = .89]. Presentation time also increased linearly with the number of targets that needed to be encoded into WM in both the pop-out and the non pop-out conditions (linear fits: R^2 = .989 for pop-out and R^2 = .992 for non pop-out), and these increases were significant [*F*(4, 60) = 70.43, *p* < .001, *η*² = .82] (Figure 3.4, lower panel). The slope relating the average presentation time to the number of targets was steeper for non pop-out (3338 ms) than for pop-out (2918 ms), leading to the significant interaction between attentional demand and number of targets [$F(4, 60) = 4.84$, $p < .01$, $\eta^2 = .24$]. In the pop-out condition, these slopes were not significantly different from Experiment 1 [$t(50) = 0.53$, $p = .60$], whereas in the non pop-out condition the average difference of 732 ms approached statistical significance [$t(50) = 1.67$, $p = .10$]. The offset in the presentation time between pop-out and non pop-out conditions increased from 2145 ms, for one target, to 3554 ms, for five targets (Table 3.1). Thus, as predicted, the presentation time was reduced in the non pop-out conditions with smaller numbers of targets as compared to the presentation time in Experiment 1 (in particular with one and two targets, see Figure 3.4, lower panel). With the memory loads 4 and 5 presentation time was indistinguishable across the two experiments [$t(50) = 0.13$, $p = .90$], and this was the case for each number of targets in the pop-out condition [$F(1, 50) = 0.04$, $p = .85$][1].

Next it was investigated whether the presentation time in the non pop-out condition equaled the sum of the encoding time in the pop-out condition plus the time needed to select the target location(s) by a single-sweep search. If this was the case for any of the five memory loads, evidence would be provided that, for that load condition, participants first searched and then immediately encoded the information into WM. To conduct this analysis, it was necessary to estimate first the expected number of array items that needed to be searched for the presence of a target at each WM load, k, which, if the targets are positioned randomly, is given by the following equation: $k = A - A / (N + 1)$ (equation 1), where A and N represent the array size and the number of targets, respectively. For N = 1 to 5 targets in an array of $A = 9$ items, the expected numbers of items searched were 4.5, 6, 6.75, 7.2, and 7.5. These values were then multiplied by the expected

---

[1] The differences in performance between Experiments 1 and 3 should not be due to the articulatory suppression task used only in Experiment 3. This is because object naming would be the most likely advantage of verbalisation and this could be used equally well in the pop-out and in the non pop-out condition. Therefore, participants would have to be better also in the pop-out condition of Experiment 3 compared to Experiment 1. However, the results showed that, in the pop-out condition the performance across these two experiments was identical. This was also the case in the most difficult condition (non pop-out with five targets) when investigated individually. Although articulatory suppression on verbal memory recall is not necessarily dramatic (Baddeley, 2000, 2003), this similarity in presentation time between Experiment 1 without articulatory suppression and Experiment 3 with articulatory suppression indicates that encoding was based to a considerable degree on visual processing.

search time per single item, which according to the results from Experiment 2, was 324 ms. The resulting theoretical values are plotted in Figure 3.5A together with the offset in the presentation time between non pop-out and pop-out conditions obtained empirically.

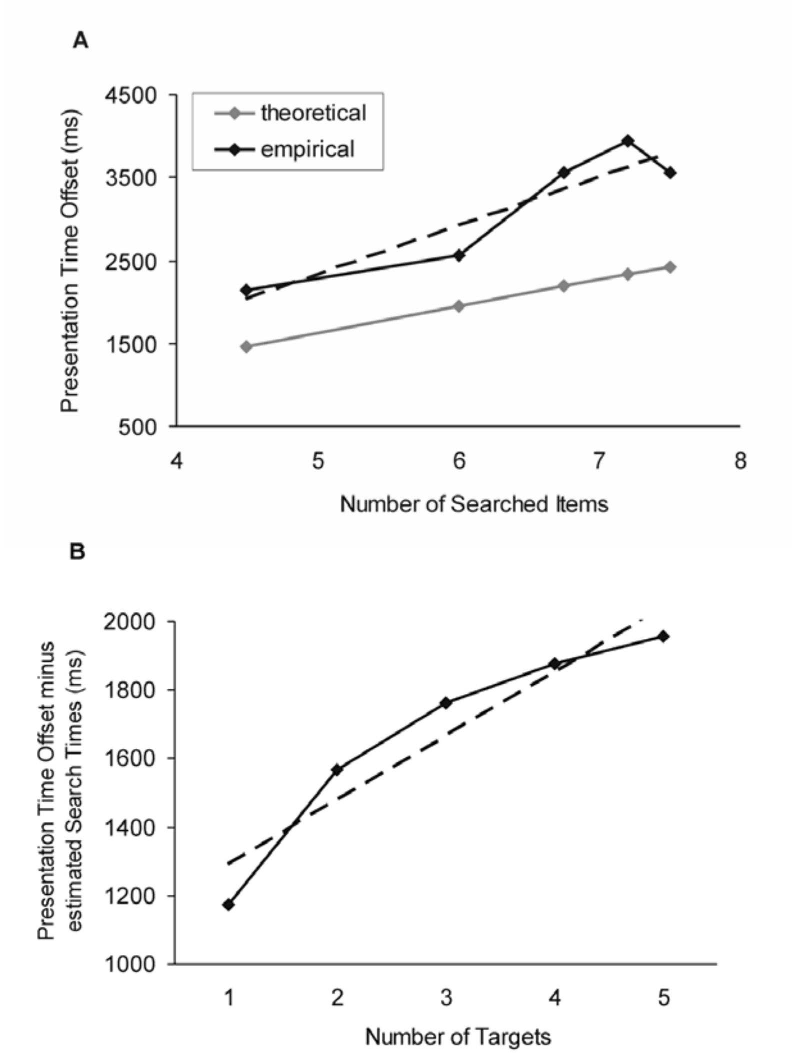


Figure 3.5 A. Empirically obtained offset in the presentation time produced by the lack of pop-out in Experiment 3 and theoretically predicted offset assuming a single-sweep search. X-axis: the average numbers of items that needed to be searched if one to five targets are presented in the array. Dashed line: linear fit (parameters reported in the main text). B. The difference between the two offsets in A, expressed as a function of number of target items. Dashed line: linear fit (parameters reported in the main text).

One can see that the theoretical and empirical values do not match. The empirical offset in the presentation time was, already with WM load 1 (i.e., 4.5 items searched), considerably larger than that predicted by a single-sweep search. This difference increased further with the higher WM loads as the slope with which the empirical values increased was much steeper than expected by simple search for target items (585 vs. 324 ms; 81% higher slope; linear fit: R^2 = .86). The difference between the two, expressed as a function of the number of target items, accumulated to over 1.8 s with WM load 5 (Figure 3.5B) whereas the large positive intercept of the resulting function (slope 188 ms, intercept 1104 ms, linear fit: R^2 = .903) indicated that with the lack of pop-out participants needed a constant time of 1104 ms irrespectively of the number of targets. These results suggest that, simple serial search does not account for the slowdown in the presentation time caused by the lack of pop-out even when the participants know the number of targets presented in the array and this result holds for all five memory load conditions.

Taken together, the results of Experiments 1 to 3 indicate an excess in the costs on presentation time produced by the lack of perceptual pop-out, and this cost cannot be explained fully by simple visual search or by repeated searches for targets. Thus, the presentation time does not simply represent a sum of the two task components and hence, are not consistent with a search-and-encode strategy that would interleave the search process with the WM encoding. Instead, the results revealed a super-additive increase of the times for encoding and determination of target locations, indicating that participants used another, time consuming strategy. One possibility, as suggested by the finding that WM and attention interfere (Smyth and Scholey, 1994; Awh et al., 1998; Jolicœur and Dell'Acqua, 1998, 1999; Oh and Kim, 2004; Woodman and Luck, 2004; Barrouillet et al., 2007) as well as by the subjective reports of our participants, is that they invested the additional time in the process of memorising all target locations prior to encoding their shapes. This two-step strategy was investigated more directly in Experiments 4 and 5.

3.3.2.3 Reported encoding strategies

All 16 participants reported using the same two-step strategy as described by the majority of participants in Experiment 1.

3.4 Experiment 4 - Visual search and encoding of *locations* into WM

In this experiment and the next I explicitly tested the strategy that was reported by the participants during the debriefing procedure. The majority of participants reported that, in the non pop-out condition of Experiments 1 and 3, they memorised first the locations of all the targets and only then did they encode the shapes of the associated objects. To search for experimental evidence supporting this claim, participants were presented with the same stimuli as in Experiment 1 but asked to memorise the locations of the target items only. If participants used the reported strategy, the time they need to search and memorise the target locations (e.g. the offsets in the presentation time between non pop-out and pop-out conditions) should correspond to the presentation time offsets between non pop-out and pop-out conditions in Experiment 1.

3.4.1 Method

3.4.1.1 Participants

Sixteen students and employees of the University of Frankfurt M. (8 males, 8 females) participated. The mean age was 27.1 years (range: 19-39). Eight participants took part also in Experiment 1 and two of them took part also in Experiment 2.

3.4.1.2 Apparatus, stimuli, procedure, and design

The apparatus, stimuli, procedure, and design were the same as those in Experiment 1, except for the following two differences. Participants were instructed to determine and memorise the locations of the target items only and thus, to ignore the shapes of the associated objects. In order to probe WM for target locations, the original stimulus array was presented at the test phase without the centre items and with one of the shapes missing. Participants needed to indicate whether the location of the missing shape matched one of the target locations. After each response feedback was given (Figure 3.6).

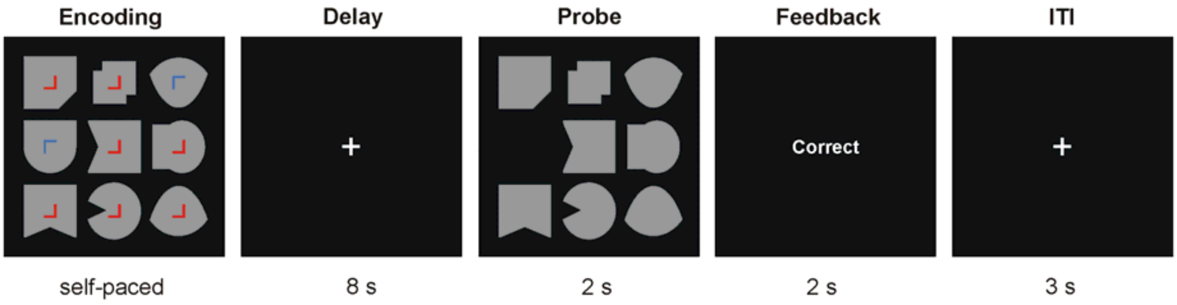


Figure 3.6 The procedure used in Experiment 4. Participants determined the locations of the target items and memorised these locations. After an interval of 8 s, participants judged whether the location of the missing item in the test array matched one of the target locations. ITI: Inter-trial interval.

3.4.2 Results and discussion

3.4.2.1 Accuracy at test

Overall, response accuracy was again high (on average 93% correct). A repeated measures ANOVA revealed only a significant main effect of number of targets [$F(4, 60) = 5.75$, $p < .01$, $\eta^2 = 0.27$]. Neither attentional demand nor the interaction between the two factors was significant [$F(1, 15) = 0.01$, $p = .96$ and $F(4, 60) = 0.68$, $p = .57$, respectively]. Thus, similarly to Experiment 1, response accuracy decreased with the number of targets whose locations needed to be encoded and, again, did not differ between pop-out and non pop-out conditions (Figure 3.7, upper panel). Participants responded more accurately in Experiment 4 than in Experiment 1 [$F(1, 50) = 9.76$, $p < .01$, $\eta^2 = .16$, for pop-out and $F(1, 50) = 10.57$, $p < .01$, $\eta^2 = .18$, for non pop-out; on average 93% correct, range: 89% to 97% in Experiment 4; on average 85% correct, range: 75% to 93 % correct in Experiment 1], indicating that their memory for locations was better than their memory for shapes. The eight participants who took part also in Experiment 1 were no more accurate than the eight new participants. Instead, it was the new participants who tended to be more accurate (95% vs. 90% correct); however, the difference did not reach the level of significance [$F(1, 14) = 4.18$, $p = .06$, $\eta^2 = .23$]. Also, task exposure did not interact with attentional demand or WM load (repeated measures ANOVA; all F-values < 2.1, all p-values > .12). Therefore, in this experiment there was no evidence that improvement due to perceptual learning had taken place among participants who took part in multiple experiments of the study.

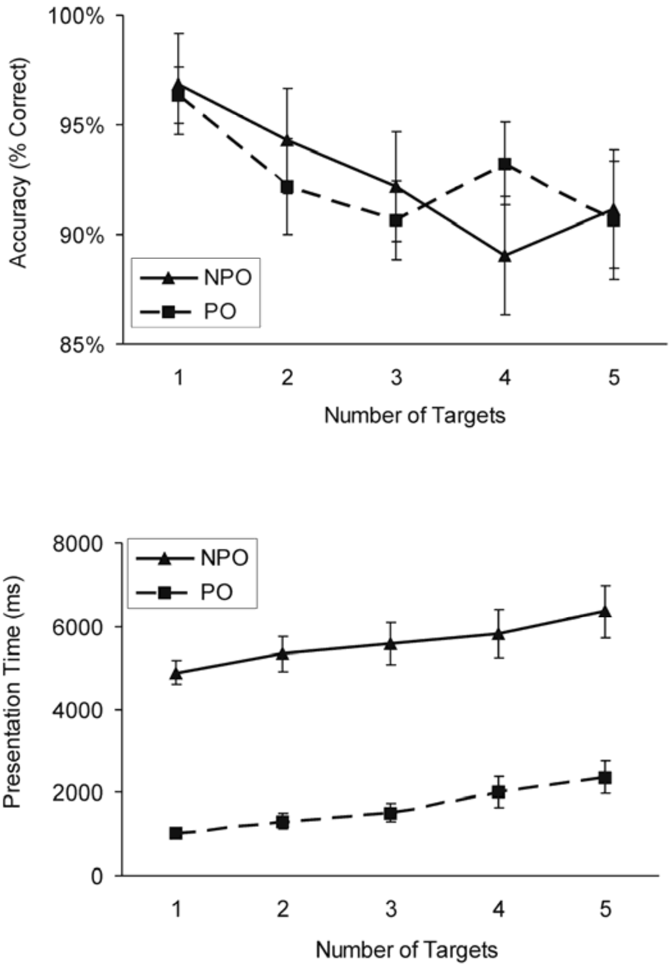


Figure 3.7 Results from Experiment 4. Mean response accuracy at test and mean presentation time as a function of number of targets and attentional demand (PO: pop-out; NPO: non pop-out). Vertical bars: the standard error of the mean

3.4.2.2 Presentation time

Similarly to Experiment 1, participants were slower without than with perceptual pop-out [$F(1, 15) = 193.9$, $p < .001$, $\eta^2 = .93$]. Presentation time increased linearly with the number of targets that needed to be encoded into WM, in both, the pop-out and non pop-out conditions (linear fits: $R^2 = .976$ for pop-out and $R^2 = .978$ for non pop-out), and these changes were significant [$F(4, 60) = 11.23$, $p < .001$, $\eta^2 = .43$] (Figure 3.7, lower panel). The interaction between attentional demand and number of targets was not significant [*F*(4, 60) = 0.54, *p* = .71] indicating again almost identical slopes relating the average presentation time to the number of targets across the two levels of attentional demand.

The slopes were much shallower in the present experiment than in Experiment 1. On average, participants needed 342 ms for encoding each additional location of a target item in the absence and 336 ms in the presence of perceptual pop-out (compared to 2606 ms and 2706 ms for encoding shapes in Experiment 1). Thus, locations were encoded much faster than shapes. A repeated measures ANOVA with the factors attentional demand, WM load, and task exposure (new vs. old participants) revealed no significant effect either for the factor task exposure or for its interaction with the other two factors (all *F*-values < 1.11, all *p*-values > .31). Thus, again no evidence that improvement due to perceptual learning had taken place among the eight participants who took part also in Experiment 1 was found.

Similarly to Experiment 1, the offsets between pop-out and non pop-out conditions were practically constant across different WM loads. Although the offsets were smaller in magnitude compared to those in Experiment 1 (*M* = 3950 ms, range: 3800 ms to 4085 ms in Experiment 4 compared to *M* = 4490 ms, range: 4008 ms to 4853 ms in Experiment 1) these differences were not significant [*F*(1, 50) = 1.50, *p* = .23] (Table 3.1). Thus, the results indicate additivity between the presentation time in the pop-out condition of Experiment 1 and the time offset between pop-out and non pop-out conditions in Experiment 4. In other words, when the time needed to encode the shapes is taken into account, the lack of pop-out caused similar effects on presentation time in Experiments 4 and 1. Therefore, the time needed to memorise the locations seems to be a reasonable explanation of the time offset between pop-out and non pop-out conditions in Experiment 1.

3.4.2.3 Reported encoding strategies

The majority of participants (15 of 16) reported integrating the target locations into one or two perceptual representations that could be described either as a spatial template, a shape composed of the individual locations, or as a chunk. One participant reported encoding discrete locations, one after another, without a particular perceptual organisation.

3.5 Experiment 5 - The role of repeated searches

When informed about the upcoming number of targets in Experiment 3, participants also reported using a two-step strategy. Apparently, they memorised the locations of all targets first, and only then encoded the shapes into WM. These reports, together with the results from Experiment 4, suggest that, if participants were informed about the number of target locations, the times needed to memorise those locations might explain the peculiar offsets in the presentation time between pop-out and non pop-out conditions found in Experiment 3. Therefore, in Experiment 5, participants were informed prior to each trial about the number of target items in the upcoming stimulus array, as in Experiment 3, and asked to remember the locations of the targets only, as in Experiment 4. The analysis was similar to that used in Experiment 3.

3.5.1 Method

3.5.1.1 Participants

Ten students and employees of the University of Frankfurt M. (4 males, 6 females) participated. The mean age was 25.2 years (range: 20-33). None of the participants took part in any of the previous experiments.

3.5.1.2 Apparatus, stimuli, procedure, and design

The apparatus, stimuli, procedure, and design were the same as in Experiment 4, except for the following two differences. First, this experiment involved the procedure from Experiment 3 to inform participants about the number of upcoming targets at the beginning of each trial. Second, the articulatory suppression task was implemented.

3.5.2 Results and discussion

3.5.2.1 Accuracy at test

As in the previous experiments, response accuracy was high at both levels of attentional demand (on average 94% correct), decreased as a function of WM load [$F(4, 36) = 3.33$, $p < .05$, $\eta^2 = .27$] but did not depend on the attentional demand condition [$F(1, 9) = 0.35$, $p = .58$]. As in Experiment 3, the interaction between the two factors was also significant [$F(4, 36) = 3.30$, $p < .05$, $\eta^2 = .27$] (graph not shown). Response accuracy in the present experiment did not differ from that obtained in Experiment 4 [$F(1, 24) = 0.75$, $p = .39$, for pop-out and $F(1, 24) = 0.47$, $p = .83$, for non pop-out (on average, 94% correct, range: 84% to 99% correct, in Experiment 5; on average, 93% correct, range: 89% to 97% correct, in Experiment 4]. As in Experiment 3, the finding that articulatory suppression did not impair participants' ability to memorise the locations indicates that the memory of locations was based to a high degree on visual processing. This conclusion was further supported by the lack of significant differences between the presentation time obtained in Experiments 5 and 4 (see presentation time section).

3.5.2.2 Presentation time

Similarly to Experiment 3, participants were slower without than with perceptual pop-out [$F(1, 9) = 145.42$, $p < .001$, $\eta^2 = .94$]. Presentation time again increased linearly with the number of targets that needed to be encoded into WM in both the pop-out and non pop-out conditions (linear fits: $R^2 = .976$ for pop-out and $R^2 = .987$ for non pop-out) and these changes were highly significant [$F(4, 36) = 66.58$, $p < .001$, $\eta^2 = .88$] (Figure 3.8A). As it would be expected from the results of Experiment 3, the slope relating the average presentation time to the number of targets was steeper for non pop-out (681 ms) than for pop-out (229 ms), leading to a significant interaction between number of targets and attentional demand [$F(4, 36) = 24.65$, $p < .001$, $\eta^2 = .73$]. The offset between the pop-out and non pop-out conditions increased gradually from 1721 ms, for one target, to 3563 ms, for five targets. In the pop-out conditions the presentation time did not significantly differ from those in Experiments 4 in which articulatory suppression was used [$F(1, 24) = 2.71$, $p = .11$]. Also, no difference was found when only the responses given in the most difficult condition (non pop-out with five targets) were investigated [*t*(24) = 1.38, *p* = .18].

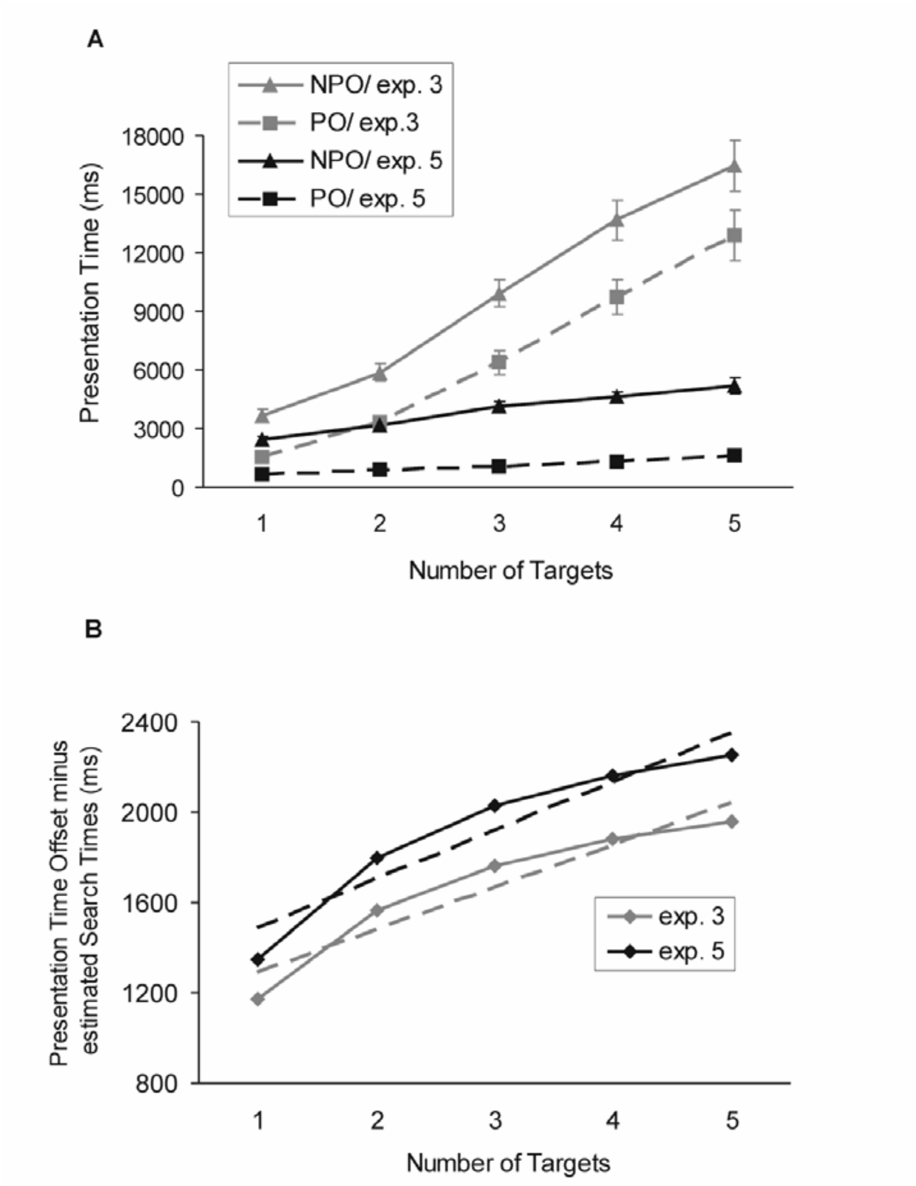


Figure 3.8 Results from Experiment 5 compared to the results from Experiment 3. A. Mean presentation time as a function of number of targets and attentional demand (PO: pop-out; NPO: non pop-out). Vertical bars: the standard error of the mean. B. Presentation time offset minus estimated search times expressed as a function of number of target items. Dashed lines: linear fit (parameters reported in the main text).

Most importantly, the offsets in Experiment 5 did not significantly differ from those obtained in Experiment 3 [range: 2145 to 3937 ms; *F*(1, 24) = 0.84, *p* = .37] (Table 3.1). Next, the degree to which the presentation time offset between non pop-out and pop-out conditions could be explained by the model of repeated serial searches discussed and tested in Experiment 3 was investigated. To this end, I estimated the amount of time spent on each target that could not be explained by the visual search by conducting analyses similar to those described in Experiment 3. Thus, the estimated offsets in the serial search time (obtained in Experiment 2) were subtracted from the offsets in the presentation time obtained in the present experiment. To compare directly the results from the present experiment with those from Experiment 3, the results shown in Figure 3.8B contain also those from Figure 3.5B (Experiment 3). This comparison revealed high similarity of the results. As in Experiment 3, again a large positive intercept of the resulting function was found, which indicates that with the lack of pop-out participants needed a constant time of 1269 ms irrespectively of the number of targets and additional 216 ms to process each target item (slope 216 ms, intercept 1269 ms, linear fit: R^2 = .903). Neither the slopes nor the intercepts differed significantly from the corresponding ones obtained in Experiment 3 [*t*(24) = 0.26, *p* = .797 for slope; *t*(24) = 0.25, *p* = .798 for intercept]. Therefore, as across Experiments 1 and 4, the presentation time was also highly consistent across Experiments 3 and 5.

These results indicate that memory for locations plays an important role in the present paradigm even when repeated searches for the relevant locations are prevented. The time needed to encode the shapes of complex objects into WM in the non pop-out condition corresponds closely to the sum of the time needed to encode the shapes in the pop-out condition and the time needed to memorise the locations of the targets. This behavioral evidence is highly consistent with the subjective reports on the two-step strategy obtained during the debriefing procedures in Experiments 1 and 3.

3.5.2.3 Reported encoding strategies

The majority of participants (9 of 10) reported using the same chunking strategy as described by the majority of participants in Experiment 4.

3.6 Discussion of experiments 1-5

One goal of the behavioural study was to investigate whether and how participants can encode complex objects into WM while engaging spatial attention for a visual search task. Attentional demand and WM load were manipulated by changing either search efficiency in the visual search component of the task or the number of shapes to be encoded in the memory component of the task. Based on the participant-chosen presentation time it was sought to isolate the processes participants used to perform the task successfully.

The data provided evidence for the two-step encoding strategy. In the non pop-out condition of Experiment 1, participants required longer presentation time than what would be expected based on the simple addition of the search time (as measured in Experiment 2) and the time needed for WM encoding. Experiment 3 ruled out that repeated searches of the same location could explain the additional costs on presentation time in the non pop-out condition. Experiments 4 and 5 demonstrated a close match between the times participants needed to memorise the locations only, and the differences in the presentation time between pop-out and non pop-out conditions when participants needed to memorise the shapes of the targets. This match remained well across different memory loads even when repeated searches at relevant target locations were strongly reduced. These results were highly consistent with the participants' subjective reports about the strategy that they used to achieve the objectives of the task.

It might be argued that other processes than those related to the memorising of target locations contribute to the additional time cost in the non pop-out condition. WM suffers from a time-related decay as soon as attention is switched away and captured by concurrent activities (Barrouillet et al., 2007). Thus, the additional time cost in the non pop-out condition might be also related to an increased need to interleave the attention-demanding visual search with the maintenance of the already encoded shapes. This possibility was not directly tested in this study. However, the results suggest that the rehearsal of complex objects was more demanding than the rehearsal of locations. Therefore, it can be expected that the need to interleave the search with the maintenance should be higher when shapes needed to be memorised compared to locations. The present findings did not support this prediction as the additional costs on presentation time in the non pop-out conditions were comparable across WM domains. Taken together, the experimental data, in combination with subjective reports, seemed to be most consistent with the two-step strategy that involves memorising the locations of all the targets before memorising the associated shapes.

Why would participants need to memorise target locations? One possibility is that this is how they cope with the interference between WM and attention that would otherwise take place. Interference between selective attention and the storage of information in spatial WM has been well documented and interpreted in terms of common cognitive resources shared by these processes (Smyth and Scholey, 1994; Awh et al., 1998; Oh and Kim, 2004; Woodman and Luck, 2004). The present findings suggest that interference between selective attention and WM encoding may not be restricted to the spatial domain, unlike the findings for WM maintenance (Woodman et al., 2001; Oh and Kim, 2004). Instead, it seems likely that in the non pop-out condition of the present experiment, interference occurred between the attentional resources needed for determination of the target locations (Treisman and Gormican, 1988; Treisman, 1998) and the WM resources needed for encoding of targets' shapes.

What is the common mechanism that is required by the visual search and the encoding of object information into WM? Selective attention seems to be that mechanism. Representations of spatial locations are maintained in WM by keeping the spotlight of attention at these locations (Awh et al., 1998; Awh and Jonides, 2001). According to this account, selective attention is recruited in the service of a rehearsal-like function that maintains information active in WM and prevents its decay. A similar mechanism might come into play during WM encoding because of the necessity to verify the success of information transfer into WM, especially when multiple objects are presented simultaneously at different locations and need to be encoded. Another reason why selective attention should be involved both in the visual search and in WM encoding is related to the stimulus complexity. Complex objects, as those used in the present task, consist of multiple elementary features. Different features are bound into an integrated objects through focused attention (Treisman and Gelade, 1980) and the storage of such information in WM requires capacity-limited attentional mechanisms as well (Wheeler and Treisman, 2002).

The implication of the behavioural study is that the memory for locations may provide a coping mechanism for interference between search and memory. In the pop-out condition the unique elementary features attract the spotlight of attention by “automatic” bottom-up mechanisms (Treisman and Gelade, 1980). Along similar lines, the locations in the non pop-out condition, once memorised, might guide the attentional spotlight in an automatic-like fashion. Consistent with this notion, it has been proposed that in order to search for multiple targets efficiently, participants use spatial WM to keep track of identified targets (Horowitz and Wolfe, 2001).

It is possible that this storage of target locations was based on visual LTM because LTM is, in general, a tool for coping with capacity limitations. LTM is used during the chunking processes in WM (short-term memory) tasks (Miller, 1956; Chase and Simon, 1973; Cowan, 2001; Gobet et al., 2001) and is responsible for the development of skills and expertise in general (Shiffrin and Schneider, 1977; Hasher and Zacks, 1979; Chase and Ericsson, 1981). The main advantage of maintaining information in LTM, as opposed to WM, is that such storage does not seem to rely on limited-capacity resources (Phillips and Cristie, 1977; Ericsson and Kintsch, 1995). It has been recently shown that, in a task similar to the present one, participants can readily store target locations into LTM when they need to memorise a number of locations that exceeds greatly the capacity of visual WM for such locations (Nikolić and Singer, 2007).

Real-life situations in which interference between WM and attention occurs may require similar coping mechanisms. One example of a cluttered visual scene, in which not only serial search but also other forms of spatial processing are needed, is map reading (e.g., Thorndike and Hayes-Roth, 1982; Garden et al., 2002). To find a desired route, one needs to identify first the key locations (e.g., the origin and destination), and only then the rest of the route can be explored. If the route is non-trivial (multiple locations in-between and turns are involved), there might be at first interference between the memory for the examined part of the route and the search for the rest of the route. However, over time, as the route is being studied, knowledge will be acquired (including information about the sequence of landmarks along the route or about metric distances and angles that are integrated into a configural cognitive map), and the access to the route should become gradually easier. Similar processes should apply to other activities that involve visual WM and attention such as navigating through complex technical drawings or within one’s environment (Garden et al., 2002; Foo et al., 2005; van Asselen et al., 2006). In general, memory for locations might be the very mechanism that allows us to extract and encode relevant information from complex visual scenes when obvious cues that automatically draw attention are not available.

Chapter 4 - Common neural substrates for encoding into visual WM and selective attention

This dissertation aimed at testing the hypothesis that the capacity limitation of visual WM is due to common limited-capacity neural resources shared by visual WM and selective attention. In fMRI Experiment 1, visual search was combined with delayed discrimination of complex objects. The demands on selective attention and WM encoding were independently modulated. This design allowed identifying the brain regions that were selectively responsive to either attentional demand or the encoding into visual WM and those involved in both processes.

4.1 Materials and methods

4.1.1 Participants

18 healthy participants (nine females, mean age 28.2 ± 6.6, range: 20 - 44) were recruited from an academic environment. Participants reported normal or corrected-to-normal visual acuity, normal colour vision, and no history of neurological or psychiatric illness. The study was approved by the local ethics committee. All participants gave written informed consent.

4.1.2 Stimuli, task, and procedure

The stimuli and task were the same as in the behavioural study (Experiment 1), except for the following differences. To provide constant visual stimulation across experimental conditions the search array was presented for the same amount of time in each trial. Thus, each 30-s trial began with the presentation of the search array for 8 s. Participants needed to memorise the *objects* marked with an L in 90° orientation (target items) while the objects associated with Ls of other orientations could be ignored (distractor items). The search array contained either one or three targets (WM load 1 and 3). After an 8-s delay interval, a probe that consisted of a single object appeared for 2 s at the centre position of the array. Participants responded with a left- or right-hand button press to indicate whether the probe did or did not match in the form and orientation of one of the memorised objects. They were instructed to respond as accurate and as fast as possible. Feedback was not provided. Half of the trials were matches. The inter-trial interval lasted 12 s (Figure 4.1). Each fMRI run (four runs per session) included six iterations of each of the four trial types (load1/PO, load3/PO, load1/NPO, load3/NPO). Pop-out and non pop-out search conditions were presented in separate blocks of six trials (two blocks for each condition per run) in a pseudo-randomised order across runs. WM load conditions were fully randomised within each block.

The duration of the encoding phase was determined based on the results of the behavioural experiment 1 (see Figure 3.2, lower panel). With regard to the experimental conditions implemented in the fMRI experiment (load1/PO, load3/PO, load1/NPO, load3/NPO) the mean presentation times ranged from 1.9 s in the easiest condition (load1/PO) to 10.8 s in the most difficult search condition (load3/NPO). It was reasoned that an encoding phase of 8 s should be long enough to enable successful encoding of the objects into WM, both in the pop-out and the non pop-out conditions. Assuming that visual search difficulty would not have an impact on processing after the array had disappeared (during maintenance and retrieval of the WM task) it was expected that performance at test would not differ between pop-out and non pop-out search conditions in the fMRI experiment.

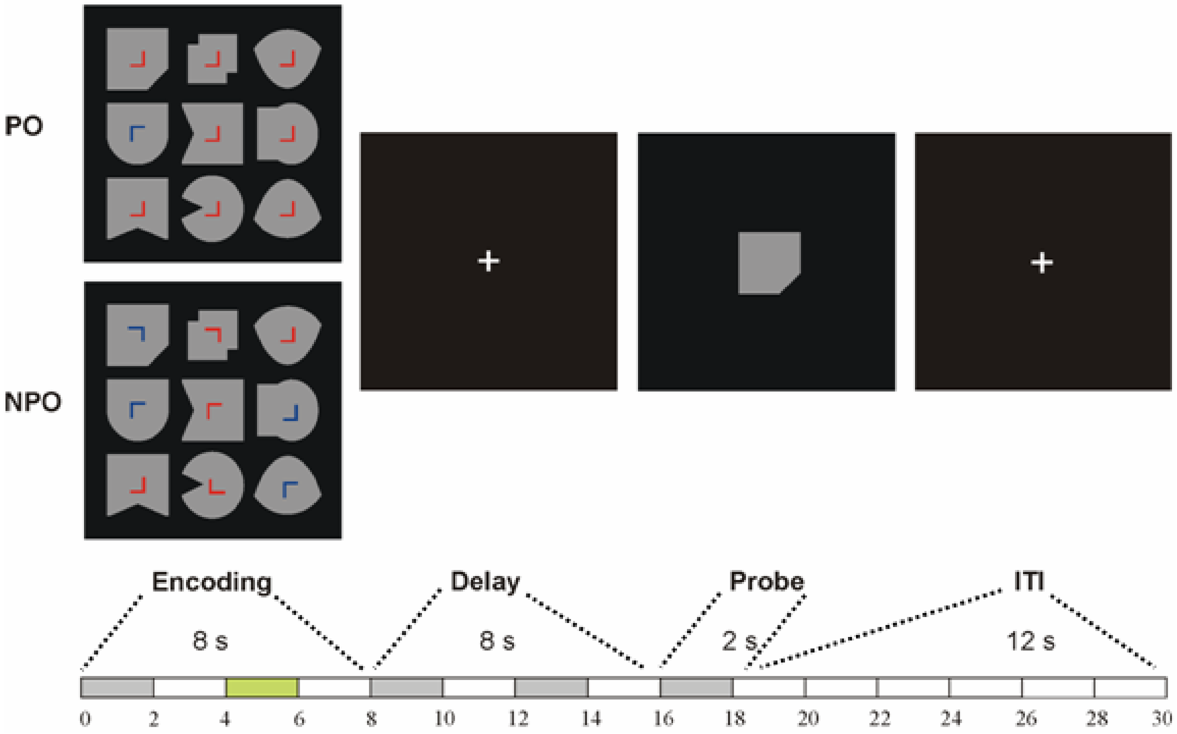


Figure 4.1 Trial design used in fMRI experiment 1. The search array was presented for 8 s and WM load was either 1 or 3. The analysis focused on the late encoding predictor (green bar, grey: additional predictors). PO: Pop-out, NPO: non pop-out, ITI: Intertrial interval.

4.1.3 Image acquisition and analyses

Anatomical three-dimensional T1-weighted images (voxel size: 1.00 x 1.00 x 1.00 mm^3) and functional images were acquired on a 3 T Magnetom Trio scanner (Siemens Medical Systems, Erlangen, Germany) equipped with a standard head coil. Functional images were collected using 34 slices (3 mm thickness with 3.4 x 3.4 mm in-plane resolution) covering the whole brain with a BOLD-sensitive EPI sequence (TR = 2 s, TE = 30 ms, FA = 80°; FOV = 220 mm, matrix = 64 x 64; duration of each run = 780 s). Trials were triggered by scanner pulses and presented with the ERTS software (Experimental Run-Time System, Berisoft, Frankfurt, Germany). Stimuli were back-projected from an LCD projector on to a screen viewed through a mirror by the supine subject in the MR scanner.

Image analyses were performed with BrainVoyager QX, version 1.4.9 (Brain Innovation, Maastricht, The Netherlands). Data preprocessing included slice scan time correction with the first scan time within a volume used as a reference for alignment by sinc interpolation, three-dimensional motion correction, spatial smoothing with an 8 mm Gaussian kernel (full width at half-maximum), temporal high pass filtering with a cut-off of 260 s to remove low-frequency non-linear drifts of three or fewer cycles per time course, and linear trend removal. Talairach transformation was performed for the complete set of functional data of each subject, yielding a 4-D data representation (volume time course: 3 x space, 1 x time). A multi-subject statistical analysis was performed by multiple linear regression of the BOLD response time course in each voxel. The general linear model of the experiment was computed for 72 z-normalised volume time courses (18 participants x 4 runs). For each of the four experimental conditions, five task phases were defined representing early encoding (0-4 s) and late encoding (4-8 s), early delay (8-12 s) and late delay (12-16 s) and retrieval (16-18 s). The different task phases were modelled by predictors of 2 s duration in order to avoid contamination by variance in the fMRI signal attributable to neural activity that occurred in the preceding or subsequent task phases (Figure 4.1) (Zarahn et al., 1997). The signal values during these phases were considered the effects of interest. The corresponding predictors were obtained by convolution of an ideal box-car response with a gamma function model of the haemodynamic response (Friston et al., 1998). All error trials were collapsed on a separate predictor.

3D group statistical maps were generated by associating each voxel with the F-value corresponding to the specific set of predictors and calculated on the basis of the least mean squares solution of the general linear model with a random-effects model. The obtained beta weights of each predictor served as input for the second-level whole-brain random-effects analysis including a 2 x 2 factorial design. Thus, the beta values of participants were treated explicitly as realisations of the two within-subjects factors attentional demand (level 1: PO, level 2: NPO) and WM load (level 1: load 1, level 2: load 3), which allowed to directly test for an interaction between the two based on F-statistics. To compare activations between experimental conditions within one task phase, linear contrasts were performed using t-statistics. Multi-subject statistical maps were thresholded at $q < 0.05$, corrected for false discovery rate (Genovese et al., 2002) and visualised on a surface reconstruction of the MNI template brain (courtesy of the Montreal Neurological Institute). FMRI time courses were shown for selected regions of interest (ROI) where the effects of WM load and attentional demand appeared most prominently. ROIs were functionally defined based on the multi-subject statistical maps overlaid on the cortical surface map of the MNI template brain. Starting from the voxel showing peak activation in the multi-subject map, a surface patch of 30 mm^2 (4 mm thickness) was marked. Representative time courses for each experimental condition were obtained by averaging the percent signal changes of the individual voxels within the obtained volume across all participants and repetitions.

Supplementary analysis 1

Hemispheric differences in activation related to WM encoding and visual search were statistically tested by comparing late encoding activation (4-6 s) against the baseline in homologous areas according to our a priori hypotheses. Because the literature on object WM (e.g., D'Esposito et al., 1998; Fletcher and Henson, 2001) and visual attention (e.g., Rosen et al., 1999; Pollmann and von Cramon, 2000; Nobre et al., 2003) suggest hemispheric asymmetries predominantly in the prefrontal cortex, the analysis focussed on the frontal lobes. ROIs were functionally defined based on the WM load and attentional demand contrasts [(load3/PO + load3/NPO) – (load1/PO + load1/NPO); (load3/NPO + load1/NPO) – (load3/PO + load1/PO)]. Peak activation for these contrasts (see Table 4.2) defined the centres of ROIs that comprised a cuboid with a total volume of 512 mm^3 each. In a second step, z-normalised individual beta values were averaged separately for each task condition in these ROIs and their homologous regions in the opposite hemisphere and entered into separate 3-way ANOVAs with the factors search difficulty (PO vs. NPO), WM load (1 vs. 3) and hemisphere (left vs. right). This allowed to test 2-way interactions between search difficulty x hemisphere and WM load x hemisphere.

Supplementary analysis 2

A supplementary analysis was performed in order to quantify differences in the latency of the peak activation between PO and NPO search conditions during WM encoding in selected ROIs. For each experimental condition, the time-of-peak point was extracted from each individual time course within a time window of two to eight volumes after stimulus onset. In those cases where there was more than one peak during the defined time window the first peak was always used to define the time-of-peak point. As this procedure was applied evenly across conditions, it should not have biased the resulting latency differences. Latency differences between PO and NPO search conditions were then compared using t-statistics.

Supplementary analysis 3

To assess the influence of differential search speed in the PO and NPO conditions on the interaction between search difficulty and WM load, the encoding phase (0-8 s) was divided into four phases of 2 s duration each (E1: 0-2 s, E2: 2-4 s, E3: 4-6 s, E4: 6-8 s). Each encoding phase was modelled separately by predictors of 2 s duration that were obtained by shifting an ideal box-car response function (assuming a value of 1 for the volumes of the respective encoding phase and a value of 0 for the remaining time points) by 4 s to account for the haemodynamic delay. Linear contrasts representing 3-way interactions between the factors search difficulty (PO vs. NPO), WM load (3 vs.1) and encoding phase (E1 vs. E2; E1 vs. E3; E1 vs. E4; E2 vs. E3; E2 vs. E4; E3 vs. E4) were calculated separately to test whether search speed had an effect on the interaction between search difficulty and WM load.

4.2 Results

4.2.1 Behavioural performance

An ANOVA tested the effects of search difficulty (PO vs. NPO) and WM load (load 1 vs. load 3) on response accuracy and RT at test. Participants' WM performance at test was equally good under PO and NPO search conditions [WM load 1, 96.1% & 93.8% correct, respectively; WM load 3, 81.5% & 80.6% correct; ANOVA, $F(1, 17) = 2.68$, $p = 0.12$]. RTs to the probe object did not differ between the PO and NPO search conditions either [WM load 1, 807 ms & 769 ms, respectively; WM load 3, 998 ms & 1022 ms; $F(1, 17) = 0.27$, $p = 0.61$] (Figure 4.2). A strong main effect was observed only for WM load. In both search conditions response accuracy declined from WM load 1 to WM load 3 [on average by 13.9 percentage points, $F(1, 17) = 70.37$, $p < 0.001$], and RTs were significantly slower on average by 222 ms [$F(1, 17) = 202.14$, $p < 0.001$]. The interaction between search difficulty and WM load reached significance only for RTs [$F(1, 17) = 6.88$, $p < 0.05$]. The finding that memory performance at test did not differ between PO and NPO search conditions indicates that the presentation time of the search array (8 s) was indeed sufficiently long to ensure that participants were able to complete the encoding process even in the most demanding condition (load3/NPO). Therefore, the task was suitable for probing common and selective activations for visual search and WM encoding with event-related fMRI.

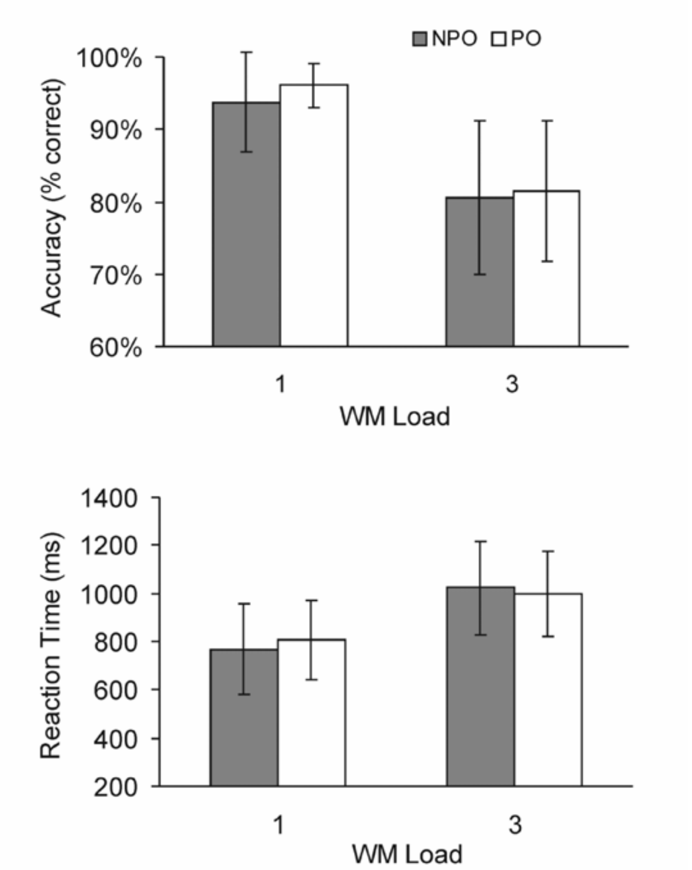


Figure 4.2 Mean response accuracy and reaction times in the four experimental conditions. Bars represent standard errors of the mean. PO: pop-out, NPO: non pop-out.

4.2.2 Brain systems for visual attention and WM encoding

FMRI analysis focused on the late encoding phase (4-8 s) because the behavioural study indicated that encoding times increased by about 4 s when the search changed from PO to NPO (Figure 3.2, lower panel). Therefore, effects of attentional demand were expected during the later part of the stimulus presentation phase. The contrast analyses of fMRI data for the late encoding predictor (4-6 s after stimulus onset) revealed a high degree of overlap in the brain areas that showed higher activation for difficult compared to easy search [(load3/NPO + load1/NPO) – (load3/PO + load1/PO)] and higher activation for WM load 3 compared to WM load 1 [(load3/PO + load3/NPO) – (load1/PO + load1/NPO)]. Overlap in activation was observed in the occipito-temporal cortex, the lateral and medial parietal cortex (inferior parietal sulcus, precuneus), along the precentral sulcus (PrcS), in the frontal midline, the insula, and the thalamus (Table 4.1 and Figure 4.3). The fronto-parietal activation patterns were similar to those reported previously in studies that compared activation induced by attention and WM tasks (LaBar et al., 1999; Pollmann and von Cramon, 2000; Corbetta et al., 2002). In the present study, the common pattern of brain regions involved during both visual search and WM encoding also included the prefrontal cortex (PFC), with overlapping activations restricted to a part of the right middle frontal gyrus (MFG). The left MFG and inferior frontal gyrus (IFG) were selectively responsive to WM load as were regions in the left anterior inferior parietal lobule (IPL) and bilateral inferior temporal cortex (IT). In contrast, areas selectively responsive to high attentional demand were found within the right PFC (MFG and IFG) and occipital cortex.

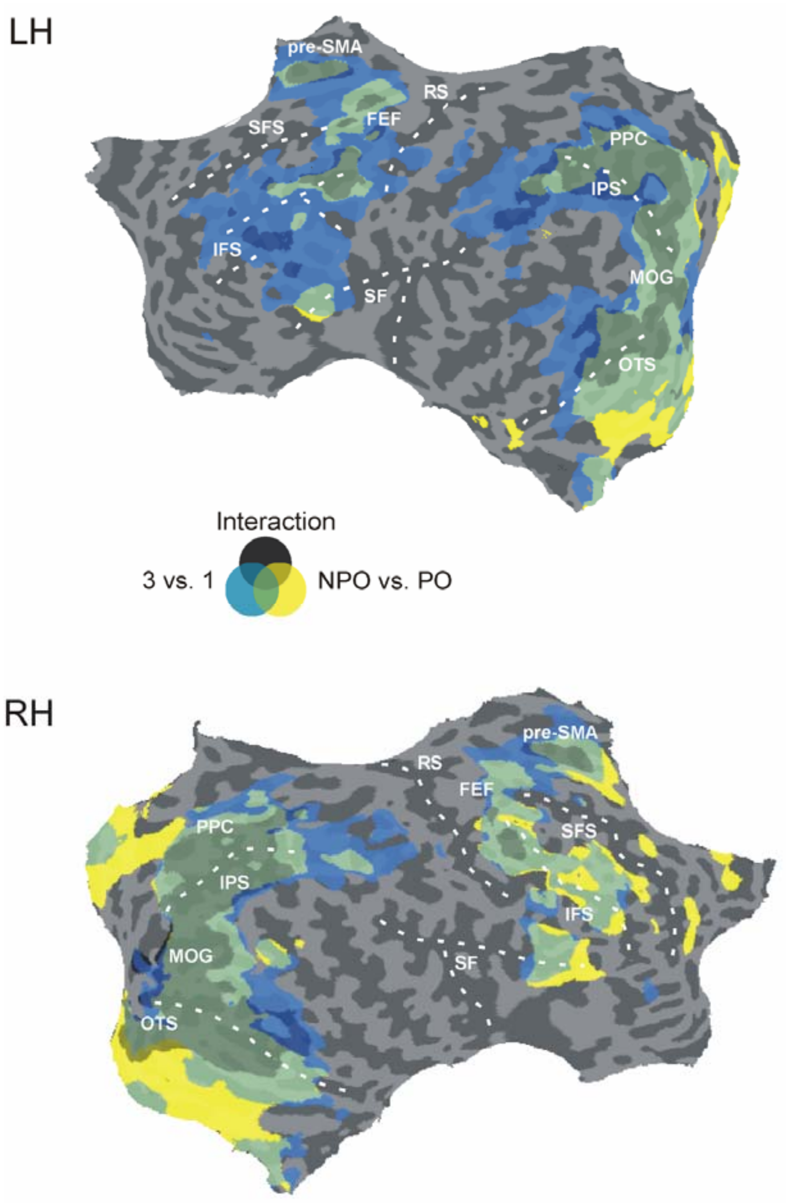


Figure 4.3 Group results for the late encoding predictor (4-6 s). Statistical maps of the contrasts NPO vs. PO (yellow), WM load 3 vs. 1 (blue), and the significant 2-way interaction of search difficulty x WM load (black) are projected on the flattened surface reconstruction of the MNI template brain (courtesy of the Montreal Neurological Institute) (LH: left hemisphere, RH: right hemisphere). Activations are those exceeding a whole-brain false discovery rate threshold of *q*(FDR) < 0.05. FEF: frontal eye field, IFS: inferior frontal sulcus, IPS: inferior parietal sulcus, MOG: middle occipital gyrus, OTS: occipito-temporal sulcus, PPC: posterior parietal cortex, pre-SMA: pre-supplementary motor area, RS: rolandic sulcus, SF: Sylvian fissure, SFS: superior frontal sulcus.

Table 4.1 Brain regions showing significant activation in the contrasts for encoding

					Contrast			
Brain region	BA	x	y	z	WM load	Attentional demand	2-way interaction	3-way interaction
Common activation								
R MFG	9	44	8	38	*	*		
R MFG	46	45	32	27	*	*		
R MFG	46	39	29	18	*	*		
R IFG	9/44	53	9	26	*	*		
L dlPrcS	6	-21	-12	53	*	*	*	
R dlPrcS	6	29	-7	58	*	*		
L FEF	6	-42	-10	48	*	*	*	
R FEF	6	40	-5	55	*	*	*	
L vlPrcS	6	-45	-4	41	*	*	*	*
R vlPrcS	6	47	5	38	*	*	*	
L Pre-SMA	6	-5	5	53	*	*	*	*
R Pre-SMA	6	4	7	55	*	*	*	*
R insula	13	32	15	13	*	*		
L insula	13	-29	23	7	*	*		
L SPL	7	-19	-65	56	*	*	*	
R SPL	7	28	-64	48	*	*	*	
L IPL	40	-38	-41	40	*	*	*	
R IPL	40	34	-54	40	*	*	*	
L precuneus	7	-16	-66	51	*	*	*	
L precuneus	7	-20	-72	35	*	*	*	*
R precuneus	7	25	-66	33	*	*	*	*
L cuneus	18	-24	-89	4	*	*	*	
R cuneus	18	26	-87	0	*	*	*	
L MOG	19	-40	-74	-10	*	*	*	
L FG	37	-40	-58	-7	*	*	*	
R FG	37	40	-59	-11	*	*	*	*
L MOG	19	-24	-80	17	*	*	*	*
R MOG	19	34	-80	17	*	*	*	*
L cuneus	19	-21	-81	32	*	*	*	*
R IOG	19	37	-75	0	*	*	*	*
R lingual Gyrus	18	10	-83	-3	*	*	*	*
L thalamus		-8	-24	1	*	*		
R thalamus		7	-21	3	*	*		
WM-selective								
L MFG	46	-40	32	24	*		*	
L MFG/ IFS	9	-47	21	28	*		*	*
L IFG	45	-52	19	7	*		*	
L IPL	40	-51	-37	40	*		*	
L ITG	19	-48	-61	-2	*		*	*
R ITG	37	53	-45	-14	*		*	
L FG	37	-44	-49	-12	*		*	
R FG	37	51	-46	-15	*		*	
L MTG	37	-50	-47	-8	*			
R MTG	37	49	-49	-1	*			

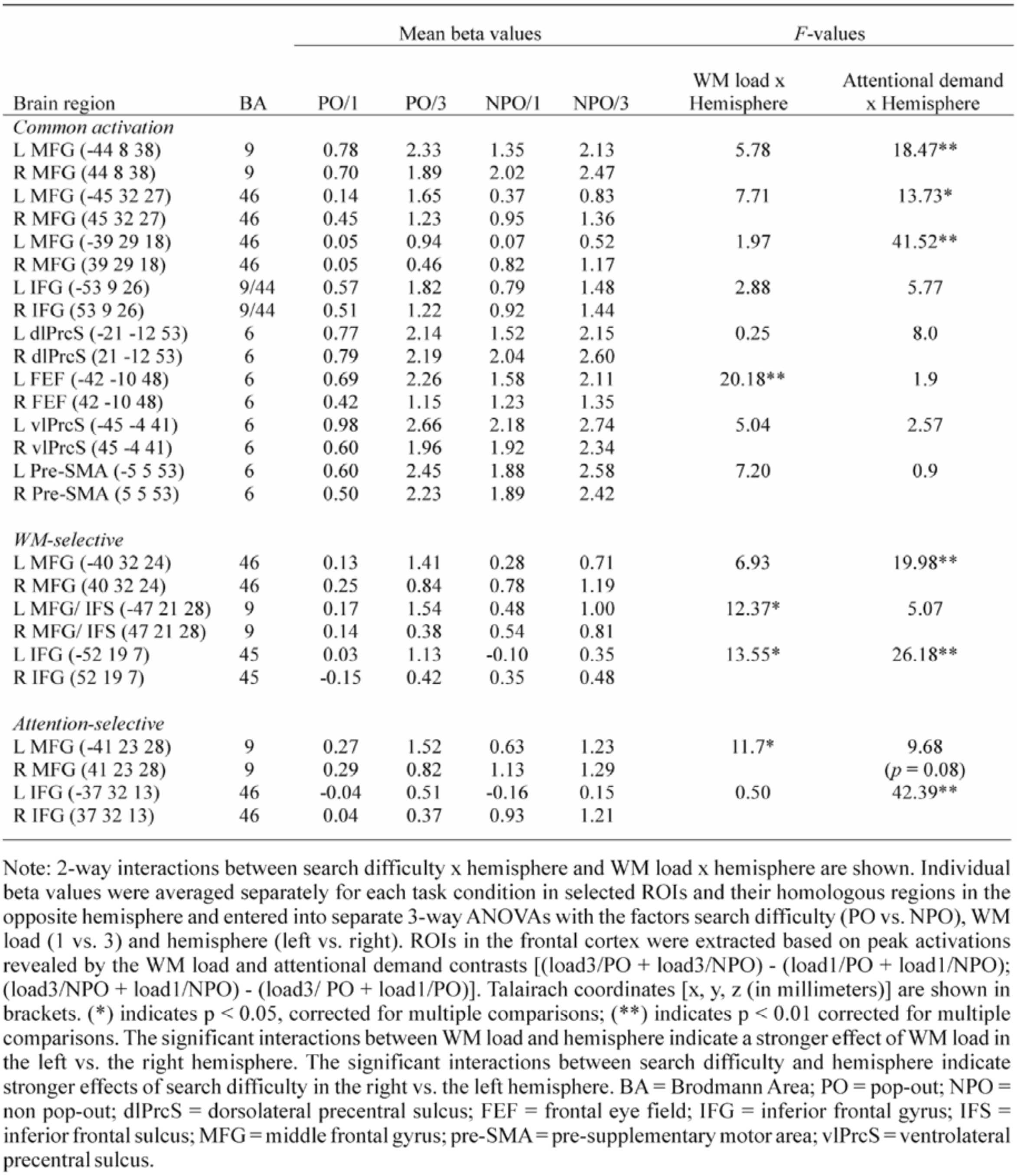
Table 4.1 (continued)

					Contrast			
Brain region	BA	x	y	z	WM load	Attentional demand	2-way interaction	3-way interaction
Attention-selective								
R MFG	9	41	23	28		*		
R IFG	46	37	32	13		*		
L PHG	19	-21	-48	-4		*		
R PHG	19	19	-46	4		*		
L Lingual Gyrus	19	-15	-54	0		*		
R Lingual Gyrus	19	13	-58	3		*		
R Cuneus	31	17	-70	8		*		

Note: Significant contrasts (whole brain random-effects analysis) for the late encoding predictor (4-6 s) are shown. (*) indicates q(FDR) < 0.05. Talairach coordinates [x, y, z (in millimeters)] of the activation maxima are shown. BA = Brodmann Area; WM load = load 3 vs. load 1($t = 2.68$); Attentional demand = NPO vs. PO ($t = 2.71$); 2-way interaction = load 3 vs. load 1 x NPO vs.PO ($F = 11.28$); 3-way interaction = load 3 vs. load 1 x NPO vs. PO x E1 vs. E3 ($t = 4.19$); dlPrcS = dorsolateral precentral sulcus; FEF = frontal eye field; FG = fusiform gyrus; IFS = inferior frontal sulcus; IFG = inferior frontal gyrus; IOG = inferior occipital gyrus; IPL = inferior parietal lobule; ITG = inferior temporal gyrus; MFG = middle frontal gyrus; MOG = middle occipital gyrus; MTG = middle temporal gyrus; PHG = parahippocampal gyrus; pre-SMA= pre-supplementary motor area;SPL = superior parietal lobule; vlPrcS= ventrolateral precentral sulcus; E1 = encoding phase 1 (0-2 s); E3 = encoding phase 3 (4-6 s).

The results indicated that regions within the MFG and IFG were selectively responsive to WM load and attentional demand in the left and right hemispheres, respectively. To test statistically whether this pattern of prefrontal activation reflected a hemispheric dissociation, late encoding activation (4-6 s) was compared against the baseline in homologous areas (Supplementary analysis 1, see Materials and methods). 3-way ANOVAs with the factors search difficulty (PO vs. NPO), WM load (1 vs. 3), and hemisphere (left vs. right) computed separately for each ROI revealed significant interactions between WM load and hemisphere in the MFG, IFG, and the FEF (all F-values > 12.37, all p-values < 0.05, corrected for multiple comparisons), indicating a stronger effect of WM load in the left vs. the right hemisphere. For several regions in the MFG and IFG the effect of search difficulty was stronger in the right vs. the left hemisphere (significant interaction search difficulty x hemisphere, all F-values > 13.73, all p-values < 0.05, corrected for multiple comparisons). None of the other frontal regions showed significant interactions with the factor hemisphere (all F-values < 8.0, all p-values > 0.16, corrected for multiple comparisons) (see Table 4.2). These findings suggest a prefrontal hemispheric specialisation with left PFC selectively responsive to WM load and right PFC selectively responsive to attentional demand.

Table 4.2 Prefrontal hemispheric dissociation in activation related to WM encoding and visual search



		Mean beta values				*F*-values	
Brain region	BA	PO/1	PO/3	NPO/1	NPO/3	WM load x Hemisphere	Attentional demand x Hemisphere
Common activation							
L MFG (-44 8 38)	9	0.78	2.33	1.35	2.13	5.78	18.47**
R MFG (44 8 38)	9	0.70	1.89	2.02	2.47		
L MFG (-45 32 27)	46	0.14	1.65	0.37	0.83	7.71	13.73*
R MFG (45 32 27)	46	0.45	1.23	0.95	1.36		
L MFG (-39 29 18)	46	0.05	0.94	0.07	0.52	1.97	41.52**
R MFG (39 29 18)	46	0.05	0.46	0.82	1.17		
L IFG (-53 9 26)	9/44	0.57	1.82	0.79	1.48	2.88	5.77
R IFG (53 9 26)	9/44	0.51	1.22	0.92	1.44		
L dlPrcS (-21 -12 53)	6	0.77	2.14	1.52	2.15	0.25	8.0
R dlPrcS (21 -12 53)	6	0.79	2.19	2.04	2.60		
L FEF (-42 -10 48)	6	0.69	2.26	1.58	2.11	20.18**	1.9
R FEF (42 -10 48)	6	0.42	1.15	1.23	1.35		
L vlPrcS (-45 -4 41)	6	0.98	2.66	2.18	2.74	5.04	2.57
R vlPrcS (45 -4 41)	6	0.60	1.96	1.92	2.34		
L Pre-SMA (-5 5 53)	6	0.60	2.45	1.88	2.58	7.20	0.9
R Pre-SMA (5 5 53)	6	0.50	2.23	1.89	2.42		
WM-selective							
L MFG (-40 32 24)	46	0.13	1.41	0.28	0.71	6.93	19.98**
R MFG (40 32 24)	46	0.25	0.84	0.78	1.19		
L MFG/ IFS (-47 21 28)	9	0.17	1.54	0.48	1.00	12.37*	5.07
R MFG/ IFS (47 21 28)	9	0.14	0.38	0.54	0.81		
L IFG (-52 19 7)	45	0.03	1.13	-0.10	0.35	13.55*	26.18**
R IFG (52 19 7)	45	-0.15	0.42	0.35	0.48		
Attention-selective							
L MFG (-41 23 28)	9	0.27	1.52	0.63	1.23	11.7*	9.68
R MFG (41 23 28)	9	0.29	0.82	1.13	1.29		($p = 0.08$)
L IFG (-37 32 13)	46	-0.04	0.51	-0.16	0.15	0.50	42.39**
R IFG (37 32 13)	46	0.04	0.37	0.93	1.21		

Note: 2-way interactions between search difficulty x hemisphere and WM load x hemisphere are shown. Individual beta values were averaged separately for each task condition in selected ROIs and their homologous regions in the opposite hemisphere and entered into separate 3-way ANOVAs with the factors search difficulty (PO vs. NPO), WM load (1 vs. 3) and hemisphere (left vs. right). ROIs in the frontal cortex were extracted based on peak activations revealed by the WM load and attentional demand contrasts [(load3/PO + load3/NPO) - (load1/PO + load1/NPO); (load3/NPO + load1/NPO) - (load3/ PO + load1/PO)]. Talairach coordinates [x, y, z (in millimeters)] are shown in brackets. (*) indicates $p < 0.05$, corrected for multiple comparisons; (**) indicates $p < 0.01$ corrected for multiple comparisons. The significant interactions between WM load and hemisphere indicate a stronger effect of WM load in the left vs. the right hemisphere. The significant interactions between search difficulty and hemisphere indicate stronger effects of search difficulty in the right vs. the left hemisphere. BA = Brodmann Area; PO = pop-out; NPO = non pop-out; dlPrcS = dorsolateral precentral sulcus; FEF = frontal eye field; IFG = inferior frontal gyrus; IFS = inferior frontal sulcus; MFG = middle frontal gyrus; pre-SMA = pre-supplementary motor area; vlPrcS = ventrolateral precentral sulcus.

Behavioural evidence (see chapter 3) indicated that the two search conditions differed in the degree of search efficiency as reflected by slower processing in the NPO compared to the PO condition. Specifically, the process of encoding was delayed by about 4 s (Figure 3.2, lower panel). Two further analyses were conducted to examine whether the differences in activation for NPO vs. PO search during the late encoding phase were mainly driven by the varying duration of the search process (fast search with PO vs. slow search with NPO). First, the contrast between NPO search during late encoding (4-6 s after stimulus onset) and PO search during early encoding (0-2 s after stimulus onset) [(NPO/load3/late encoding + NPO/load1/late encoding) – (PO/load3/early encoding + PO/load1/early encoding] was calculated. The results were similar to those obtained for the effect of attentional demand during late encoding, albeit less widespread (Figure 4.4). Overlapping activation for the two contrasts appeared in several lateral frontal, parietal and occipital regions. Therefore, at least for these regions, the effect of attentional demand during late encoding was not likely to be overemphasised because of fast search in the pop-out condition.

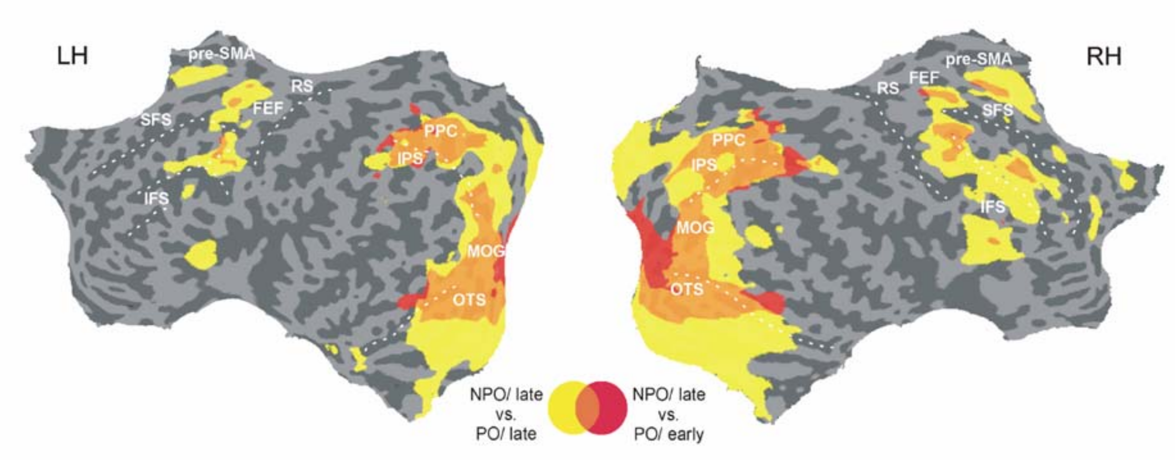


Figure 4.4 Group results obtained by the contrast between NPO search during late encoding (4-6 s after stimulus onset) and PO search during early encoding (0-2 s after stimulus onset). Statistical maps of the contrast [(NPO/load3/late encoding + NPO/load1/late encoding) - (PO/load3/early encoding + PO/load1/early encoding)] are projected in red on the flattened surface reconstruction of the MNI template brain (NPO: non pop-out, PO: pop-out, LH: left hemisphere, RH: right hemisphere). Activations are those exceeding a whole-brain false discovery rate threshold of *q*(FDR) < 0.05. The results were similar to those obtained for the effect of attentional demand during late encoding (in yellow).

In a second analysis, the mean beta estimates for the early and late encoding regressors were extracted from selected ROIs. ROIs were functionally defined based on the contrasts NPO/late encoding vs. PO/late encoding and NPO/late encoding vs. PO/early encoding. In those regions where the interaction effect between WM load and attentional demand was most prominent (left dorsal PrcS and bilateral visual cortex, see next paragraph) a surface patch of 30 mm² (4 mm thickness) was marked, starting from the voxels with the highest activation revealed in the two contrasts maps. Mean beta estimates were extracted from the obtained volumes and entered into ANOVAs with the factors search difficulty (ANOVA 1: NPO/late encoding vs. PO/late encoding; ANOVA 2: NPO/late encoding vs. PO/early encoding) and WM load (load 1 vs. load 3 for both ANOVAs). For each ROI, the effect sizes of the two main effects of attentional demand (all F-values > 10.29, all p-values < 0.01) were compared. If the attention effects during the late encoding phase were overemphasised in these ROIs it was expected to find stronger effect sizes for NPO/late encoding vs. PO/late encoding than for NPO/late encoding vs. PO/early encoding. The results did not support this hypothesis. For the left dorsal PrcS the effect sizes of the attention effects were comparable ($\eta^2 = 0.62$ for NPO/late encoding vs. PO/late encoding; $\eta^2 = 0.54$ for NPO/late encoding vs. PO/early encoding). For the left and the right occipital cortex (OC) the effect size was even stronger for NPO/late encoding vs. PO/early encoding ($\eta^2 = 0.81$ for the left OC; $\eta^2 = 0.79$ for the right OC) than for NPO/late encoding vs. PO/late encoding ($\eta^2 = 0.58$ for the left OC; $\eta^2 = 0.38$ for the right OC). These findings provided also support that in those regions where the interaction effect (see next paragraph) was most prominent it was not attributable to the faster search in the PO condition.

4.2.3 Interference between attention and WM encoding

Brain areas reflecting functional interference between attention-demanding visual search and WM encoding were identified by the interaction contrast [(load3/PO - load1/PO) - [load3/NPO - load1/NPO)]. Significant activation was found only in a subset of the regions with overlapping activations for the attention and WM load contrasts. These regions included the occipito-temporal and posterior parietal cortex as well as the medial frontal cortex and the PrcS of both hemispheres (Table 4.1 and Figure 4.3, dark green colour). Time course analyses of these regions showed a smaller increase in BOLD signal with increasing WM load for NPO compared to PO search (Figure 4.5A, purple circles). This type of interaction was most pronounced in early and higher visual areas (middle occipital gyrus, cuneus) and in the left dorsal PrcS. Here, the BOLD response was always the lowest in the PO condition when participants needed to memorise only one object, but increased to the same degree in the remaining three conditions (Figure. 4.5A, red circles). Thus, in these brain areas, the BOLD response could not exceed the plateau of activation that was reached already with load3/PO or with load1/NPO in order to respond to joint demands on WM and attention. In contrast, in adjacent brain regions that showed an overlap in activation but no interaction, the BOLD signal further increased in the most difficult condition (load3/NPO) (Figure 4.5B, black circles).

Areas preferentially sensitive to WM load (but not to attentional demand) also emerged in the analysis of 2-way interactions. These included the left MFG, IFG and anterior IPL and the IT cortex, bilaterally (Table 4.1). Here, the time course of BOLD activation peaked later under conditions of NPO vs. PO search (Figure 4.5A, orange circle) with significant differences in the latency of the peak amplitudes in the left MFG ($t = 2.8$, $p < 0.01$ for load 1; $t = 2.7$, $p < 0.01$ for load 3), the left IT cortex ($t = 4.02$, $p < 0.001$ for load 1; $t = 3.1$, $p < 0.01$ for load 3) and the left anterior IPL ($t = 4.5$, $p < 0.01$ for load 1; $t = 1.8$, $p = 0.08$ for load 3) (Supplementary analysis 2, see Materials and methods). Consistently with the delay in encoding times revealed in the behavioural study (Figure 3.2, lower panel), the time shift in the maximum amplitude of BOLD activation was about 4 s and appeared without any compromise on the size of the WM load effect in the NPO condition.

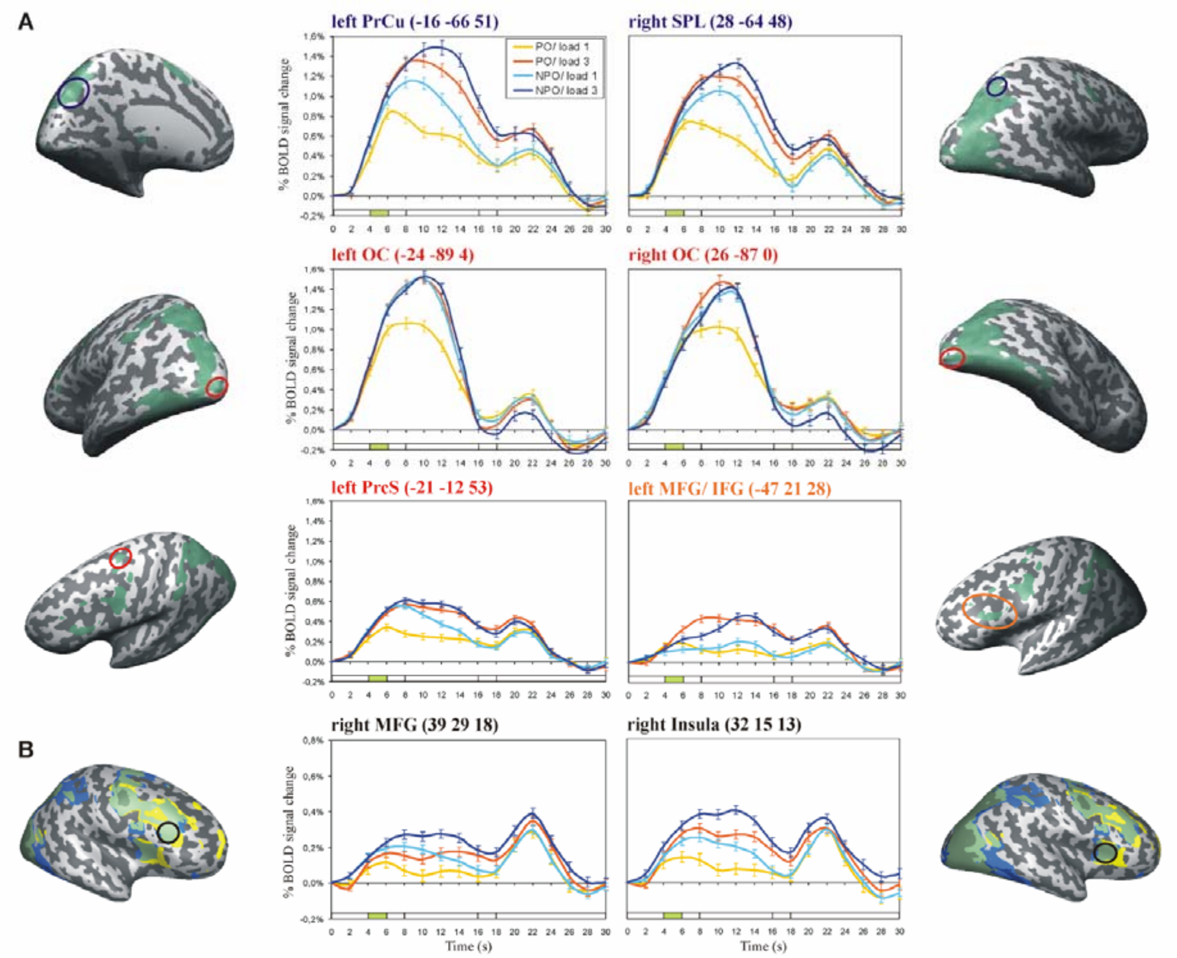


Figure 4.5 Averaged time courses of the BOLD response in the four conditions. A. Statistical group maps of the interaction contrast of search difficulty x WM load during late encoding are shown. The maps are projected on inflated surface reconstructions of the MNI template brain (dark green). During encoding time courses indicated a smaller increase in BOLD signal with increasing WM load for NPO vs. PO in parietal regions (purple circles, PrCu: precuneus, SPL: superior parietal lobule). In the occipital cortex (OC) and the PrcS the BOLD response did not exceed a plateau of activation that was reached already with load3/PO and load1/NPO (red circles). A delayed WM load effect for NPO vs. PO was revealed in the left MFG and IFG (orange circle). B. Statistical maps of the contrasts NPO vs. PO (yellow), WM load 3 vs. 1 (blue), and the significant 2-way interaction of search difficulty x WM load (black) are shown. Regions in the right MFG and insula (black circles) showed an additive increase in activation with increased WM load and search difficulty. Bars represent standard errors of the mean.

The smaller effect of WM load under high attentional demand in posterior areas and the PrcS might also be driven by temporal differences between the two search conditions (slow vs. fast search). In subsequent whole-brain analyses the influence of search speed on the interaction between search difficulty and WM load was assessed by calculating the interactions between the factors search difficulty, WM load and encoding phase (Supplementary analysis 3, see Materials and methods). Significant activation reflecting a 3-way interaction was found only between the factors search difficulty (NPO vs. PO), WM load (load 3 vs. load 1), and encoding phase (E3 vs. E1) bilaterally in a distributed network of occipital, temporal, and parietal areas [q(FDR) < 0.05]. Frontal activations included the medial frontal cortex, the left ventral PrcS, and the left MFG/ IFG (Table 4.1 and Figure 4.6), the same regions where the latency of the peak amplitude between PO and NPO search had appeared most strongly (Fig. 4.5A). These results, again, indicated a time shift in activation produced by the difficult NPO search. However, the observed regions differed from the areas that showed strong plateau effects during late encoding (left dorsal PrcS and bilateral visual cortex). The regions in the lateral parietal cortex that were associated with the 2-way interaction contrast did not emerge in this 3-way interaction contrast either. Thus, the decreased WM load effect under high attentional demand observed in these areas could not be explained by time shifts in peak activation between the two search conditions.

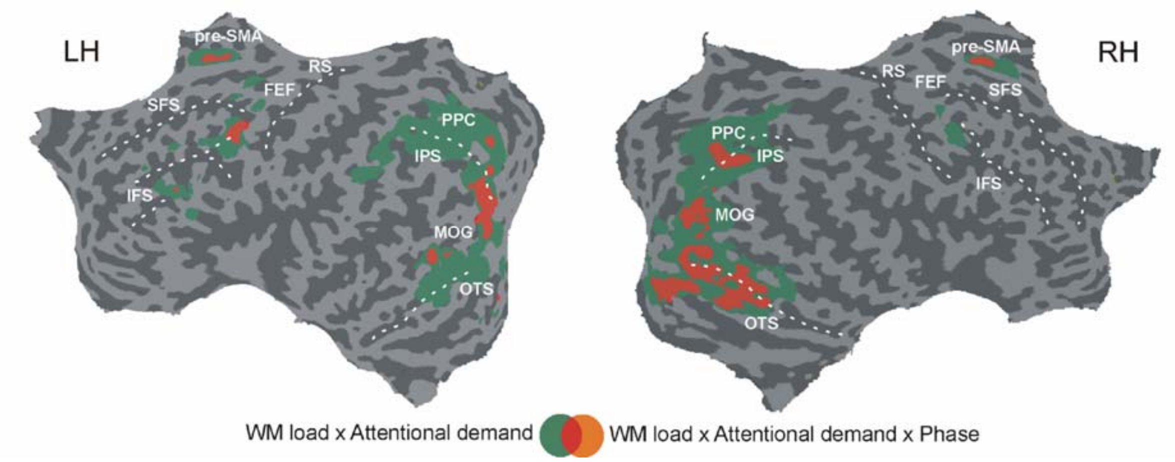


Figure 4.6 Influence of search speed on the interaction of search difficulty x WM load. Results of the 3-way interaction analysis between search difficulty (NPO vs. PO), WM load (3 vs. 1) and encoding phase (E3 vs. E1) are shown superimposed in orange on the results of the 2-way interaction of search difficulty x WM load (dark green). Activations are those that exceeded a whole-brain false discovery rate threshold of q(FDR) < 0.05. Encoding phase E1: 0-2 s, encoding phase E3: 4-6 s. NPO: non pop-out, PO: pop-out.

4.2.4 Load effects during WM maintenance

One goal of this study was to investigate neurophysiological interactions between attentional processes involved in visual search and the encoding of information into visual WM. It was reasoned that if participants successfully performed the WM task despite the concurrent demands on attentional resources, the observed effect of interference between search difficulty and WM load should be restricted to the encoding phase. Interference between the two processes should not be observed during the subsequent delay phase. Consistent with this hypothesis, no significant activation was found for the interaction contrast between search difficulty and WM load for the late delay predictor (12-14 s after stimulus onset). Neither did the NPO vs. PO contrast yield significant activation. Thus, the process of active maintenance of objects in WM was not limited by attentional processing required by difficult visual search. The increase in the number of objects maintained in WM (load 1 vs. load 3) was associated with significant activation mainly around the IPS, extending into both superior and inferior parietal lobules, the lateral prefrontal, medial frontal and premotor cortex, the temporal cortex and the insula. These activation foci were almost identical to those observed during the encoding phase, which revealed additional activation in early and higher visual areas (Figure 4.7, Table 4.3).

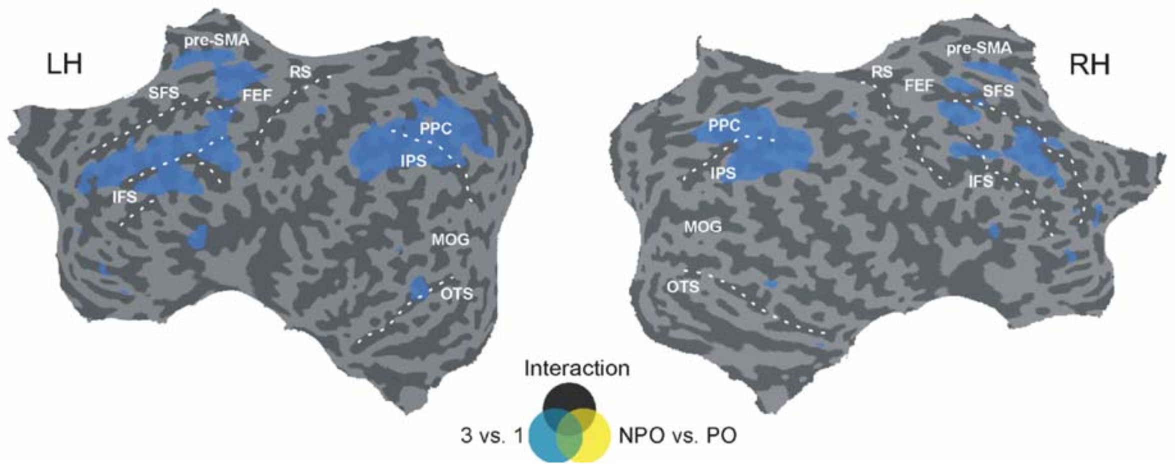


Figure 4.7 Group results for the late delay predictor (12-14 s). Significant activations were found only for the WM load contrast (load 3 vs. load 1) (blue). Activations are those exceeding a whole-brain false discovery rate threshold of *q*(FDR) < 0.05.

Table 4.3 Brain regions showing significant activation in the contrasts for delay

					Contrast		
Brain region	BA	x	y	z	WM load	Attentional demand	2-way interaction
WM-selective							
L SFG	9	-28	45	29	*		
R SFG	9	27	32	33	*		
L MFG	9	-38	17	29	*		
R MFG	9	39	31	34	*		
L IFG	44	-45	9	31	*		
L dlPrcS	6	-22	-7	57	*		
R dlPrcS	6	28	-6	56	*		
L FEF	6	-39	-8	44	*		
R FEF	6	31	-5	52	*		
L vlPrcS	6	-39	-1	38	*		
R vlPrcS	6	44	3	34	*		
L Pre-SMA	6	-5	1	60	*		
R Pre-SMA	6	3	3	59	*		
L Insula	13	-31	22	9	*		
R Insula	13	32	22	11	*		
L SPL	7	-35	-57	48	*		
R SPL	7	30	-58	42	*		
L IPL	40	-33	-48	39	*		
R IPL	40	37	-55	43	*		
L Precuneus	7	-12	-69	51	*		
R Precuneus	7	12	-72	50	*		
L ITG/ FG	37	-38	-51	-6	*		
R ITG	37	56	-51	-12	*		

Note: Significant contrasts (whole brain random-effects analysis) for the late delay predictor (12-14 s) are shown. (*) indicates q(FDR) < 0.05. Talairach coordinates [x, y, z (in millimeters)] of the activation maxima are shown. BA =Brodmann Area; WM load = load 3 vs. load 1 (t = 2.85); Attentional demand = NPO vs. PO; 2-way interaction = load 3 vs. load 1 x NPO vs.PO; BA = Brodmann Area; PO = pop-out; NPO = non pop-out; dlPrcS = dorsolateral precentral sulcus; FEF = frontal eye field; FG = fusiform gyrus; IFG = inferior frontal gyrus; IPL = inferior parietal lobule; ITG = inferior temporal gyrus; MFG = middle frontal gyrus; pre-SMA = pre-supplementary motor area; SFG = superior frontal gyrus; SPL = superior parietal lobule; vlPrcS = ventrolateral precentral sulcus.

4.3 Discussion

In this experiment, visual search and delayed discrimination of complex objects were combined within one single task and the demands on selective attention and WM encoding were independently modulated. The goal was to identify the brain regions that were selectively responsive to either WM or attentional demand and those involved in both processes. It was hypothesised that if visual WM and selective attention were subserved in part by common areas with limited neural processing capacity, activation in these areas under conditions of joint demand on both processes should reach a plateau or at least be less than additive, as reflected in a statistical interaction between attention and WM. Conversely, an additive increase in BOLD activation under simultaneous WM and attentional demands was expected in regions whose processing capacity was not exceeded. The BOLD signal in these overlap regions should increase to the same degree with WM load under low and high attentional demand. It was important to observe this pattern in at least some brain areas in order to rule out the possibility that the capacity constrained pattern observed in other areas was an effect of haemodynamic saturation or time spent on task components.

4.3.1 Common activation for visual attention and encoding into visual WM

Overlapping activation for attention-demanding visual search and encoding into visual WM was observed in distributed posterior and frontal regions. Consistent with the hypotheses a subset of these regions, in the right prefrontal cortex and bilateral insula, showed an additive increase in BOLD activation associated with increased WM load and attentional demand. These results are in agreement with the view that the processes underlying attention-demanding visual search and the encoding into visual WM require access to common neural and cognitive resources. The additive increase in BOLD activation suggests that the demands on these frontal regions were well within their processing limits even in the condition where high WM load was combined with difficult search. Conversely, the analysis revealed an interaction effect between the two task manipulations for visual, parietal, and premotor cortex. Activation increased from WM load 1 to WM load 3 but this increase was significantly smaller in the difficult (NPO) compared to the easy (PO) search condition. In contrast, activation associated with increased WM load in the left PFC was delayed rather than reduced under high attentional demand. These results indicate that competition for processing resources that are shared by the WM and attention systems can lead to a severe limitation of neural processing capabilities.

The brain areas mediating these common processing limitations of visual WM and attention included regions that are classically considered to support goal-directed visuospatial attention (Kanwisher and Wojciulik, 2000; Corbetta et al., 2002; Pessoa et al., 2003) and have been implicated in the capacity limitation of visual

WM (Linden et al., 2003; Todd and Marois, 2004; Marois and Ivanoff, 2005; Xu and Chun, 2006). Indeed, a survey of the neural substrates that support top-down mechanisms for visual WM showed a striking degree of overlap with those of selective attention (Pessoa and Ungerleider, 2004). The design characteristics of the combined task allowed assigning functional consequences to the overlap in activation by testing for interactions between the two task components which were found in a subset of the brain regions that supported both attentional selection and WM encoding. Thus, the demonstration of interference between the processes involved in attention-demanding visual search and WM encoding strongly suggests that the two cognitive domains tap into common neural resources.

4.3.2 Evidence for a neural bottleneck of visual attention and encoding into visual WM?

Capacity limits of information processing traditionally have been interpreted in terms of bottlenecks that occur if the same two cognitive operations act upon a single capacity-limited channel (Broadbent, 1958). As a result, one or both operations will be delayed or otherwise impaired (Pashler, 1994; Jolicœur and Dell'Acqua, 1999; Sigman and Dehaene, 2005). It has been shown that processing bottlenecks can operate at different stages in the flow of information from perception to memory and action (Pashler, 1998; Marois and Ivanoff, 2005). With regard to the present findings it thus can be proposed that the distributed regions in the posterior, but not prefrontal cortex form a neural bottleneck for joint demand on attention and WM resources during the stage of WM encoding.

On the basis of the present data it cannot be decided whether the bottleneck reflects capacity limitations at a particular set of regions or constraints of the capacity for cooperation among multiple regions. Nevertheless, the results show that event-related fMRI can detect interactions in activity patterns in response to increased attentional and WM demands within distributed cortical regions. The findings indicate common capacity limitations for visual WM and attention in the occipito-temporal and posterior parietal cortex, the PrcS, and the pre-SMA in both hemispheres. This limitation was manifested in a reduced WM load effect under conditions of NPO versus PO search and was pronounced most strongly in early and higher visual areas and in the left dorsal PrcS.

It might be argued that the plateau of activation that was already reached with load3/PO and load1/NPO was a result of a haemodynamic saturation of the neurovascular system. BOLD activation in visual areas showed an increase of up to 1.5 % signal change and in the left dorsal PrcS a plateau of activation was reached at 0.6% signal change. This activation is unlikely to have reached the physiological plateau because checkerboard stimulation with similar scanning parameters can lead to BOLD signal changes of up to 4% in the occipital cortex (Uludag et al., 2004) which is about three-fold larger activation than the presently observed saturation point. Moreover, several regions associated with an overlap in activation but no interaction showed a further increase in BOLD activity from load1/NPO to load3/NPO and from load3/PO to load3/NPO. Such an additive increase appeared in regions adjacent to those showing a strong plateau effect, for instance in the right MFG and bilateral insula, which suggests that the latter effect also resulted from differential processing induced by the task manipulations rather than from haemodynamic saturation.

It furthermore does not seem plausible that the plateau effect observed in the visual cortex is owed to limitations on perceptual rather than memory processes. It has been proposed that in inefficient visual search tasks when targets and distractors are highly similar attention is shifted serially from one item or one group of items to the next (Treisman and Gormican, 1988). At most one group might comprise about 4 items (Pylyshyn and Storm, 1988; Cavanagh and Alvarez, 2005). As participants were not informed about the number of targets presented in the upcoming array they had to serially scan the entire array of nine items in order to find the single target in the NPO condition. In contrast, in the PO condition they immediately focused the target items. Thus in load3/PO only three (or one group of three items) as opposed to nine items (or three groups of three items) in load1/NPO had to be processed. If memory processing had not played a role, a further increase in activation for load1/NPO compared to load3/PO due to a higher perceptual load, would have been expected, which, however, was not observed. Thus, activation in the visual cortex was not solely a result of limitations on perceptual processing but rather reflected both perceptual and WM-related processing.

In the present task, the two search conditions differed in the degree of search efficiency as indicated in the behavioural study by slower processing times (about 4 s) in the NPO compared to the PO condition. However, in light of the behavioural performance in the fMRI experiment and the absence of search difficulty effects on the delay activity it is unlikely that insufficient time available for WM encoding in the most demanding condition produced the smaller effect of WM load in the NPO vs. the PO condition. Response accuracy and RTs at test were equally high in the two search conditions and delay activity increased to the same degree from WM load 1 to load 3, irrespective of search difficulty. Moreover, the additive increase in activation with high WM load and difficult search in several regions rules out that the observed interaction effect was owed to incomplete encoding or prolonged search in the more demanding conditions (load3/NPO, load3/PO and load1/NPO) compared to the less demanding condition (load1/PO). Taken together, these results indicate that even under difficult visual search participants efficiently engaged into the process of encoding into WM, which is a prerequisite for successful WM maintenance.

The influence of temporal differences across search conditions was further addressed in two subsequent analyses. Taking the faster search process in the PO condition into account by contrasting NPO/late encoding (4-6 s after stimulus onset) versus PO/early encoding (0-2 s after stimulus onset) the effect of attentional demand that was observed during the late encoding phase could be replicated in those regions that showed strong plateau effects associated with processing limitations during encoding (left dorsal PrcS and bilateral visual cortex) (see Figure 4.4). Moreover, 3-way interaction analyses between the factors attentional demand, WM load and encoding phase did not yield significant activation in those regions as well (Figure 4.6). Of course, caution is warranted in interpreting non-effects because of potentially insufficient statistical power. However, the BOLD response functions which showed little differences in latency or slope across conditions in these regions (Figure 4.5A) suggest that the plateau effects were not a result of fast versus slow visual search.

4.3.3 Selective activations for visual attention and encoding into visual WM

As participants applied attentional and WM processes to the same stimulus displays the present experiment allowed to identify the brain areas that were selectively responsive to either WM encoding or attentional demand. Areas specifically sensitive to WM load appeared in the left lateral PFC, the left anterior inferior parietal lobule, and bilaterally in the inferior temporal cortex. Interestingly, the prefrontal areas showed a time shift in activation associated with the increase in WM load between the PO and the NPO condition (Figure 4.5A). The delay of about 4 s reflected accurately the delay in encoding times estimated in the behavioural study (Experiment 1). As attention-demanding visual search and WM encoding shared a large portion of their neural resources in posterior regions the delayed WM load-related activation in the left PFC might be a consequence of this neural bottleneck. In the light of equal memory performance at test across search conditions, it can be proposed that this delay in activation reflects a mechanism that allowed participants to compensate for the common demands on limited neural resources shared by attention and WM processes in the posterior cortex. The interplay between the PFC and posterior regions was not in the direct focus of the present study. Nevertheless, the present data indicate that successful encoding into visual WM requires joint processing across encoding-selective areas and areas that are also called upon by demands on selective attention. The availability of neural resources mediating selective attention, thus, seems to be a critical factor for constraining the process of encoding information into visual WM.

Interestingly, the PFC showed a hemispheric asymmetry with left MFG and IFG selectively responsive to WM load and right MFG and IFG selectively responsive to attentional demand which might point to a functional dissociation of the PFC. In line with this finding, prefrontal hemispheric specialisation has been reported in previous imaging studies showing right-dominant activation during conditions of inefficient visual search (Pollmann and von Cramon, 2000; Nobre et al., 2003) and visuo-spatial orienting (Rosen et al., 1999). WM for non-spatial material such as objects, colours, and faces has been associated particularly with the left PFC in contrast to spatial material which is represented predominantly in the right hemisphere (D'Esposito et al., 1998; Fletcher and Henson, 2001; Munk et al., 2002; Manoach et al., 2004; Mohr et al., 2006). Therefore, the left-hemispheric dominance of WM load-related activation in the present task might reflect content-specific encoding processes. The question of content-specific encoding processes was further addressed in the second fMRI experiment.

4.4 Excursus: Task-induced deactivations are process-specific

Typically, fMRI studies focus on reporting only increases in BOLD activity caused by the neural response to an externally controlled stimulus or task. These task-induced changes (“activations”) are detected when comparisons are made between a task state, designed to place specific demands on the brain, and a baseline or control state, with a set of demands that are uniquely different from those of the task state. However, researchers have also frequently encountered decreases in BOLD activity during the performance of cognitive tasks (“deactivations”). Task-induced deactivations indicate higher levels of blood flow during the baseline state than during the task of interest.

The basis of these decreases in BOLD activation is incompletely understood at present. Deactivation has been discussed in terms of a direct haemodynamic compensatory response in the vascular system (Devor et al., 2005). According to the so-called “vascular-steal hypothesis” decreases may result from a redistribution of cerebral blood flow to areas that are active from adjacent areas. However, increasing evidence suggests that there is a significant neuronal contribution to the negative BOLD signal (Shmuel et al., 2006). Also, fMRI studies using cross-modal stimuli have shown deactivation of the auditory cortices during visual stimulation and in the visual cortices during auditory stimulation (Laurienti et al., 2002). In this case, the specific areas that decrease in activity are dependent on the characteristics of the task (“task-dependent decreases”). However, what is most remarkable about deactivations is that they have been found consistently in several brain regions across a wide variety of cognitive tasks (e.g., attention, memory, language processing, and motor tasks) and sensory modalities (e.g., visual and auditory) (Shulman et al., 1997; Mazoyer et al., 2001). Common regions typically include the ventromedial prefrontal cortex, the anterior and posterior cingulate cortex, medial and lateral parietal regions (precuneus, supramarginal and angular gyri), and the medial temporal lobe. Thus, task-induced deactivations can occur independent of the characteristics of the task. This finding has led to the assumption of an organised mode of brain function that is present as a baseline ("default mode") reflecting a set of operations that are spontaneously employed when people are mentally unconstrained, and that is attenuated during various goal-directed behaviours. According to this account, default activity is an inverse function of task demand, where higher demands reduce activity in the default network because the mental resources used for various internal processes have been reallocated to performing the task (Gusnard and Raichle, 2001; McKiernan et al., 2003). The default mode theory has gained much interest in the study of human brain function. Many candidates for ongoing processes have been proposed among them mind-wandering, self-awareness, monitoring the external environment, the body image, the mental or emotional state (Gusnard and Raichle, 2001; McKiernan et al., 2003; Mason, 2007).

The investigation of task-induced deactivation in response to attentional demand and WM load aimed at further testing the hypothesis of a default mode of brain function. According to the default mode theory I predicted an increase in BOLD deactivation as a function of task demand in a highly similar set of regions for both the attention and WM components of the task. More specifically, these overlap areas should consistently show an additive increase in BOLD deactivation as a consequence of an increase in the demands on WM and visual search difficulty. In contrast, if the results on task-induced BOLD deactivations revealed exclusive main effects for either attentional demand or WM load at different cortical sites this finding would be inconsistent with the default mode theory. Rather, this finding would indicate that task-induced deactivations are associated with neural processing that is specific to each task component.

4.4.1 Results

During the encoding phase of the present task decreases in BOLD activity in comparison to the baseline activity were observed in several areas that belong to the default network described above. The contrast between the late encoding predictors (PO1/late encoding + PO3/late encoding + NPO1/late encoding +

NPO3/ late encoding) and the baseline revealed deactivations bilateral in the posterior cingulate cortex, in posterior lateral cortices (angular gyrus, supramarginal gyrus), and the anterolateral temporal cortex including the posterior end of the Sylvian fissure and the superior and middle temporal gyri. Regions in the frontal cortex included the medial frontal gyrus, the ventral anterior cingulate cortex, and parts of the lateral superior frontal and inferior frontal gyri (Figure 4.8).

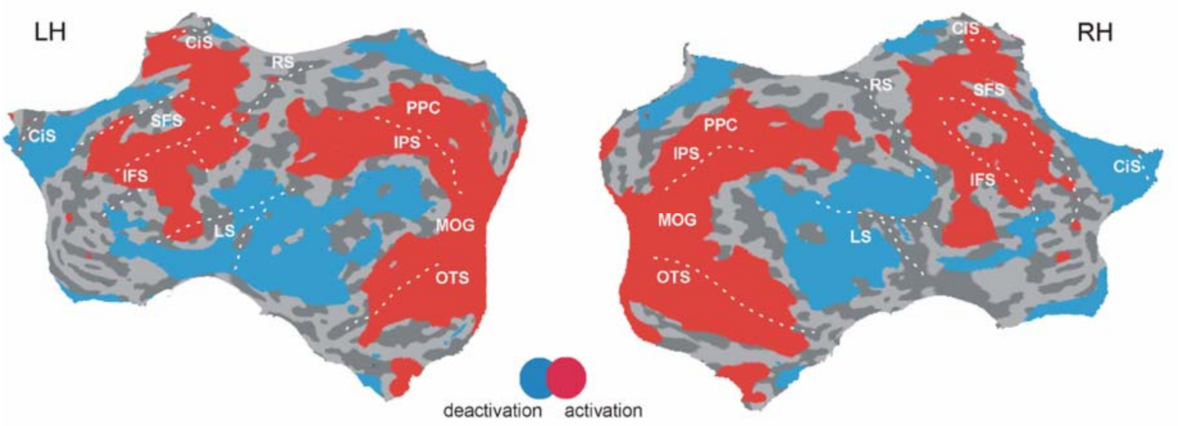


Figure 4.8 Activation (red) and deactivation (blue) compared to the baseline revealed by the late encoding predictors (4-6 s). Activations and deactivations are those exceeding a whole-brain false discovery rate threshold of q(FDR) < 0.05. (LH: left hemisphere, RH: right hemisphere). CiS: cingulate sulcus, IFS: inferior frontal sulcus, IPS: inferior parietal sulcus, LS: lateral sulcus, MOG: middle occipital gyrus, OTS: occipito-temporal sulcus, PPC: posterior parietal cortex, RS: rolandic sulcus, SFS: superior frontal sulcus.

Further analyses were conducted to test whether these deactivations were modulated by the demands on attention and/or WM load. For this purpose, separate masks were defined based on the group statistical maps that reflected a significant main effect of attention (NPO vs. PO), WM load (load 3 vs. load 1) and the interaction between the two factors (including activations and deactivations) and the GLM was calculated separately for each mask. The design matrix was the same as for the whole brain analysis, though the GLM was restricted to the voxels of the functionally defined masks. Significant decreases from baseline were then tested by contrasting BOLD activity during the late encoding phase (PO1/late encoding + PO3/late encoding + NPO1/late encoding + NPO3/ late encoding) against baseline activity. As depicted in Figure 4.9, distributed regions showed stronger deactivations for both NPO compared to PO conditions and for WM load 3 compared to load 1. These regions included the left posterior medial cortex (posterior cingulate cortex, precuneus), parts of the superior and middle temporal gyri, bilateral insula, the right PrcS, and the left medial frontal gyrus. Overall, deactivation in response to attentional demand appeared to be stronger in the left, and to WM load in the right hemisphere. Areas selectively suppressed by high attentional demand were found in the left lateral prefrontal cortex (SFG, MFG, IFG) and the medial frontal cortex (medial frontal and anterior cingulate gyrus). Posterior regions included the left middle and bilateral superior temporal gyrus, the left angular gyrus and posterior parts of the right insula. In contrast, regions in the right hemisphere were selectively suppressed by high WM load, e.g, the precuneus, the supramarginal gyrus, and temporal regions. Frontal regions included the right IFG, bilateral anterior medial frontal gyrus and anterior cingulate cortex. Deactivation reflecting a significant interaction between search difficulty and WM load was found in the left posterior cingulate cortex and precuneus, the left posterior insula, and the left superior temporal gyrus. In the right hemisphere interaction effects were observed in the anterior cingulate cortex, the medial and the superior frontal gyrus. Time course analyses of these regions showed a smaller decrease in BOLD signal with increasing WM load for NPO compared to PO search. BOLD deactivation was always the lowest in the PO condition when participants needed to memorise only one object, but increased to the same degree in the remaining three conditions. In contrast, in brain regions that showed an overlap in deactivation but no interaction, the BOLD signal additively decreased with increased attentional demand and WM load (Figure 4.9).

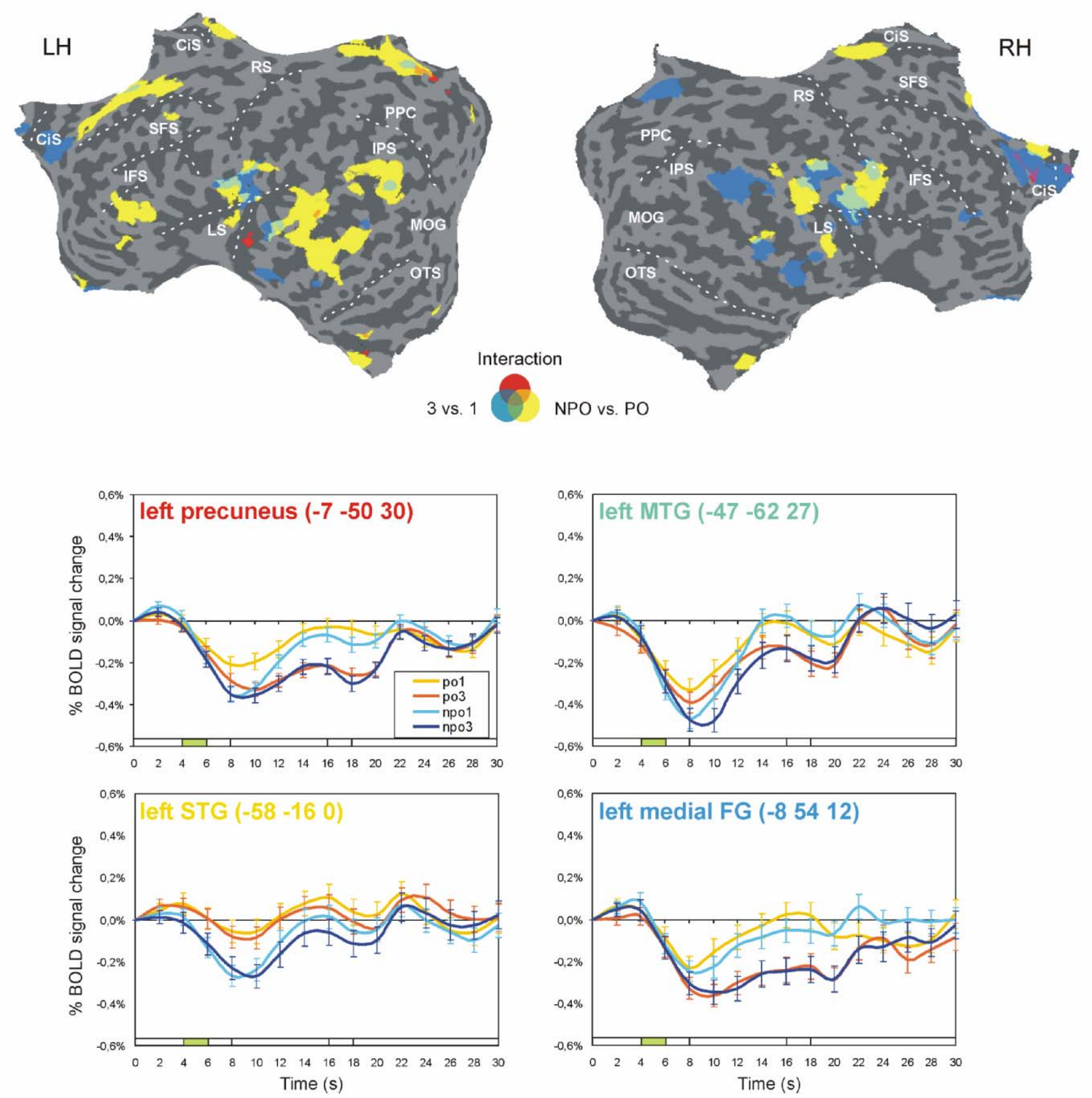


Figure 4.9 Significant deactivations in the four experimental conditions during encoding (4-6 s). Statistical maps of the contrasts NPO vs. PO (yellow), WM load 3 vs. 1 (blue), and the significant 2-way interaction of search difficulty x WM load (red) are shown. Deactivations are those exceeding a whole-brain false discovery rate threshold of *q*(FDR) < 0.05. (LH: left hemisphere, RH: right hemisphere). CiS: cingulate sulcus, FG: frontal gyrus, IFS: inferior frontal sulcus, IPS: inferior parietal sulcus, LS: lateral sulcus, MOG: middle occipital gyrus, MTG: middle temporal gyrus, OTS: occipito-temporal sulcus, PPC: posterior parietal cortex, RS: rolandic sulcus, SFS: superior frontal sulcus, STG: superior temporal gyrus.

Table 4.3 Brain regions showing significant deactivation in the contrasts for encoding

Brain region	BA	x	y	z	Brain region	BA	x	y	z
Common deactivation					*Interaction*				
L SFG	9	-9	30	51	R ant. cingulate	32	5	38	13
R PrcS	4	52	-4	16	R medial FG	10	5	57	21
L insula	13	-36	-15	19	R SFG	9	5	53	32
R insula	13	45	-13	17	L insula	13	-43	-8	-5
L STG	22	-52	-15	7	L STG	22	-56	-33	8
L STG	39	-47	-56	31	L PHG	35	-23	-17	-11
L MTG	39	-49	-61	27	L precuneus	31	-6	-50	29
R IPL	40	55	-30	23	L post. cingulate	31	-12	-55	22
L precuneus	31	-11	-43	35					
L cingulate	31	-9	-43	36					
Attention-selective					*WM-selective*				
L ant. cingulate	24	-8	32	-3	L ant. cingulate	24	-7	34	9
L medial FG	6	-2	-18	51	R ant. cingulate	32	6	41	7
R medial FG	9	5	37	31	L medial FG	10	-6	48	14
L SFG	9	-8	50	36	R medial FG	9	4	47	29
R SFG	8	5	36	43	R IFG	46	47	39	3
L MFG	6	-32	11	51	L STG	38	-43	12	-18
L IFG	47	-37	29	-6	R MTG	22	66	-31	4
R insula	13	50	-33	18	R supramarginal gyrus	40	51	-56	32
R STG	22	51	-10	9	R precuneus	31	6	-47	29
L MTG	21	-58	-18	-7					
L angular gyrus	39	-41	-65	36					
L paracentral lobule	31	-5	-30	47					
L paracentral lobule	6	8	-24	46					
L cingulate	31	-4	-35	41					

Note: Significant contrasts (whole brain random-effects analysis; *q*(FDR) < 0.05) for the late encoding predictor (4-6 s) are shown. WM load: load 3 vs. load 1 ($t = 2.20$); Attentional demand = NPO vs. PO ($t = 2.25$); Interaction = load 3 vs. load 1 x NPO vs. PO ($t = 2.15$); Talairach coordinates [x, y, z (in millimeters)] of the activation maxima are shown. BA = Brodmann Area; FG = frontal gyrus; IFG = inferior frontal gyrus; IPL = inferior parietal lobule; MFG = middle frontal gyrus;MTG = middle temporal gyrus; PHG = parahippocampal gyrus; PrcS = precentral sulcus; SFG = superior frontal gyrus; STG = superior temporal gyrus.

4.4.2 Discussion

The analysis of deactivation indicates three important findings. First, during WM encoding significant deactivation occurred in areas of the default network and this deactivation was modulated by task demand. This result is consistent with previous studies showing that the magnitude of deactivation is sensitive to the difficulty level of attention and WM tasks (McKiernan et al., 2003; Tomasi et al., 2006). Second, there was a dissociation between regions where task-induced deactivation was an inverse linear function of task demand with the strongest increase in deactivation when both WM load and search difficulty where high (left medial and lateral parietal cortex, posterior insula, temporal, and medial frontal regions) and regions where the amount of deactivation reached a limit already with load3/PO or with load1/NPO (right medial parietal cortex, left posterior insula, medial and superior frontal cortex). These increases in deactivation in response to WM load and attentional demand mirrored the observed increases in activation (Figure 4.10).

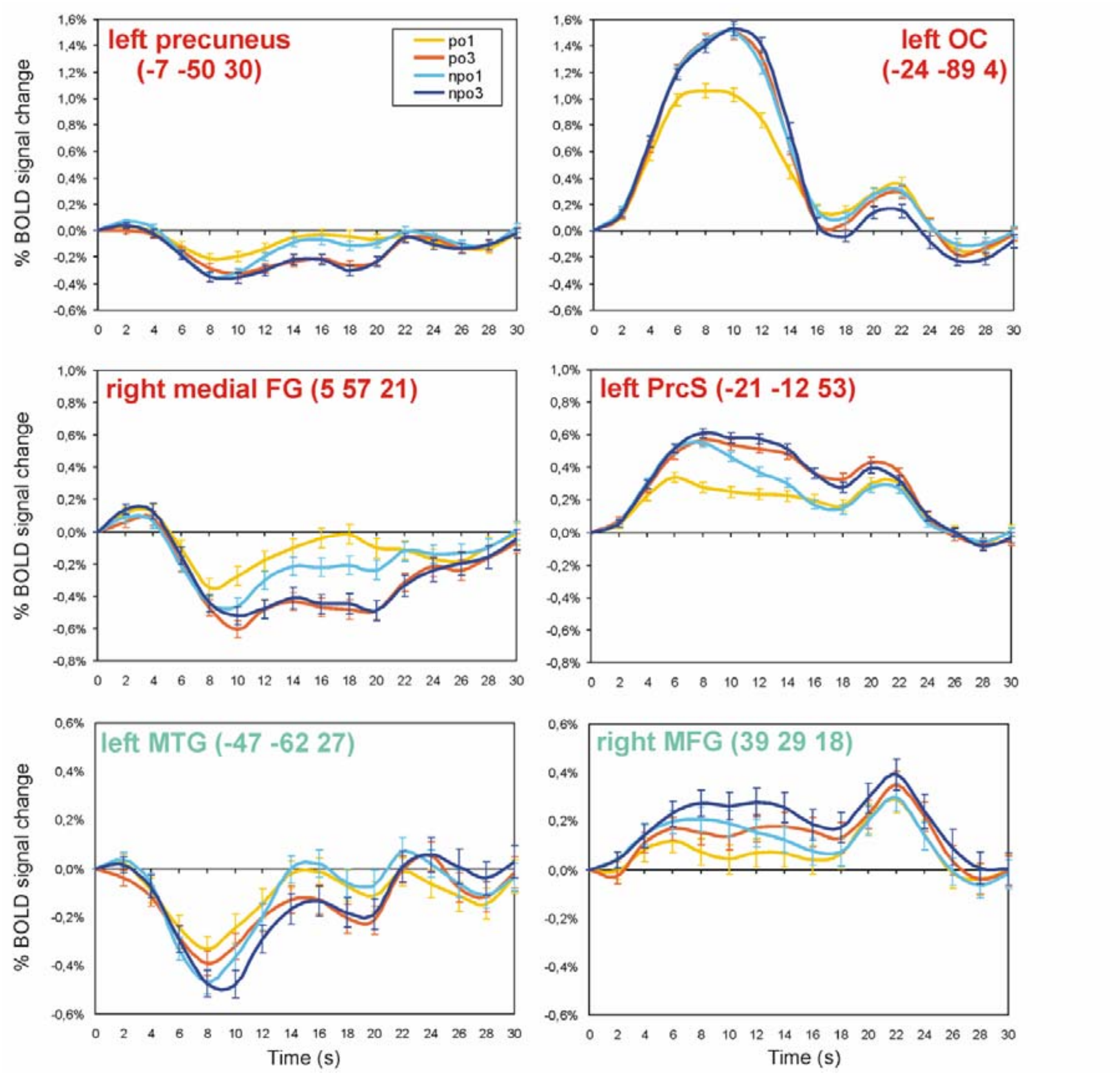


Figure 4.10 Significant deactivations and activations in the four experimental conditions during encoding (4-6 s). Time courses from areas that showed an interaction effect (red) and effects of both WM load and attentional demand (green) are shown. FG: frontal gyrus, OC: occipital cortex, MFG: middle frontal gyrus, MTG: middle temporal gyrus, PrcS: precentral sulcus.

Third, the patterns of deactivation in response to WM load and attentional demand overlapped in distributed regions but also showed a considerable degree of selectivity. A hemispheric differentiation was found with regions in the right and left hemisphere selectively deactivated in response to high WM load and attentional demand, respectively. The selectivity in location indicates that these decreases are dependent on the task characteristics. Task-dependent decreases have been most often found in sensory cortices (Laurienti et al., 2002). The present results suggest that they can also occur in frontal and parietal regions that are involved in higher cognitive processing.

What are the mechanisms underlying the observed deactivations? From a cognitive perspective it has been proposed that demand-dependent increases of BOLD deactivation reflect the reallocation of processing resources in task-irrelevant or default network regions (McKiernan et al., 2003; Tomasi et al., 2006; Chun and Turk-Browne, 2007). However, the present finding of attention-selective and WM-selective deactivations is inconsistent with the default mode theory. According to this account one would expect to find an increase in BOLD deactivation as a function of task demand in a highly similar set of regions for both the attention and WM components of the task. The present results raise the possibility that deactivations are functionally related to cognitive processing required in specific tasks. Supporting evidence for this idea has been provided by recent fMRI studies that successfully exploited the effects of BOLD deactivation on behavioural performance (Hampson et al., 2006; Shulman et al., 2007). For instance, when subjects monitored a stream of distracter objects for a target, the BOLD deactivation preceding the target in the right temporal-parietal junction, a region that overlaps with some of the default mode regions, was stronger on trials in which the target was detected than missed (Shulman et al., 2007). The authors argued that this deactivation might reflect the efficient filtering of distracting information, a process that might operate during both active tasks and resting states. Another approach investigating functional connectivity has been used by Hampson et al. (2006). Previously, functional connectivity between nodes in the default mode network at rest has been taken as evidence for the default mode theory that they function together during rest (Fox, 2005). Consistently, Hampson et al. (2006) demonstrated that two default mode regions, the posterior cingulate cortex and portions of the medial frontal gyrus and ventral anterior cingulate cortex, were functionally connected at rest. However, in contradiction to the default mode theory, functional connectivity between the two regions was also found during a verbal WM task and the strength of this functional connection was positively correlated with WM performance. The authors suggested a functional role of deactivation in facilitating or monitoring cognitive performance. However, assigning the specific processes to those regions that show task-induced deactivations remains rather speculative. Therefore, other explanations in terms of a neuronal or a haemodynamic origin need to be considered.

A recent study in human cerebral cortex (Shmuel et al., 2002) demonstrated prolonged BOLD deactivation in the medial occipital cortex in response to partial visual field stimulation. The deactivations were correlated with decreases in cerebral blood flow and a commensurately smaller decrease in the oxygen consumption rate. This reduction in the oxygen consumption rate was interpreted as evidence for a decrease in neuronal activity that triggers a reduction in cerebral blood flow. Additional support for a neuronal origin of BOLD deactivation was later provided by a monkey study showing that BOLD deactivation was associated with comparable decreases in local field potentials and multiunit activity (Shmuel et al., 2006). Importantly, these studies also indicated that the BOLD deactivation was tightly coupled with task-induced activations in neighbouring regions. Based on this coupling it has been suggested that lateral suppression mediated by inhibitory connections within striate and extrastriate cortex might account for the reduction in neuronal activity. It seems possible that long-range inhibitory interconnections might exist as well as there is evidence for suppressive influences that operate even across hemispheres (Sack et al., 2005). Therefore, the attention- and WM-selective patterns of deactivation observed in the present study may also be explained by specific inhibitory connections of task-relevant regions.

It has also been proposed that BOLD deactivation is due to the local shunting of cerebral blood flow to areas that are active from adjacent areas (Devor et al., 2005). However, in the present study deactivation occurred at great distance and across different vascular distributions from the activated brain regions. Also, activations in response to attentional demand appeared to be stronger in the right and to WM load in the left hemisphere. In contrast, deactivation in response to attentional demand appeared to be stronger in the left and to WM load in the right hemisphere. Therefore, an explanation in terms of neuronal processing seems to be likely although a haemodynamic component cannot be completely ruled out.

Taken together, the present study indicates that task-induced deactivations in WM and attention tasks including fronto-parietal regions rely on both the demand on cognitive processing and the characteristics of the specific task. Although the exact mechanisms underlying BOLD deactivations remain to be determined, these findings seem to be inconsistent with the hypothesis of a default mode of brain function.

Chapter 5 - Common neural substrates for encoding into spatial WM and selective attention

In fMRI experiment 2 the hypothesis was tested that the capacity limitation of spatial WM is due to common limited-capacity neural resources shared by spatial WM and selective attention. Visual search was combined with delayed-discrimination of locations and the demands on selective attention and encoding into spatial WM were independently modulated. The brain regions that were selectively responsive to either attentional demand or the encoding into spatial WM and those involved in both processes were identified.

5.1 Materials and methods

5.1.1 Participants

20 healthy participants (ten females, mean age 28.5 ± 3.7, range: 22-35) were recruited from an academic environment. Participants reported normal or corrected-to-normal visual acuity, normal colour vision, and no history of neurological or psychiatric illness. The study was approved by the local ethics committee. All participants gave written informed consent.

5.1.2 Stimuli, task, and procedure

The stimuli, task, and procedure were the same as in fMRI experiment 1, except for the following differences. Participants were instructed to memorise the *locations* of the objects marked with an L in 90° orientation (target items) while the objects associated with Ls of other orientations could be ignored (distractor items). In order to probe WM for target locations, after an 8-s delay interval the original stimulus array was presented without the centre items and with one of the shapes missing (Figure 5.1). Participants responded with a left- or right-hand button press to indicate whether the location of the missing shape did or did not match one of the target locations. WM load conditions included load 1, 3, or 5.

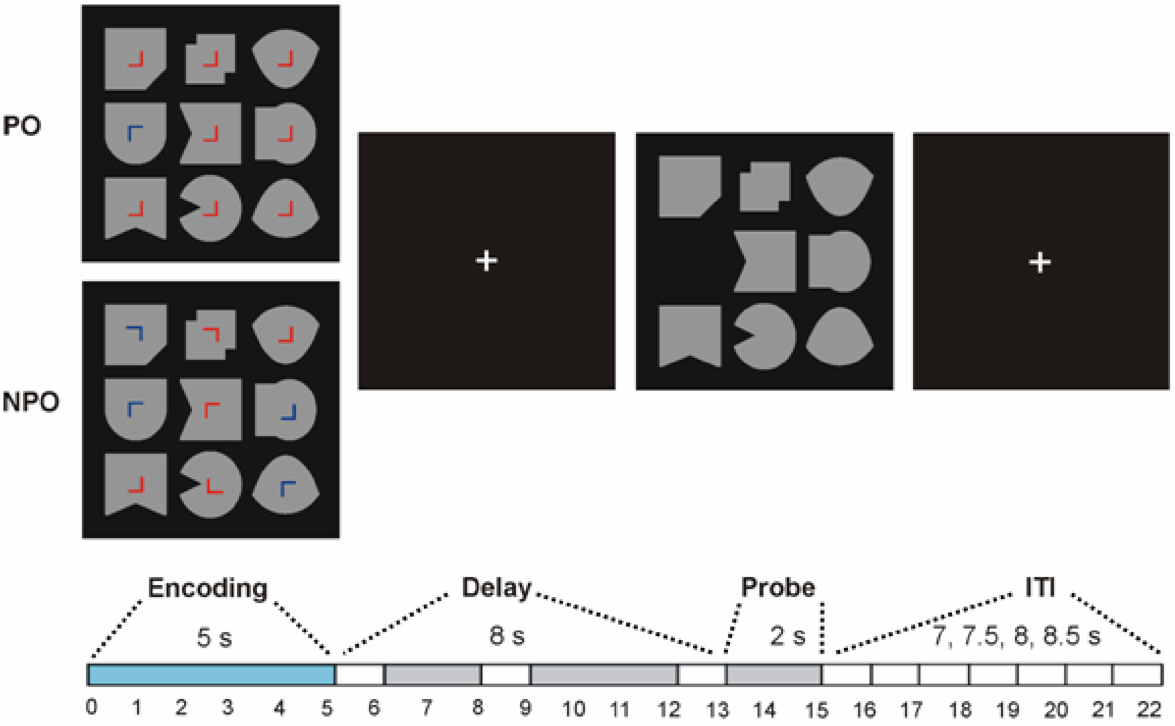


Figure 5.1 Trial design used in fMRI experiment 2. The search array was presented for 5 s and WM load was either load 1, 3, or 5. The analysis focused on the encoding predictor (blue bar, grey: additional predictors). PO: Pop-out, NPO: non pop-out, ITI: Inter-trial interval.

Each trial began with the presentation of the search array for 5 s. With regard to the individual presentation times derived from the behavioural study (Experiment 4, Figure 3.7, lower panel) it was expected that this time would be long enough to enable successful encoding of the locations into WM even in the most demanding condition (load5/NPO). The inter-trial interval lasted 7, 7.5, 8 or 8.5 s. The inter-trial interval was jittered in order to increase the effective sampling rate and the separability of overlapping functions modelling the different task phases (Ollinger et al., 2001). The experiment consisted of four runs with 30 experimental trials each, resulting in 20 trials for each of the six trial types (load1/PO; load3/PO; load5/PO; load1/NPO; load3/NPO; load5/NPO). PO and NPO conditions were presented in separate blocks of seven or eight trials (two blocks for each condition per run) in a pseudo-randomised order across runs. Within each block, WM load conditions were presented in a pseudo-randomised order to equal the number of WM load 1, 3, and 5 trials. Half of the trials were matches. 20% of the trials (4 trials of each condition) were partial trials that ended after the encoding phase without informing participants in advance. Thus, in these trials participants were required to encode the target locations into WM but not to maintain them during the delay and to retrieve them during the probe phase. These trials were pseudo-randomly interspersed and were included to compensate the overlap of the haemodynamic responses to successive neural events associated with the encoding and maintenance phases (Ollinger et al., 2001).

5.1.3 Image acquisition and analyses

Anatomical three-dimensional T1-weighted images (voxel size: 1.00 x 1.00 x 1.00 mm^3) and functional images were acquired on a 3 T Magnetom Trio scanner (Siemens Medical Systems, Erlangen, Germany) equipped with a standard head coil. Functional images were collected using 17 axial slices (5 mm thickness with 3.6 x 3.6 mm in-plane resolution) with a BOLD-sensitive EPI sequence (TR = 1 s, TE = 30 ms, FA = 80°; FOV = 220 mm, matrix = 64 x 64; duration of each run = 667 s). Trials were triggered by scanner pulses and presented with the ERTS software (Experimental Run-Time System, Berisoft, Frankfurt, Germany). Stimuli were back-projected from an LCD projector on to a screen viewed through a mirror by the supine subject in the MR scanner.

Image analyses were performed with Brainvoyager QX, version 1.8 (Brain Innovation, Maastricht, The Netherlands). Data preprocessing included slice scan time correction with the first scan time within a volume used as a reference for alignment by sinc interpolation, three-dimensional motion correction, spatial smoothing with an 4 mm Gaussian kernel (full width at half-maximum), temporal high pass filtering with a cut-off of 222 s to remove low-frequency non-linear drifts of three or fewer cycles per time course, and linear trend removal. Talairach transformation was performed for the complete set of functional data of each subject, yielding a 4-D data representation (volume time course: 3 x space, 1 x time). A multi-subject statistical analysis was performed by multiple linear regression of the BOLD response time course in each voxel. The general linear model of the experiment was computed for 77 z-normalised volume time courses. The data of three runs were excluded from the analysis due to technical problems with response recording during the scanning procedure. For the design matrix, four time points were defined per experimental condition, representing the different periods of the experiment (encoding: 0-5 s after stimulus onset; early delay: 6-8 s; late delay: 9-12 s; retrieval: 13-15 s). The signal values during these phases were considered the effects of interest. The early delay predictor was included to avoid that the activity captured by the late delay predictor was contaminated by encoding activity (Zarahn et al., 1997). The corresponding predictors were obtained by convolution of an ideal box-car response with a gamma function model of the haemodynamic response (Friston et al., 1998). All error trials were collapsed on a separate predictor.

3D group statistical maps were generated by associating each voxel with the F-value corresponding to the specific set of predictors and calculated on the basis of the least mean squares solution of the general linear model with a random-effects model. The obtained beta weights of each predictor served as input for the second-level whole-brain random-effects analysis including a 2 x 3 factorial design. Thus, the beta values of participants were treated explicitly as realisations of the two within-subjects factors attentional demand (level 1: PO, level 2: NPO) and WM load (level 1: load 1, level 2: load 3, level 3: load 5), which allowed to directly test for main and interaction effects between the two based on F-statistics. To compare activations between experimental conditions within one task phase, linear contrasts were performed using t-statistics. Multi-subject statistical maps were thresholded at $q < 0.05$, corrected for false discovery rate (Genovese et al., 2002) and visualised on a surface reconstruction of the MNI template brain (courtesy of the Montreal Neurological Institute). FMRI time courses were shown for selected regions of interest (ROI) where the effects of WM load and attentional demand appeared most prominently. ROIs were functionally defined based on the multi-subject statistical maps overlaid on the cortical surface map of the MNI template brain. Starting from the voxel showing peak activation in the multi-subject map, a surface patch of 30 mm^2 (4 mm thickness) was marked. Representative time courses for each experimental condition were obtained by averaging the percent signal changes of the individual voxels within the obtained volume across all participants and repetitions.

5.2 Results

5.2.1 Behavioural performance

An ANOVA tested the effects of search difficulty and WM load on response accuracy and RT at test (Figure 5.2). Participants' WM performance at test was equally good under PO and NPO search conditions [WM load 1, 95.6% & 95.2% correct, respectively; WM load 3, 91.6% & 93.9% correct; WM load 5, 90.0% & 92.6%, ANOVA, $F(1, 19) = 3.13$, $p = 0.09$]. There was a main effect of the factor search difficulty on RTs [$F(1, 19) = 8.7$, $p < 0.01$]. However, post hoc t-tests revealed a significant difference between PO and NPO conditions only within WM load 1 [WM load 1, 754 ms & 695 ms, respectively, $t(19) = 4.1$, $p < 0.01$; WM load 3, 920 ms & 904 ms, $t(19) = 1.3$, $p = 0.19$; WM load 5, 1050 ms & 1038 ms, $t(19) = 0.57$, $p = 0.57$]. Differences in performance between the WM load conditions were significant for accuracy [$F(2, 38) = 4.79$, $p < 0.05$] and for RTs [$F(2, 38) = 108.62$, $p < 0.001$]. Response accuracy declined from WM load 1 to load 3 and from WM load 3 to load 5 in both search conditions. Post hoc t-tests of differences between successive levels of WM load indicated that accuracy was significant lower for WM load 5/PO vs. load 1/PO [$t(19) = 2.75$, $p < 0.05$, all other t-values < 1.67, p-values > 0.11]. In both search conditions, RTs were significantly slower for WM load 5 vs. load 1, for WM load 3 vs. load 1 and for WM load 5 vs. load 3 (all t-values > 6.41, all p-values < 0.001]. There were no significant interactions between search difficulty and WM load [$F(2, 38) = 0.84$, $p = 0.42$ for accuracy; $F(2, 38) = 2.76$, $p = 0.08$ for RTs].

The finding that memory performance at test did not differ between PO and NPO conditions indicates that the presentation time of the search array (5 s) was indeed sufficiently long to ensure that participants were able to complete the encoding process even in the most demanding condition (load 5/NPO). Therefore, the task was suitable for probing common and selective activations for visual search and spatial WM encoding with event-related fMRI. Overall response accuracy was high (on average 92.6% and 93.9% correct for PO and NPO conditions, respectively). Therefore, it seems likely that the lack of significant differences in response accuracy between WM load conditions can be explained by a ceiling effect.

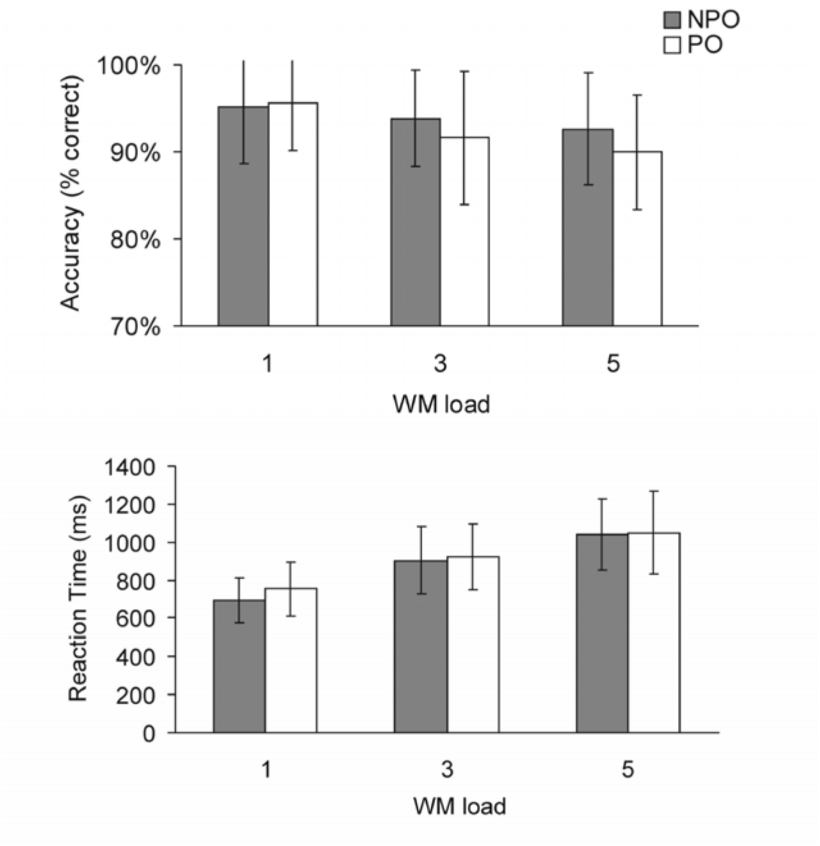


Figure 5.2 Mean response accuracy and reaction times in the six experimental conditions. Bars represent standard errors of the mean. PO: pop-out, NPO: non pop-out.

5.2.2 Brain systems for visual attention and encoding into spatial WM

The analyses of fMRI data for the encoding predictor (0-5 s after stimulus onset) revealed a high degree of overlap in the brain areas that showed a significant main effect of visual search difficulty and those that showed a significant main effect of WM load. Post-hoc contrasts indicated that the latter effect was mainly driven by the difference in activation between WM load 5 vs. load 1 [(load5/PO + load5/NPO) – (load1/PO + load1/NPO)] and therefore, the results for this contrast are shown (Figure 5.3, Table 5.1). Overlap in activation associated with higher activation for NPO vs. PO and higher activation for WM load 5 vs. load 1 was observed bilateral in the lateral and medial occipito-temporal cortex (middle and inferior occipital gyrus, middle temporal gyrus, fusiform gyrus, and parahippocampal gyrus). Parietal areas were activated around the IPS, extending into both the inferior and superior parietal lobules and precuneus. Frontal activation occurred along the PrcS including the intersections between precentral and middle and inferior frontal gyri, in the frontal midline, and anterior insula. Subcortical activations were found in the thalamus. The common pattern of brain regions involved during both visual search and spatial WM encoding also included parts of the mid-MFG. Overlapping activations in these regions were more pronounced in the right hemisphere. In contrast, anterior parts of the MFG and IFG bilateral were selectively responsive to increased WM load. These regions showed a time shift in the maximum amplitudes of BOLD activation with a later peak in the NPO compared to the PO condition (Figure 5.4, blue). Posterior parts of the MFG responded selectively to high attentional demand. In the posterior cortex, areas selectively responsive to WM load were found in the right posterior insula, the left anterior IPL, and bilateral in the temporal cortex. Here again, BOLD activation peaked later in the NPO compared to the PO condition without any compromise on the size of the WM load effect. Activation in the occipital cortex preferentially increased in the NPO condition.

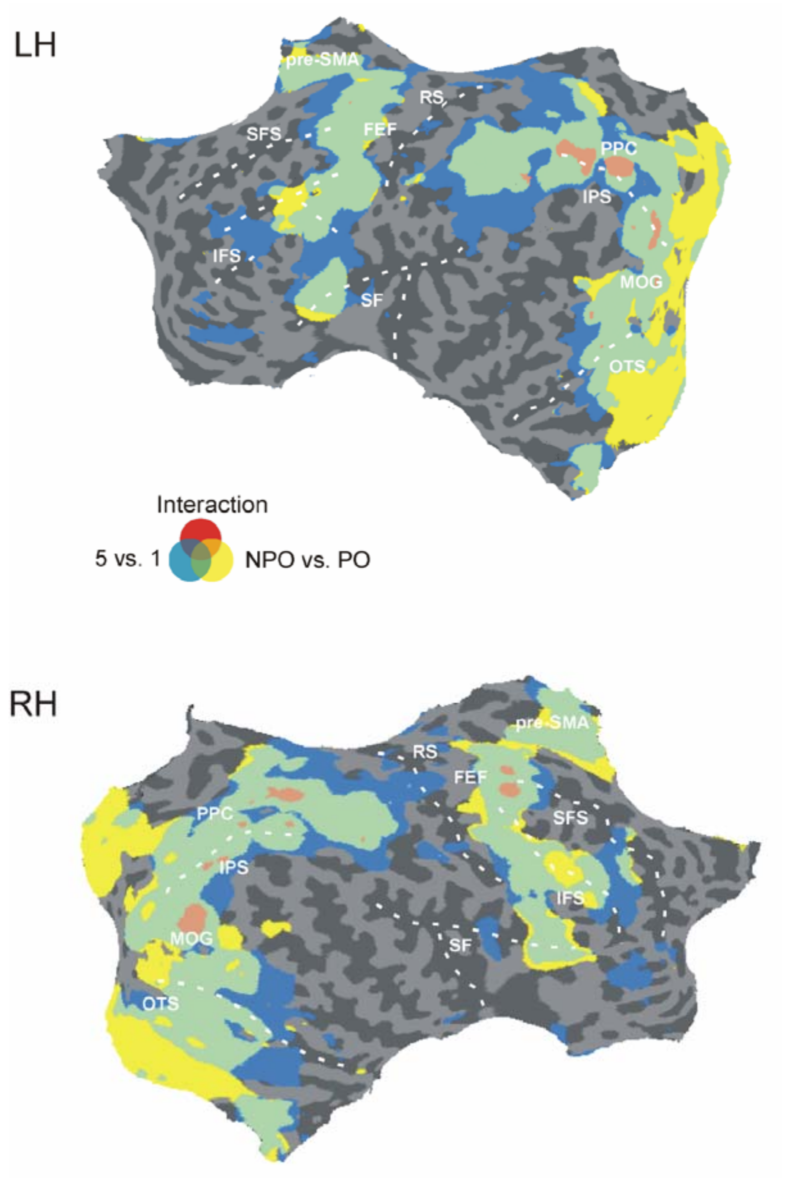
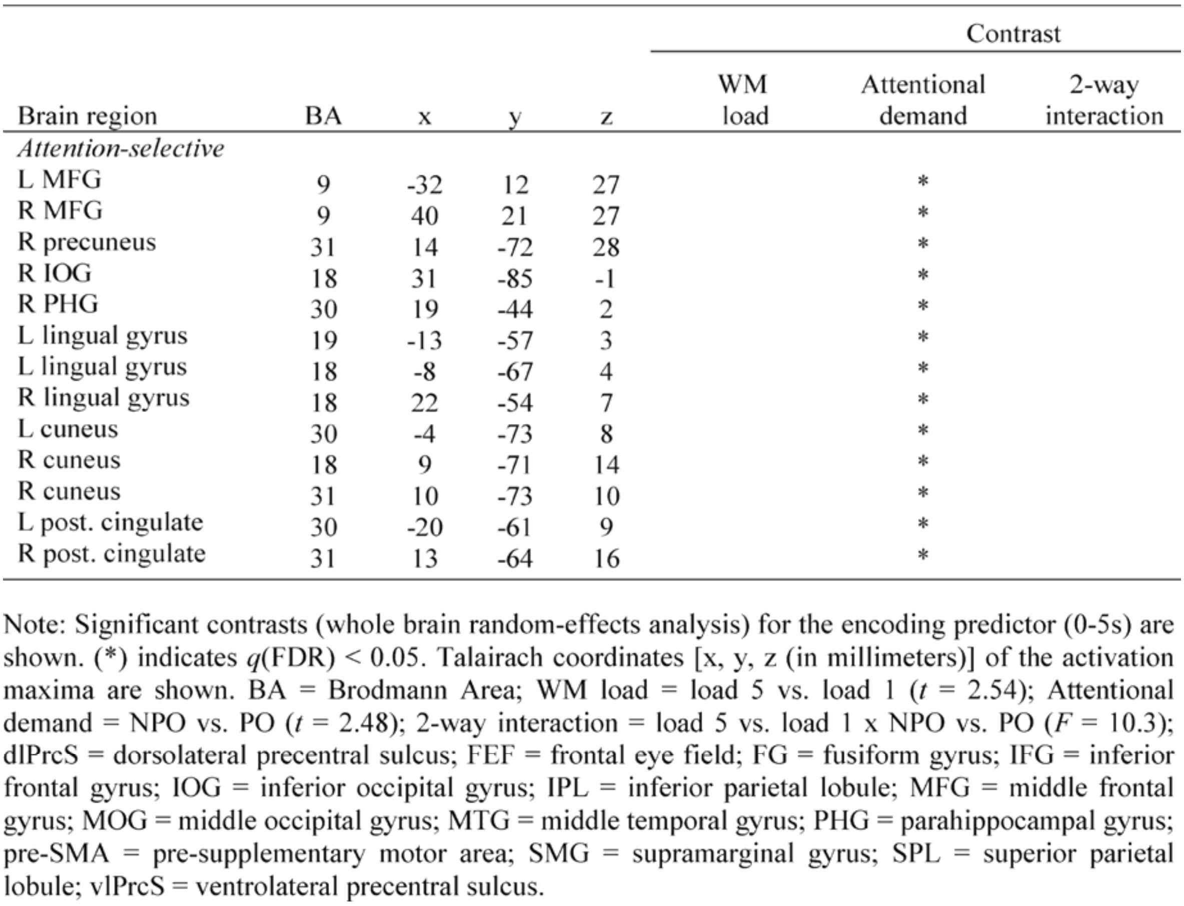


Figure 5.3 Group results for the encoding predictor (0-5 s). Statistical maps of the contrasts NPO vs. PO (yellow), WM load 5 vs. 1 (blue), and the significant 2-way interaction of search difficulty x WM load (red) are projected on the flattened surface reconstruction of the MNI template brain (courtesy of the Montreal Neurological Institute) (LH: left hemisphere, RH: right hemisphere). Activations are those exceeding a whole-brain false discovery rate threshold of q(FDR) < 0.05. FEF: frontal eye field, IFS: inferior frontal sulcus, IPS: inferior parietal sulcus, MOG: middle occipital gyrus, OTS: occipito-temporal sulcus, PPC: posterior parietal cortex, pre-SMA: pre-supplementary motor area, RS: rolandic sulcus, SF: Sylvian fissure, SFS: superior frontal sulcus.

Table 5.1 Brain regions showing significant activation in the contrasts for encoding

					Contrast		
Brain region	BA	x	y	z	WM load	Attentional demand	2-way interaction
Common activation							
L MFG	9	-38	20	29	*	*	
R MFG	46	40	29	27	*	*	
L dlPrcS	6	-21	-3	57	*	*	*
R dlPrcS	6	26	-4	61	*	*	
L FEF	6	-27	-9	47	*	*	
R FEF	6	27	-4	48	*	*	*
L vlPrcS	6	-44	2	29	*	*	
R vlPrcS	6	48	6	31	*	*	
L pre-SMA	6	-6	2	54	*	*	
R pre-SMA	6	5	4	54	*	*	
L insula	13	-30	18	12	*	*	
R insula	13	34	20	10	*	*	
L IPL	40	-40	-30	40	*	*	
R IPL	40	36	-37	45	*	*	
L IPL	40	-38	-40	41	*	*	*
R IPL	40	45	-29	40	*	*	*
L SPL	7	-19	-60	52	*	*	*
R SPL	7	24	-61	56	*	*	*
L precuneus	7	-17	-64	48	*	*	*
R precuneus	19	28	-68	31	*	*	*
L MOG	19	-26	-77	23	*	*	*
R MOG	19	33	-74	18	*	*	*
R MOG	31	43	-68	4	*	*	
L IOG	19	-35	-79	1	*	*	
R IOG	19	33	-76	-2	*	*	
L FG	37	-37	-60	-6	*	*	
R FG	37	43	-54	-6	*	*	
L MTG	37	-46	-60	0	*	*	*
R MTG	37	45	-57	1	*	*	
L PHG	19	-27	-45	-5	*	*	
R PHG	19	8	-44	-4	*	*	
L lingual gyrus	18	-26	-71	-11	*	*	
R lingual gyrus	19	26	-65	-2	*	*	
L thalamus		-15	-28	0	*	*	
R thalamus		14	-26	1	*	*	
WM-selective							
L MFG	9	-43	29	26	*		
L MFG	10	-37	45	22	*		
R MFG	46	41	37	22	*		
L IFG	47	-23	31	-10	*		
R insula	13	37	-3	15	*		
L IPL	40	-54	-29	33	*		
L SMG	40	-44	-42	36	*		
R MTG	21	51	-49	-1	*		
L FG	37	-29	-39	-12	*		
R PHG	36	28	-34	-11	*		

Table 5.1 (continued)



					Contrast		
Brain region	BA	x	y	z	WM load	Attentional demand	2-way interaction
Attention-selective							
L MFG	9	-32	12	27		*	
R MFG	9	40	21	27		*	
R precuneus	31	14	-72	28		*	
R IOG	18	31	-85	-1		*	
R PHG	30	19	-44	2		*	
L lingual gyrus	19	-13	-57	3		*	
L lingual gyrus	18	-8	-67	4		*	
R lingual gyrus	18	22	-54	7		*	
L cuneus	30	-4	-73	8		*	
R cuneus	18	9	-71	14		*	
R cuneus	31	10	-73	10		*	
L post. cingulate	30	-20	-61	9		*	
R post. cingulate	31	13	-64	16		*	

Note: Significant contrasts (whole brain random-effects analysis) for the encoding predictor (0-5s) are shown. (*) indicates q(FDR) < 0.05. Talairach coordinates [x, y, z (in millimeters)] of the activation maxima are shown. BA = Brodmann Area; WM load = load 5 vs. load 1 (t = 2.54); Attentional demand = NPO vs. PO (t = 2.48); 2-way interaction = load 5 vs. load 1 x NPO vs. PO (F = 10.3); dlPrcS = dorsolateral precentral sulcus; FEF = frontal eye field; FG = fusiform gyrus; IFG = inferior frontal gyrus; IOG = inferior occipital gyrus; IPL = inferior parietal lobule; MFG = middle frontal gyrus; MOG = middle occipital gyrus; MTG = middle temporal gyrus; PHG = parahippocampal gyrus; pre-SMA = pre-supplementary motor area; SMG = supramarginal gyrus; SPL = superior parietal lobule; vlPrcS = ventrolateral precentral sulcus.

5.2.3 Interference between attention and spatial WM encoding

Brain areas reflecting functional interference between attention-demanding visual search and spatial WM encoding were identified by the interaction between the factors attentional demand (NPO, PO) and WM load (load 1, 3, and 5). Activation associated with a significant interaction effect was found in a subset of the regions with overlapping activations for the attention and WM load contrasts. These regions included the lateral MOG, the lateral and medial parietal cortex (IPL, SPL, precuneus), and the dorsal PrcS including the FEF. Activation in the latter region was more pronounced in the right hemisphere. Activation in the temporal cortex appeared only in the left hemisphere (Table 5.1 and Figure 5.3, red colour). Time course analyses of these regions showed a smaller increase in the BOLD signal with increasing WM load for NPO compared to PO search (Figure 5.4, red). In both search conditions the BOLD response increased from WM load 1 to WM load 3. Activation further increased when participants needed to memorise five locations, however this increase was smaller in the NPO condition than the PO condition. In contrast, in adjacent brain regions that showed an overlap in activation but no interaction, the increase in activation across WM load conditions did not differ between NPO and PO search conditions (Figure 5.4, green).

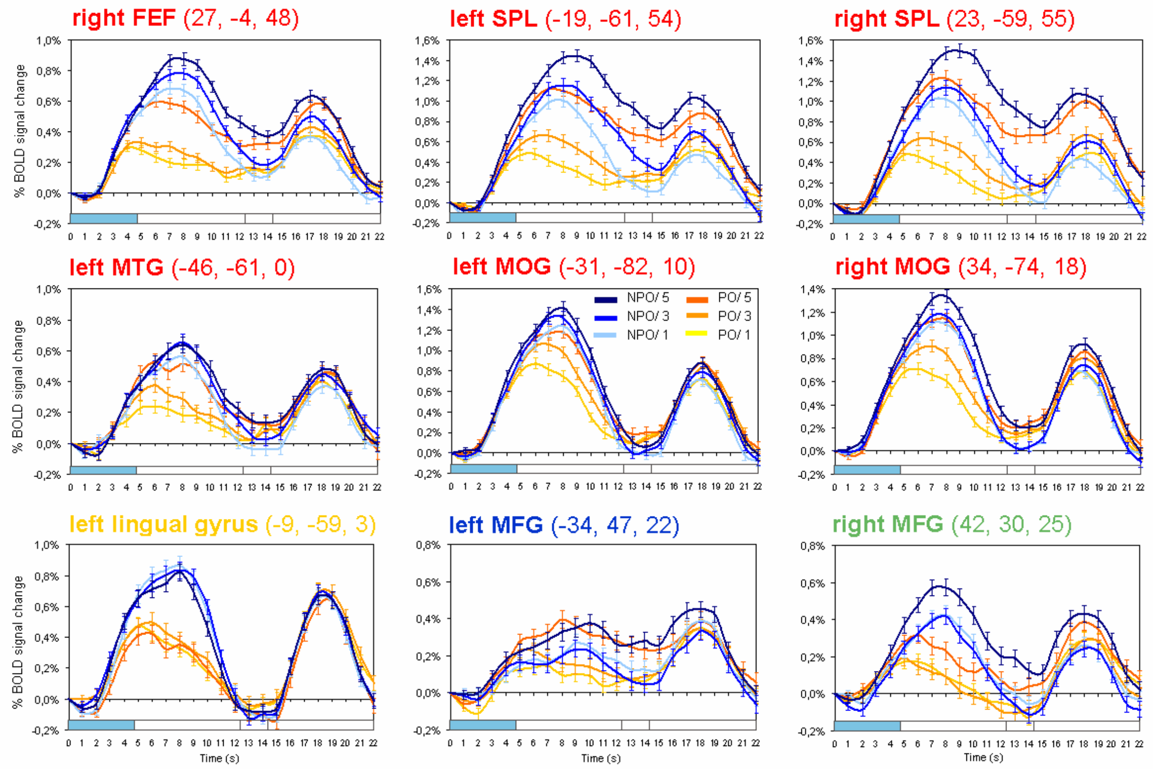


Figure 5.4 Averaged time courses of the BOLD signal in the six experimental conditions from selected regions showing an interaction effect (red), an effect of attentional demand (yellow), an effect of WM load (blue) and effects of both manipulations (green). Bars represent standard errors of the mean. FEF: frontal eye field, MFG: middle frontal gyrus, MOG: middle occipital gyrus, MTG: middle temporal gyrus, SPL: superior parietal lobe.

5.2.4 Load effects during spatial WM encoding

It was further examined how activity in different cortical regions within the spatial WM circuit varied with different WM loads. To investigate effects of WM load that were not influenced by concurrent demands on attentional processing, post-hoc contrasts included only PO conditions [(load 3/PO – load 1/PO); (load 5/PO - load3 /PO)]. The results indicated that activity increased from WM load 1 to WM load 3 and from WM load 3 to WM load 5 bilateral in the occipito-parietal cortex (MOG, SPL, IPL, and precuneus), along the PrcS, and the left pre-SMA (Figure 5.5, overlap). Within these regions there was a segregation of areas where BOLD activation increased monotonically with increasing WM load (cuneus, precunues, MOG, SPL) (Figure 5.5, left panel) and areas where the increase in BOLD activation was stronger between high WM load conditions (load 5 vs. load 3) compared to low WM load conditions (load 3 vs. load 1) (right IPL, PrcS, left pre-SMA) (Figure 5.5, middle panel). The latter activation pattern appeared even more pronounced in the PFC, right pre-SMA, anterior insula, left anterior IPL, and the occipito-temporal cortex. Here, BOLD activation did not differ between WM load conditions 1 and 3, increasing only in the highest WM load condition (Figure 5.5, dark blue and right panel).

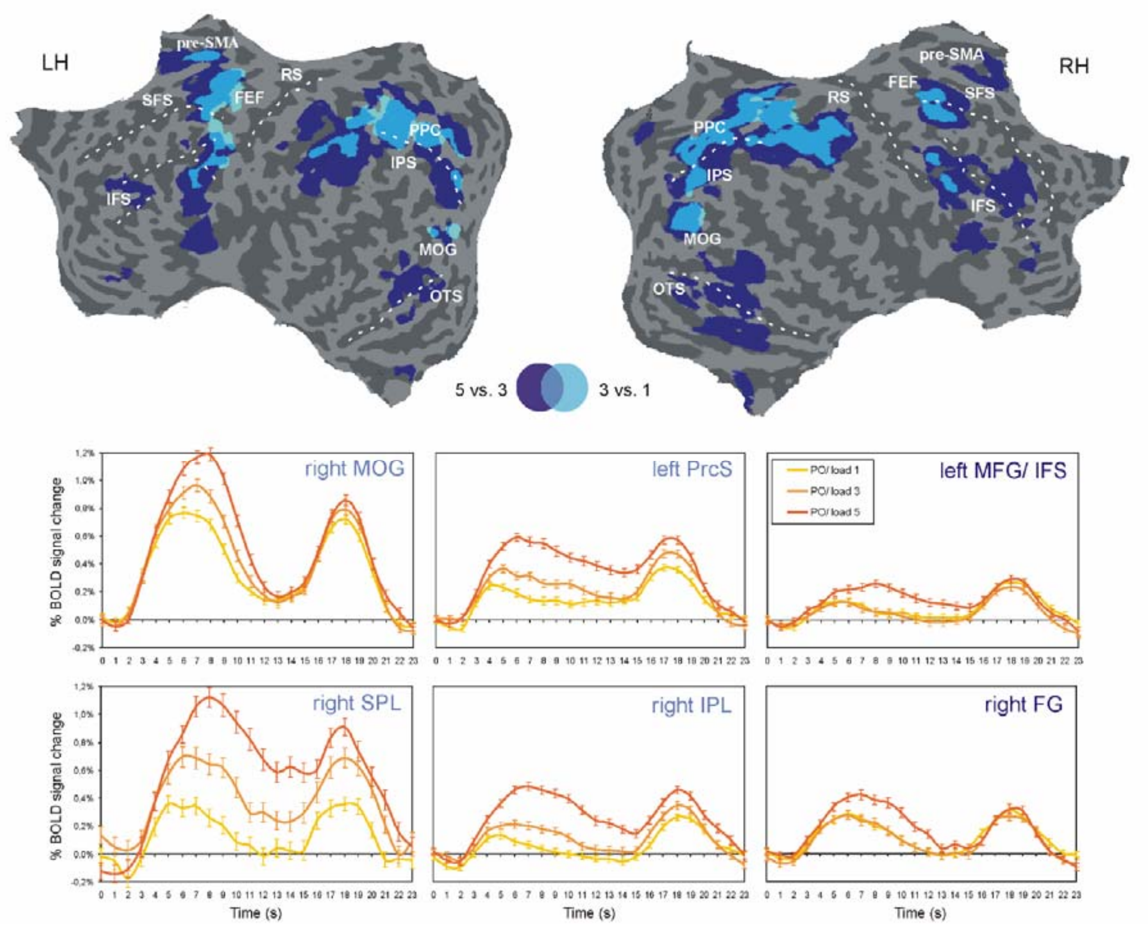


Figure 5.5 WM load effects during encoding (0-5 s). Statistical maps of the contrasts WM load 5 vs. load 3 (dark blue) and WM load 3 vs. 1 (bright blue) (overlap in middle blue) are projected on the flattened surface reconstruction of the MNI template brain. Activations are those exceeding a whole-brain false discovery rate threshold of q(FDR) < 0.05.

5.2.5 Load effects during spatial WM maintenance

It was expected that if participants successfully performed the WM task despite the concurrent demands on attentional resources the observed effect of interference between search difficulty and WM load should be restricted to the encoding phase. Consistent with this prediction, the interaction contrast between search difficulty and WM load did not yield significant activation during the late delay phase (9-12 s after stimulus onset). Neither did delay activation increase in the NPO condition compared to the PO condition. Thus, the process of active maintenance of locations in WM was not limited by attentional processing required by difficult visual search. Linear contrasts indicated significant stronger activation in frontal (MFG, PrcS) and parietal regions (SPL, IPL, precuneus) when participants needed to maintain 5 vs. 1 location. Delay activity did not differ between WM loads 1 and 3 and WM loads 3 and 5.

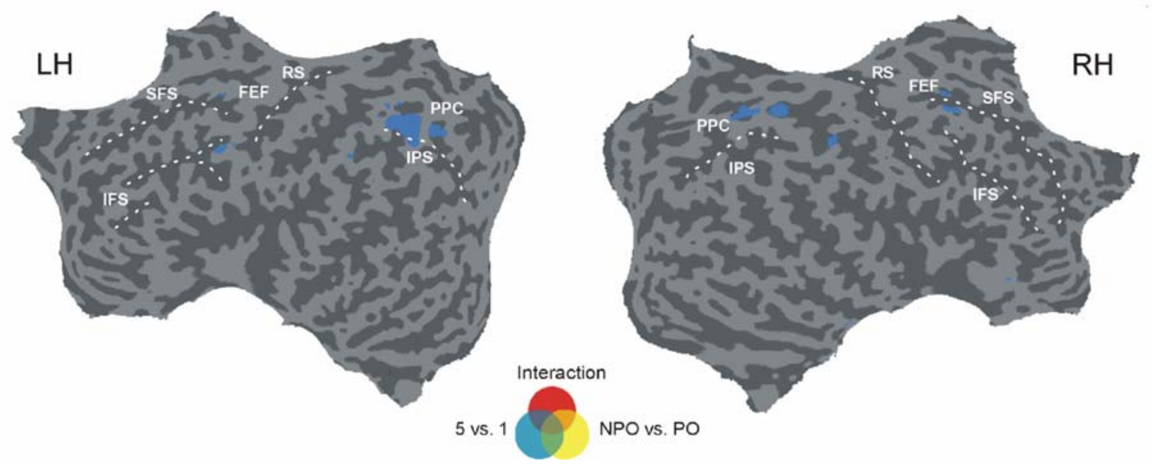


Figure 5.6 Group results for the late delay predictor (9-12 s). Significant activations were found only for the contrast WM load 5 vs. load 1. Activations are those exceeding a whole-brain false discovery rate threshold of *q*(FDR) < 0.05. FEF: frontal eye field, IFS: inferior frontal sulcus, IPS: inferior parietal sulcus, PPC: posterior parietal cortex, RS: rolandic sulcus, SFS: superior frontal sulcus.

Table 5.2 Brain regions showing significant activation in the contrasts for delay

					Contrast		
Brain region	BA	x	y	z	WM load	Attentional demand	2-way interaction
WM-selective							
L FEF	6	-28	-6	53	*		
R FEF	6	26	-1	47	*		
L vlPrcS	6	-44	-1	36	*		
L IPL	40	-45	-38	50	*		
R IPL	40	38	-28	36	*		
L SPL	7	-28	-58	51	*		
R SPL	7	24	-53	53	*		
L precuneus	7	-21	-59	51	*		
R precuneus	7	23	-60	49	*		

Note: Significant contrasts (whole brain random-effects analysis) for the late delay predictor (9-12 s) are shown. (*) indicates q(FDR) < 0.05. Talairach coordinates [x , y, z (in millimeters)] of the activation maxima are shown. WM load = load 5 vs. load 1 (t = 3.75); BA = Brodmann Area ; FEF = frontal eye field; IPL = inferior parietal lobule; SPL = superior parietal lobule; vlPrcS = ventrolateral precentral sulcus.

5.3 Discussion

In this experiment visual search was combined with delayed discrimination of spatial locations within one single task and the demands on selective attention and WM encoding were independently modulated. The goal was to identify the brain regions that were selectively responsive to either attentional demand or spatial WM load and those involved in both processes. It was hypothesised that if spatial WM and selective attention were subserved in part by common areas with limited neural processing capacity, activation in these areas under conditions of joint demand on both processes should reach a plateau or at least be less than additive, as reflected in a statistical interaction between attention and WM. Conversely, an additive increase in BOLD activation under simultaneous WM and attentional demands was expected in regions whose processing capacity was not exceeded. The BOLD signal in these overlap regions should increase to the same degree with WM load under low and high attentional demand.

5.3.1 Common activation for visual attention and encoding into spatial WM

Overlapping activation for attention-demanding visual search and encoding into spatial WM was observed in distributed posterior and frontal regions. In the majority of these regions the overlap was associated with an additive increase in

BOLD activation under high demands on attention and WM. These results are in agreement with the view that the processes underlying attention-demanding visual search and the encoding into spatial WM require access to common neural and cognitive resources (Awh and Jonides, 2001). The additive increase in BOLD activation suggests that the demands on these regions were well within their processing limits even in the condition where high WM load was combined with difficult search. Conversely, in a subset of the overlap regions including the visual, parietal, and premotor cortex, an interaction effect between the two task manipulations was found. Activation increased from WM load 1 to WM load 5 but this increase was significantly smaller in the NPO compared to the PO condition. In addition, PFC activation associated with increased WM load was delayed rather than reduced under high attentional demand. These results indicate that competition for processing resources that are shared by the spatial WM and attention systems can lead to a severe limitation of neural processing capabilities and provide evidence that the two cognitive domains tap into common neural resources.

The brain areas mediating these common processing limitations of spatial WM and attention included regions that are classically considered to support goal-directed visuospatial attention (Kanwisher and Wojciulik, 2000; Corbetta and Shulman, 2002; Pessoa et al., 2003). The occipito-parietal and the premotor cortex have been implicated in the capacity limitation of visual WM (Linden et al., 2003; Todd and Marois, 2004; Marois and Ivanoff, 2005; Xu and Chun, 2006), however comparable evidence in the spatial domain still lacks. In one study, Leung et al. (2004) asked subjects to maintain one to four locations during a 15s-delay period. Consistent with the findings from the visual domain, activity in the parietal cortex monotonically increased from load 1 to load 3 levelling off in the highest load condition. Activity in the MFG showed an increase from load 1 to load 2 and decreased from load 2 to load 3 and from load 3 to load 4. The authors interpreted both activation patterns in terms of a neural limitation of spatial WM capacity. However, because fMRI activity was not correlated with behavioural performance (as has been done in the studies on visual WM capacity), the roles of parietal and prefrontal cortex in the limitation of spatial WM capacity remained unclear. The present finding indicating common processing limitations of encoding into spatial WM and attention in posterior but not prefrontal regions points to differential contributions of these regions to the capacity constraints of spatial WM.

5.3.2 Evidence for a neural bottleneck of visual attention and encoding into spatial WM?

Similar to fMRI experiment 1, the present study revealed a reduced WM load effect under conditions of NPO versus PO in distributed regions in the visual, parietal and premotor cortex. This limitation might be interpreted in terms of a neural bottleneck for joint demand on attention and WM resources during the stage of WM encoding. As outlined in chapter 4 several arguments might be raised against this interpretation among them i) effects of haemodynamic saturation of the neurovascular system, ii) limitations on perceptual rather than memory processes in the visual cortex, and iii) insufficient time available for WM encoding in the NPO condition.

As mentioned previously, the finding that checkerboard stimulation with similar scanning parameters can lead to BOLD signal changes of up to 4% in the occipital cortex (Uludag et al., 2004) is inconsistent with the first argument as this is about three-fold larger activation than the maximum increases in BOLD activation found in the present experiment (BOLD activation increased up to 1.4 % signal change in visual areas, up to 1.5 % signal change in parietal area, and up to 0.9% signal change in the right dorsal PrcS). Moreover, regions associated with an overlap in activation but no interaction showed a further increase in BOLD activity from load3/NPO to load5/NPO and from load5/PO to load5/NPO. Such an additive increase appeared in distributed regions adjacent to those showing an interaction effect, for instance in the PFC, the ventral PrcS, the anterior parietal, and visual cortex. This suggests that the interaction effect also resulted from differential processing induced by the task manipulations rather than from haemodynamic saturation.

The temporal resolution of fMRI and the design characteristics of the present experiment including a rather long encoding period made it difficult to separate perceptual processing from memory processing. However, it seems unlikely that the interaction effect observed in the visual cortex was owed exclusively to limitations on perceptual processes. As in the fMRI experiment 1, participants were not informed about the number of targets presented in the upcoming array. Therefore, they had to serially scan the entire array of nine items in order to find the single target in the NPO condition. In contrast, in the PO condition they immediately focused the target items. Thus, in the load5/PO condition only five as opposed to nine items in the load1/NPO condition had to be processed. If memory processing had not played a role, a further increase in activation for load1/NPO compared to load5/PO due to a higher perceptual load, would have been expected, which, however, was not observed. Thus, activation in the visual cortex was not solely a result of limitations on perceptual processing but rather reflected both perceptual and WM encoding-related processes.

Although participants needed more time for target-distractor discrimination in the NPO vs. the PO condition (see behavioural study, Experiment 4) insufficient time available for WM encoding in the NPO conditions should not explain the smaller effect of WM load under difficult vs. easy visual search. First, response accuracy and RTs at test did not differ between the two search conditions. Second, delay activity increased to the same degree from WM load 1 to load 5, irrespective of search difficulty. Third, there was an additive increase in activation with high WM load and high attentional demand that appeared in distributed regions. Taken together, these results indicate that even under difficult visual search participants efficiently engaged into the process of encoding into spatial WM.

5.3.3 Selective activations for visual attention and encoding into spatial WM

As participants applied attentional and WM processes to the same stimulus displays the present experiment allowed to identify the brain areas that were selectively responsive to either the demands on visual attention or the process of encoding into spatial WM. Areas selectively sensitive to high attentional demand appeared in early and higher visual areas and in a posterior region of the MFG adjacent to the precentral gyrus. Activity in the visual cortex has been consistently shown in studies on visual search (Leonards et al., 2000; Nobre et al., 2003). The visual cortex is also almost ubiquitously activated in tasks of visual attention and might reflect the site of attentional effects that are controlled by parietal and/or frontal regions (Kastner and Ungerleider, 2000; Pessoa et al., 2003). The involvement of the lateral PFC during inefficient visual search has not been consistently reported. One possibility is that activity in DLPFC and VLPFC revealed during difficult visual search reflects WM processes that guide the allocation of selective attention (Anderson et al., 2007). This idea may be supported by the present finding of prefrontal regions that respond to both attentional demand and WM load. In addition, the present study revealed areas in PFC that were responsive to attentional demand but not WM load. This attention-selective activation indicates that the lateral PFC contributes to various processes during visual search. Candidate functions may include spatial search strategies, duration or complexity of array exploration (Gitelman et al., 2002) or sustained attention (Coull et al., 1996).

Areas selectively sensitive to WM load were found in lateral parts of the PFC (MFG, IFG) slightly more pronounced in the right hemisphere. Here, activation extended into more dorsal parts of the MFG. Regions in the posterior cortex included the right posterior insula, the left anterior inferior parietal lobule, and the temporal cortex bilateral. Similar to the findings from fMRI experiment 1, these WM-selective regions showed a time shift in activation associated with the increase in WM load between the PO and the NPO condition. These results indicate that a similar mechanism was involved when either objects or locations needed to be encoded into WM under concurrent demands on attentional processing. As discussed in chapter 4 the candidate mechanism might allow participants to compensate for the common demands on limited neural resources shared by attention and WM processes in the posterior cortex.

Activation selectively responsive to the demands on WM encoding differed with respect to the stimulus domain within the lateral PFC. Activation related to object WM was strongly lateralised to the left hemisphere (MFG, IFG), whereas prefrontal activation related to spatial WM encoding appeared slightly more pronounced in the right hemisphere. Moreover, a dorsal-ventral gradient for the encoding of spatial and object information was observed. For object WM load prefrontal activation extended into mid-ventrolateral PFC, into a region around the IFS (BA 45) adjacent to the precentral sulcus of the left hemisphere. This region comprises a part of Broca's area that is involved in language processing and speech production (Amunts et al., 1999) but has been also associated with WM storage of shapes or colour information (Manoach et al., 2004; Mohr et al., 2006). For spatial WM load activation in the right PFC extended into more dorsal and anterior parts of the MFG (BA 9, BA 10). Other regions in the mid-dorsolateral PFC (bilateral BA 46, left BA10) showed WM-selective activation for both types of stimuli. This domain-independence indicates a role of mid-dorsolateral PFC in processes required by both the encoding of objects and locations into WM rather than in stimulus-selective mnemonic processing (see next paragraph and general discussion).

Furthermore, the investigation of pure effects of WM load that were not influenced by concurrent demands on attentional processing indicated a dissociation between areas where activation increased with each increase in WM load (PrcS, occipito-parietal cortex) and areas where activation increased only at the highest load level (lateral PFC, insula, occipito-temporal cortex). The former regions largely overlapped with those that were associated with limited neural processing capacity for joint demand on WM and attentional demands. Cognitive processes related to memory encoding should be more strongly engaged with each increase in memory load (Curtis and D'Esposito, 2003). Thus, it seems likely that the processes that shared common neural resources with selective attention during the encoding phase could indeed be attributed to the mnemonic requirements of the task. Also consistent with this interpretation is the finding that activity in the PrcS and parietal cortex remained high throughout the delay period. In contrast, the mid-dorsolateral PFC, the anterior insula, and regions in the occipito-temporal cortex may not be involved in memory processing per se but may support other nonmnemonic functions necessary during WM encoding. For instance, higher levels of stimulus complexity demand greater strategic or organisational processing in order to facilitate performance (Glahn et al., 2002; Bor et al., 2003). These processes have been associated, during WM encoding, with the corecruitement of the lateral PFC and regions in the fusiform gyrus known to be involved in the processing of object-based information (Bor et al., 2003). In the present task, the formation of configural representations or chunks of information might have been especially demanding when subjects needed to encode five positions leading to stronger activation in the PFC and temporal regions in this condition. Importantly, the processes supported by lateral PFC were not limited by attentional processes that constrained the activity in fronto-parietal regions. This finding again points to a differential role of prefrontal (e.g., strategic processing) and posterior regions (e.g., attention-based mnemonic processing) during WM encoding.

Chapter 6 - General discussion

The general aim of this dissertation was to characterise interactions between visual selective attention and visual WM. The conceptual link between visual WM and attention that was addressed stemmed from one characteristic feature of visual WM and attention, namely their limitation in capacity. Specifically, fMRI was used to test the hypothesis that the capacity limitation of visual WM is due to limited-capacity cognitive and neural resources shared with the process of visual selective attention.

To determine whether visual attention and WM are represented by different or common neural substrates the demands on visual attention and the process of encoding of information into WM were manipulated orthogonally within one unitary task. It was hypothesised that if visual WM and attention shared common limited-capacity neural resources, these resources would become exhausted in conditions that make high demand on both processes, thus resulting in interference. Such interference would indicate a limitation of the neural resources available for WM encoding and attentional processing. Two fMRI experiments were conducted that required subjects to encode either objects or locations into WM. Thus, the question whether effects of interference between visual attention and WM encoding are domain-specific or generalise across different classes of stimuli could be addressed.

The fMRI results demonstrated that visual selective attention and the encoding into visual WM share, to a high degree, common neural resources but show also some degree of selectivity. Common neural resources for visual attention and WM encoding appeared in distributed posterior and frontal regions. Most importantly, the results of both experiments revealed several visual, parietal, and premotor areas that showed overlapping activation for the two task components and were severely reduced in their WM load response under the condition with high attentional demand. Regions in the PFC were selectively responsive to WM load and differed to some degree depending on the WM domain. Here, activation associated with increased WM load was delayed rather than reduced under high attentional demand. These findings indicate that competition for resources shared by visual attention and WM encoding can limit processing capabilities in distributed posterior brain regions but not PFC and suggest a crucial role of attention in the encoding of both visual and spatial information into WM.

6.1 An attention-based model of visual WM encoding

A crucial role of selective attention for WM maintenance has been well established (Awh et al., 1998; Jha and McCarthy, 2000; Awh and Jonides, 2001; Postle et al., 2004; Lepsien and Nobre, 2007). The present dissertation focused on the encoding into WM and the findings suggest that an attention-based model applies to the encoding period as well. Moreover, the similarity in the effects of interference between attention and the encoding of objects or locations into WM indicates that the attention-based model of WM encoding is valid across different WM domains.

Why would attentional mechanisms be needed during the encoding of objects and locations into visual WM? It is well established that the active maintenance of information over short periods of time requires participants to engage in some form of rehearsal. In the spatial domain, such a mechanism can be supported by a retrospective perceptual code in terms of covert shifts of attention to the memorised locations (Awh and Jonides, 2001; Postle, 2006). On the neural level, attention-based rehearsal is accomplished by allocating attention via activity in the FEF and parietal cortex to extrastriate and parietal regions responsible for the perception of location (Awh and Jonides, 2001; Jha, 2002; Postle et al., 2004). In analogy to spatial WM maintenance, internal shifts of attention towards objects may underlie the rehearsal of objects in WM (Awh et al., 2006; Lepsien and Nobre, 2007). On the behavioural level, some evidence has been provided by the finding of attentional capture by objects that have been recently held in WM (Downing, 2000). Moreover, the behavioural study presented in this dissertation indicated interference between the attentional resources needed for detection of target locations and the WM resources needed for encoding of targets' shapes. On the neural level, it has been recently shown that orienting attention toward a representation of a face or a scene held within WM selectively increases maintenance-related activity in posterior regions specialised for processing the relevant objects (e.g. in the fusiform gyrus for faces and the parahippocampal gyrus for scenes) (Lepsien and Nobre, 2007). In addition, this study revealed regions in the parietal and medial and lateral prefrontal cortex that were involved in the control of object-based attention presumably triggering the modulation in the posterior regions. In analogy to the attention-based rehearsal mechanism operating during the retention period, repeated covert scanning of multiple locations or objects might be necessary for the formation of WM representations during the encoding period and thus, determine the interference between visual WM and attention demands in posterior parts of the cortex.

In addition, complex objects, as used in the present task and as we usually encounter them in our everyday experience, consist of multiple parts, each with its own features. Different features are bound together into integrated objects by means of focused attention (Treisman and Gelade, 1980). The posterior parietal cortex has been shown to be involved in visual feature integration (Shafritz et al., 2002). In the context of WM, it can be proposed that distinct regions in the visual cortex serve as simple parallel feature stores. These stores are modulated by attentional mechanisms (Awh and Jonides, 2001; Jha 2002; Postle et al., 2004; Lepsien and Nobre, 2007) that integrate the distributed information into unified object representations (Wheeler and Treisman, 2002). These attentional modulations seem to be subserved by parietal and premotor regions (Kanwisher and Wojciulik, 2000; Corbetta and Shulman, 2002; Pessoa et al., 2003; Lepsien and Nobre, 2007). Therefore, the need for integration of information might be another possibility that determines the interference between visual WM and attention demands in posterior parts of the cortex.

However, it can not be unambiguously excluded that other factors than the demands on attention-based encoding or the binding of featural information contributed to the observed effect of interference in posterior regions. One factor concerns the processing of spatial information. In experiment 2, spatial processing was obviously necessary for the encoding of locations into WM, but may have been necessary for controlling the operation of attention in the difficult search condition as well. Thus, common demands on spatial rather than attentional processing may have caused the interference between the two task components in the posterior cortex. Spatial localisation occurs within the dorsal visual pathway, extending from the primary visual cortex to the posterior parietal cortex (Mishkin et al., 1983; Husain and Nachev, 2007). Also, bilateral parietal lesions lead to both spatial impairments and impairments in attention-demanding visual search tasks (Robertson et al., 1997) and it is difficult to determine whether the attentional deficits are a result of more general deficits in spatial processing. However, the posterior parietal cortex is a large region comprising several subregions that are likely to mediate a variety of functions, including spatial and even nonspatial perceptional, attentional, mnemonic, and action processes (Husain and Nachev, 2007). Therefore, it is possible that in the present tasks common demands on spatial and attentional processing have led to interference in distinct regions of the parietal cortex. Future research is needed to disentangle the various contributions of parietal regions to the common processes underlying visual search and visual WM.

Spatial processing might have been involved also in the object WM task. Complex objects as used in experiment 1 are composed of several edges and angles, and thus may contain both nonspatial and spatial information. Previously, it has been shown that the neural mechanisms for the maintenance of objects such as houses or three-dimensional paper-clips overlap with those that maintain spatial location information in regions of occipital, parietal, and prefrontal cortex (Pollmann and von Cramon, 2000; Sala et al., 2003). Moreover, subjects needed to encode not only the objects but also their specific orientation. In addition, as suggested by the findings of the behavioural study, participants might have used a strategy that involved memorising the positions of the targets before encoding their shapes in the NPO condition. Thus, there were several sources of spatial processing even in the object encoding task which might have interfered with the demands on spatial processing in the visual search task leading to interference in neural activation in the posterior cortex.

6.2 Object-selective and spatial-selective activation

Despite the large consistency in posterior cortex activation associated with common limitations for attention and WM encoding, experiments 1 and 2 revealed differences in activation selectively responsive to the demands on object vs. spatial WM encoding. These differences partly supported the domain-specific organisation of WM in posterior PFC (Goldman-Rakic, 1987). According to this account, a dorsal-ventral and a right-left gradient exists for the storage of spatial and object information (Courtney et al., 1996, 1998; D'Esposito et al., 1998; Munk et al., 2002; Manoach et al., 2004; Mohr et al., 2006). Consistently, prefrontal activation selectively responsive to object WM load was found in mid-ventrolateral PFC, in a region around the IFS adjacent to the precentral sulcus and this activation was lateralised to the left hemisphere. This region included BA 45 that is part of Broca’s area that is involved in language processing and speech production (Amunts et al., 1999). As most subjects reported using a verbal-associative encoding strategy in experiment 1 it cannot be ruled out that it was this strategy that was associated with left ventrolateral PFC activation in BA 45. Consistent with the domain-specific account spatial WM load-selective activation was more pronounced in the right mid-dorsolateral PFC including dorsal and anterior parts of the MFG (BA 9, BA 10). However, in other regions of the mid-dorsolateral PFC (bilateral BA 46, left BA 10) WM load-selective activation appeared for both types of stimuli. This finding might fit within the process-specific (rather than the domain-specific) account of PFC organisation whereas a dorsal-ventral dissociation exists according to the type of process (mid-dorsolateral PFC for manipulation or monitoring of information and mid-ventrolateral PFC for the maintenance of information) (Owen et al., 1999; Petrides, 2000). However, the present experiments were not designed to explicitly test these models of WM organisation. Activation related to object vs. spatial WM encoding was compared across different experiments with different subjects and thus, it can not be excluded that differences in overall difficulty of the experiments, the scanning parameters and conditions, and the subjects, contributed to differences in the observed activation patterns. Importantly, these models are not necessarily mutually exclusive and the present results might fit best with recent models integrating both accounts (see next paragraphs).

6.3 Dissociation between PFC and posterior cortex

One key finding of this dissertation was the finding that the PFC was not part of the activation pattern that reflected the common processing limitations of visual WM and attention. Prefrontal activation has been linked to a variety of control processes needed to guide performance in complex tasks based on current goals (Miller and Cohen, 2001). In the context of WM, such processes include selection (Rowe et al., 2000), monitoring and transformation of information held in WM (Owen et al., 1999; Petrides, 2000; Bor et al., 2003; Champod and Petrides, 2007), mediation of interference (Postle, 2005), or the representation of task sets (Sakai and Passingham, 2003). In addition, PFC responds to WM load beyond the capacity of the parietal-premotor network (Linden et al., 2003). Therefore, WM load-selective activation observed in the PFC might fit within the framework postulating that this brain region subserves extra-mnemonic processes of top-down control over posterior regions where information is actually stored (Curtis and D'Esposito, 2003; Passingham and Sakai, 2004; Postle, 2006). The interplay between PFC and posterior regions was not in the direct focus of the present study. Nevertheless, the present data indicate that successful encoding into WM requires joint processing across encoding-selective prefrontal regions and regions that are also called upon by demands on selective attention in the posterior cortex.

6.4 A tentative model of the neural substrate of WM

The model proposed by Curtis and D’Esposito (2003) suggests that the role of mid-dorsal PFC (BAs 46, 46/9) is domain-independent sending biasing signals to specific frontal regions that are involved in the rehearsal of specific types of information. Whereas spatial rehearsal is mediated by premotor cortex (BAs 6, 8A) probably in terms of reactivation of oculomotor programs without making overt eye movements, Broca’s area (BA 44/ 45) is involved in verbal rehearsal through sub-vocalisations of to-be-remembered items. So far, the model does not specify the location and mechanism underlying the rehearsal of object information. From the present findings a role of the left IFG can be hypothesised. Most importantly, top-down signals from the mid-dorsal PFC are supposed to bias activation not only in the frontal cortex but also in the posterior cortex where WM representations are actually stored. Evidence suggests that the posterior cortex is functionally segregated by the preferred type of material it supports in WM. In the present study, domain-specific activation was found in the left temporal cortex with activation extending slightly more anterior in the object compared to the spatial WM task. This is consistent with previous studies showing persistent activity in inferior temporal areas during maintenance of visual objects (e.g., Druzgal and D'Esposito, 2003; Postle et al., 2003; Ranganath et al., 2004). The parietal cortex was involved in the encoding of locations and objects and overlapped with activation associated with difficult visual search. Thus, the functional organisation of parietal cortex seems to be complex with contributions to spatial, non-spatial representational and attentional mechanisms (Figure 6.1).

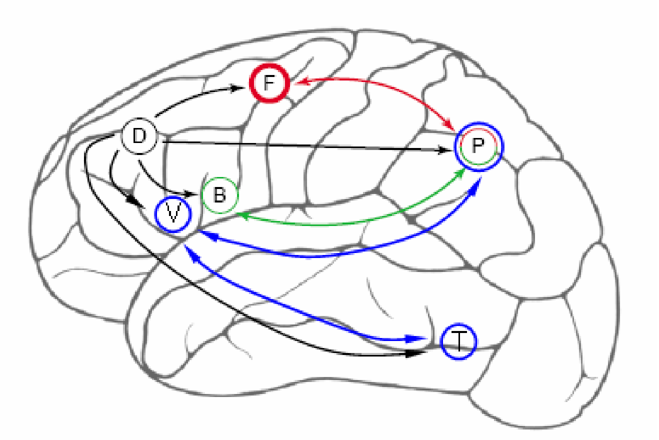


Figure 6.1 Simplified model of the cortical substrate for spatial (red), verbal (green) and object WM (blue) (modified from Curtis and D'Esposito, 2003). B: Broca's area; D: dorsal PFC; F: frontal eye-field; P: parietal cortex; T: temporal cortex; V: ventral PFC.

6.5 Capacity constraints for visual WM

Previous studies have localised the capacity limit of visual WM maintenance in the posterior occipito-parietal and premotor cortex (Linden et al., 2003; Todd and Marois, 2004; Xu and Chun, 2006) and it has been proposed that these posterior regions subserve an attention mechanism that selects and determines the maximum number of items that can be maintained (Linden et al., 2003; Xu and Chun, 2006). In these studies subjects were asked to maintain complex objects, or both the colour identity and location of objects in WM. Manipulating the demands on encoding information either into object or spatial WM and attention independently within one task, the present experiments revealed common processing limitations of attention and WM encoding in highly similar posterior regions (see Figure 6.2) and thus, provide evidence for the implicated role of attention as a cause of visual WM capacity constraints.

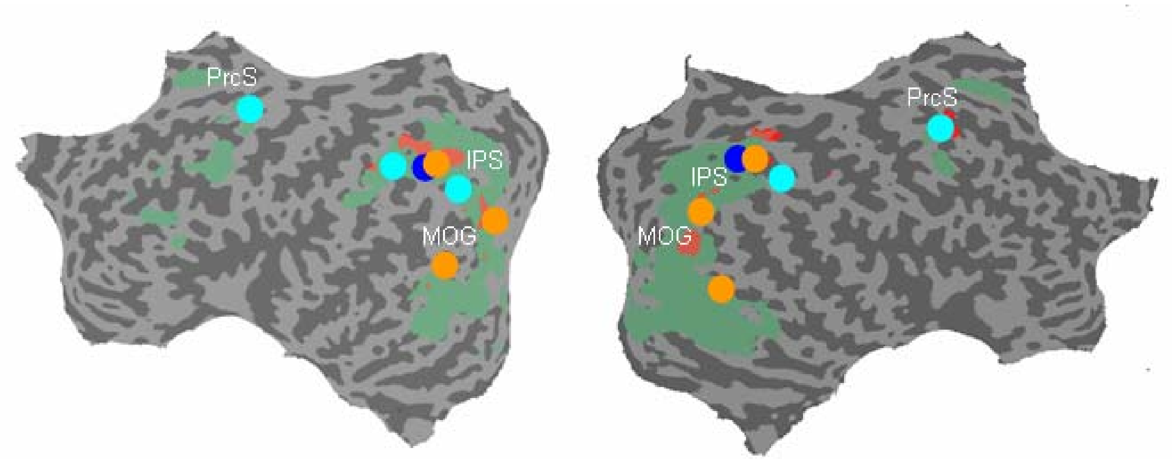


Figure 6.2 Present findings compared with previous findings on the capacity constraints for visual WM. Statistical maps of the significant 2-way interaction of search difficulty x WM load (green: object encoding, red: spatial encoding) are shown. The circles indicate regions associated with the capacity limitations as revealed in previous studies (bright blue: Linden et al., 2003; orange: Xu and Chun, 2006; dark blue: Todd and Marois, 2004.) IPS: inferior parietal sulcus, MOG: middle occipital gyrus, PrcS: precentral sulcus.

However, the finding of common capacity-limited neural mechanisms shared between object and spatial WM encoding and attention does not necessarily imply that the capacity limit of visual WM is fully reducible to that of attention (Fougnie and Marois, 2006). The capacity limit of visual WM may result from an interaction between capacity-limited attentional processes and the independent capacity of distinct features stores (Wheeler and Treisman, 2002). This would predict that distinct content-specific regions in the visual cortex may also contribute to visual WM capacity. Another capacity limiting factor that has been largely discussed is the storage format of visual WM, i.e., the question whether visual information is stored in terms of the number of integrated objects, or in terms of the object's features (Luck and Vogel, 1997; Alvarez and Cavanagh, 2004). Evidence suggests that the representation capacities of the parietal and occipital cortices differ with regard to the complexity of the objects (Xu and Chun, 2006). Whereas representations in the inferior IPS are fixed to about four objects regardless of object complexity, those in the superior IPS and lateral occipital complex are variable, tracking the number of objects held in visual WM, and representing fewer than four as their complexity increases. Furthermore, cognitive processes mediated by the PFC likely have their own capacity limitations as well, which may specifically constrain the maintenance and manipulation of information stored in visual WM (Callicott et al., 1999; Leung et al., 2007). For instance, in an n-back task that required subjects to continually encode, update, and discard the information held in WM with the presentation of a new stimulus, the amount of activation in the DLPFC increased with n up to two items, and decreased at highest WM load coincident with a significant decrement in accuracy (Callicott et al., 1999). Finally, processes specifically associated with the retrieval of information from WM might be subject to their own specific capacity limitations (Wesenick, 2003).

In conclusion, the capacity of visual WM can be limited at various stages of processing. The present dissertation illustrated that one major bottleneck of information processing arises from the common demands on neural resources shared between visual WM and selective attention during the encoding stage. A challenge for future research is to disentangle the roles of attentional, process-specific, and central limitations and to describe their complex interactions that lead to the constrained mental representation of the visual world.

6.6 Combining behavioural and fMRI results

Behavioural measures have long formed the building blocks of cognitive theory. Examining the effects of task manipulations on RT and response accuracy, the component operations constituting specified cognitive functions can be isolated and cognitive theories can be tested. Behavioural measures may comprise the most accessible objective indices of human cognition. Therefore, to be able to assign functional relevance to an observed pattern of brain activation, a necessary step in any cognitive brain imaging experiment is to ensure that it meets the behavioural criteria that confirm or constrain the range of cognitive processes that have been engaged (“behavioural-dependency criterion”, Wilkinson and Halligan, 2004). In the fMRI experiments of this dissertation it was necessary to present the search array for a fixed amount of time in order to rule out differences in brain activation owed to differences in sensory stimulation. Behavioural performance could be measured only when presenting the probe and thus, response accuracy and RTs captured only the final outcome of the task-related processes. Due to this methodological constraint it was not possible to diretely validate encoding-related effects in brain activation by corroborative behavioural effects. To meet the “behavioural-dependency criterion” (Wilkinson and Halligan, 2004), the fMRI experiments were preceded by a behavioural study in which the participant-chosen presentation time of the stimulus array was measured as a function of attentional demand and WM load, thus providing a direct index of the cognitive processes required for successful WM encoding. The demonstration of effects of search difficulty and WM load on the individual presentation times confirmed that the paradigm was suitable to induce task-dependent demands on WM encoding and attentional processing during the encoding phase. In addition, measuring the individual presentation times within the behavioural study it was possible to isolate the sub-processes that allowed participants to cope with the common processing limitations of visual attention and WM encoding. This question could not be addressed in the fMRI study due to insufficient temporal resolution. Importantly, the difference in the encoding times between NPO and PO conditions observed in the behavioural study was consistent with the time shift in WM load-related PFC activation between the NPO and PO conditions. This correspondence between the behavioural and fMRI measurements provided important implications for the fMRI data interpretation in terms of a mechanism for coping with the common limited-capacity resources for attention and WM encoding in posterior regions. Finally, the behavioural results were used to determine the duration of the encoding period needed for successful WM encoding in the fMRI experiments. Taken together, the present dissertation provides different sources of evidence that illustrate the relevance of combining neuroimaging and behavioural data.

Indeed, there is not doubt about the value of corroborative behavioural effects for the interpretation of fMRI data since functional imaging studies of cognitive function investigate the organisation and functional structure of cognitive constructs in the human brain. By choosing tasks appropriately to isolate cognitive operations based on pre-existing psychological theories, fMRI is most often used to map known cognitive processes onto particular brain regions. This has been done also in the present dissertation with the main result of common neural resources shared between visual WM encoding and attention that could be localised in the posterior but not the prefrontal cortex. This finding informs neurophysiological models of WM suggesting that WM encoding is implemented by the interplay between PFC and posterior regions that are also called upon by demands on selective attention.

However, what is the contribution of functional imaging to the understanding of the nature of cognition? Can neuroimaging data constrain cognitive models? This question has gained much controversy. Assuming that behavioural and neuroimaging studies address completely different levels of analysis some authors deny any usefulness of neuroimaging data in testing psychological models of cognitive functions (Coltheart, 2006). This view relates to Marr’s (1982) distinction between the algorithm describing a cognitive process and its physical implementation. Because the algorithm could be physically implemented in several ways, it is argued that imaging results can never confirm or disprove the operation of a given cognitive process (Wilkinson and Halligan, 2004). On the other side, based on the assumption that there is some systematic mapping from psychological function to brain structure, increasing evidence suggests that imaging data about localisation can provide supporting information for existing cognitive theories, that it can be used to generate novel hypotheses about cognitive architecture, and help to distinguish between competing theories (e.g., Henson, 2005; Jonides et al., 2006; Jack et al., 2006). Some authors even postulate that imaging data can be treated as an additional dependent variable that provides equally valid contributions as behavioural measures (Henson, 2005; Jonides et al., 2006). How does the main finding of this dissertation, i.e. that selective attention and the encoding into visual WM share common neural resources in distributed posterior brain regions but not the PFC where activation was selectively sensitive to WM load, inform cognitive models of WM? One conclusion that arises from this finding is that the encoding of information into WM is represented throughout the brain and relies critically on the very same neural and cognitive resources that support the processing of perceptual information (Slotnick, 2004; Jonides et al., 2005; Pasternak and Greenlee, 2005; Postle, 2006; Ranganath, 2006). The implication is that WM cannot be viewed as an unique or independent buffer or storage site (Baddeley and Hitch, 1974) dedicated to the representation of behaviourally relevant sensory information when it is no longer present in the external environement. Rather, the findings favour cognitive models that describe the contents of WM in terms of activated representations that are in the focus of attention (Cowan, 1988, 1999). In conclusion, the fMRI findings of this dissertation support a functional rather than a structural concept whereas WM evolves from the recruitement of attentional mechanisms the very same that act upon perceptual representations as well.

Chapter 7 - Directions for future studies

The present dissertation revealed common capacity limitations for visual WM encoding and attention in the posterior cortex indicating an attention-based model of visual WM encoding. In my presentation of the proposed model I made several claims that were only partially supported by experimental evidence or completely untested. This suggests several directions for future research. In this section, I will discuss some assumptions concerning the cognitive and neural constraints of visual WM with regard to attentional (1-2) and non-attentional mechanisms (3-4) and how they could be tested in future studies. The section closes with a proposal on the investigation of WM dysfunctions in patients with schizophrenia (5).

7.1 Modality-specific vs. modality-independent interactions between WM and attention in the posterior cortex

One direction for future studies will be to test whether interactions between attention and WM are modality-specific or generalise across WM domains. The present dissertation indicated similar effects of interference between attention and WM encoding of objects and locations in the posterior cortex. However, this effect might have been confounded by demands on spatial processing required even in the object task. Future studies combining visual search with WM encoding could use less complex stimuli with no spatial requirements such as colour, texture or brightness. If attention-demanding visual search reduced activation associated with the encoding of colours in posterior regions, this would provide evidence for i) common demands on attention (rather than spatial processing) in the posterior cortex and ii) domain-independent interactions between attention and WM. Additional support could be derived from experiments combining visual search with verbal WM encoding.

Moreover, recent evidence suggests separable neural substrates in the parietal cortex for object-based and spaced-based attentional mechanisms (Shomstein and Behrmann, 2006). Therefore, it seems possible that the neural substrate mediating interference between attention and WM could differ not only with regard to the WM modality but also with regard to the attentional modality implemented in the task. Manipulating the demands on object-based vs. spaced-based attention and the demands on object vs. spatial WM encoding within one task would be best to dissociate precisely the cognitive and neural resources that are shared between the mechanisms of selective attention and WM. I predict common processing limitations for object-based attention and object WM and for space-based attention and spatial WM in distinct posterior parietal regions.

7.2 The role of attention for the storage of integrated information in WM

A second direction will be to test the assumption that focussing attention is required to maintain bindings in visual WM. Evidence is provided by the demonstration that the retrieval of integrated information is selectively vulnerable to interference. Performance in a change detection task has been shown to be worse for changes of bindings than for features when a whole display was presented at test compared to a condition with a single probe item (Wheeler and Treisman, 2002). Comparable evidence concerning the processes involved during the maintenance phase still lacks. This issue could be addressed in a task that requires subjects to maintain either features (e.g., colour or shape) or conjunctions of features (e.g., colour and shape) in WM while performing a secondary attention demanding task during the retention period (e.g, a visual search task). If attention is required to maintain the links between colour and shape information, visual search and WM for bindings should use common attentional resources whereas visual search and WM for object features should not. Thus, three hypotheses can be made. First, in the conjunction condition, the efficiency of the search should be worse when the search task is presented during the retention period of the WM task as compared with when the search task is tested in isolation. Second, memory for bindings should be impaired in the combined task as compared with when the memory task is performed alone. In addition, the decrement in memory performance might increase as the number of items to be searched increases. Third, in the feature condition, search and memory performance should not differ whether the search and the memory task are performed together or separately.

7.3 Capacity limits of WM control in PFC

A third direction will be to investigate the capacity limitations of WM control processes mediated by PFC. Experimental manipulations that affect the demands on the coordination, transformation, or integration of information within WM should selectively influence processing capabilities in PFC and not posterior regions shared with attentional mechanisms also involved in visual perception. Moreover, a parametric manipulation of the demands on these processes may allow identifying the neural substrate associated with the capacity constraints of WM control. It can be hypothesised that neural activation increases with increasing demands on WM control reaching a plateau when the capacity is full.

Related to the question of capacity constraints of WM control in the PFC is the assumption that this region subserves extra-mnemonic processes of top-down control over posterior regions where information is actually stored (Curtis and D'Esposito, 2003; Passingham and Sakai, 2004; Postle, 2006). The interplay between PFC and posterior regions was not in the direct focus of the present dissertation and thus, the conclusion of PFC top-down signals needed to remain speculative. One possibility to address this question is to investigate the interactions between frontal and posterior brain regions by assessing the coupling between these distributed brain areas (i.e., functional connectivity) and determining the causal directionality of these interactions (i.e., effective connectivity). Ultimately, a multi-methodological approach combining fMRI with transcranial magnetic stimulation, which allows inducing temporary virtual lesions in a controlled and systematic manner, would be key for testing this model and providing direct causal evidence for PFC top-down signals in the control of WM.

7.4 The role of stimulus similarity in limiting visual WM capacity

A fourth direction will be to test the role of item similarity in the limitation of visual WM and to determine their neural correlates. High similarity between the items might lead to a reduction in the capacity to maintain and retrieve items from WM due to mutual interference. For instance, Hitch and colleagues (1988) reported poorer recognition memory for items from a set of visually similar (e.g., brush, rake, pen) vs. distinct items (e.g., pig, ball, pen). To investigate whether stimulus similarity constraints the capacity of visual WM in the frontal and posterior cortex, one could use a delayed visual discrimination task with parametric variation of WM load and independently manipulate the visual similarity of the items to be maintained. For instance, in the spatial domain a spatial rotation task could be implemented that requires subjects to memorise the rotation angle of 1 to 6 semicircles presented simultaneously. The rotation angle could be varied on a continuum in steps of 5° and visual similarity could be manipulated by presenting in a randomised manner either rotation angles that are situated adjacent (high similarity) or remote on that continuum (low similarity). An event-related fMRI design would allow dissociating activation related to the encoding, maintenance, and retrieval of spatial information as a function of visual similarity. Specifically, the neural substrates mediating similarity-based WM storage capacity should demonstrate an increase in activation with increasing WM load reaching a plateau in activation with lower WM load as similarity increases. Importantly, behavioural performance at different WM loads could be used to estimate the individual differences in WM capacity and correlate it with the fMRI data.

7.5 The neural basis of cognitive dysfunction

A fifth direction will be to investigate the neural basis of cognitive dysfunction that can be found in patients suffering from brain damage but also occurs in various psychiatric disorders such as dementia, schizophrenia, and depression. Functional imaging is not only central to increase the understanding of these diseases, it may also provide neurobiological diagnostic markers, and might be useful in assessing the efficiency of medication and other treatments.

Specifically, WM and attentional dysfunctions are core components of schizophrenia. Cognitive deficits have been linked to psychotic phenomena and contribute directly to poor social functioning in patients with this illness. They often develop before the first clinical symptoms, and affect first-degree relatives of patients. Thus, cognitive dysfunctions seem to be an inherent biological phenomenon linking schizophrenia to abnormal brain function. WM dysfunctions in schizophrenia are reflected in reduced prefrontal activity especially with high WM demands, and fronto-parietal and/or fronto-temporal disconnections (Kuperberg and Heckers, 2000; Mitchell et al., 2001; Honey, 2006). It has been proposed that impairments in WM performance are caused by deficits in selective attention for encoding and maintenance (Gold et al., 2003). The present paradigm combining the demands on both mechanisms could be used to directly test whether impairments in the limitations of WM are due to impaired processing shared by visual attention and WM encoding in the posterior cortex or whether they are caused by WM-selective dysfunctions in PFC.

Taken together, I propose future projects that are aimed at characterising the cognitive processes mediated by the neural substrates involved in visual WM. Systematically manipulating the characteristics of the to-be-encoded stimuli, such as stimulus modality and format, allows further disentangling the cognitive and neural resources that are shared between the attention and WM systems. The contribution of additional capacity-constraining factors such as cognitive control and item similarity can be empirically tested. Combining fMRI data with behavioural data, data derived from virtual lesion studies with TMS, and from patients with cognitive deficits, will provide a powerful tool for constraining the interpretation of fMRI data and testing cognitive models of WM.

Summary

Visual WM and selective attention are fundamental cognitive mechanisms, both operating at the interface between perception and action. They are related because both are concerned with the control of information, and both are postulated to have limits with respect to how much information can be processed (Miller, 1956; Pylyshyn and Storm, 1988; Cowan, 2001). However, visual WM and attention have been largely studied in isolation and interactions between the two have rarely been addressed. This dissertation aimed at investigating interactions between selective attention and the encoding of information into visual WM in the context of one common characteristic feature, namely their limitation in capacity. Specifically, fMRI was used to test the hypothesis that the capacity limitation of visual WM is due to limited-capacity cognitive and neural resources shared with the process of visual selective attention. An experimental task was used that combined visual search with delayed discrimination and the demands on selective attention and WM encoding were manipulated orthogonally. With this task it was possible to identify the brain regions that were selectively responsive to either attentional demand or the encoding into visual WM and those involved in both processes. Most interestingly, the independent manipulation of visual search difficulty and WM load allowed studying interactions between the underlying processes in terms of neural activation.

Traditional models of human information processing considered temporary memory and attention distinct, associated with separate functions. Attention and WM were thought to operate at different stages of processing, with attention taking place earlier and controlling which sensory information gets encoded into WM (Broadbent, 1958; Atkinson and Shiffrin, 1968). In this case, visual WM and attention might be represented by different neural substrates. However, recent models of WM suggest that selective attention and WM may rely on a common capacity-limited cognitive mechanism (Cowan, 1988; Baddeley, 1993). For instance, Cowan offers the view that WM is best understood as a subset of activated representations of long-term memory that is currently within the focus of attention. Selective attention has been implicated as a limiting factor for the

storage capacity of visual WM (Cowan, 1998, 2001; Wheeler and Treisman, 2002). This view predicts that visual WM and attention share common neural resources.

Frontal and parietal brain regions are the primary areas involved both in WM and visual attention (Pessoa and Ungerleider, 2004). Overlap of the cerebral networks of WM and attention has been demonstrated in targeted comparisons (LaBar et al., 1999; Pollmann and von Cramon, 2000; Corbetta et al., 2002). However, neuroanatomic overlap does not necessarily entail a functional relationship between the two cognitive domains. For example, one cannot exclude that shifts of visuospatial attention associated with activation of a given brain region are epiphenomenal to the core processes that encode and maintain information in visual WM (Awh et al., 2006). By demonstrating that memory performance declines when shifts of attention are prevented, it becomes possible to infer a true functional role of attention in visual WM (Smyth and Scholey, 1994; Awh et al., 1998; Oh and Kim, 2004; Woodman and Luck, 2004).

To determine whether visual attention and WM are represented by different or common neural substrates, the demands on visual attention and the process of encoding of information into WM were manipulated orthogonally within one unitary task. It was hypothesised that if visual WM and attention shared common limited-capacity neural resources, these resources would become exhausted in conditions that make high demand on both processes, thus resulting in interference. Such interference would indicate a limitation of the neural resources available for WM encoding and attentional processing. Two fMRI experiments were conducted that required subjects to encode either objects or locations into WM. Thus, the question whether effects of interference between visual attention and WM encoding are domain-specific or generalise across different classes of stimuli could be addressed.

The fMRI study was preceded by a behavioural study that served two purposes. First, the behavioural experiments were used to validate the engagement of the relevant attentional and WM processes by the chosen task manipulations. Second, the behavioural experiments addressed the question of whether and how subjects can cope with the common capacity limitations of visual WM and attention. The behavioural study comprised five experiments in which the study phase consisted always of identical stimuli, the tasks differing only in the instructions and in the test displays. In the main experiment participants were presented with a search array of nine grey objects and had to memorise only some of them (targets), while the others could be ignored (distractors). The discrimination between targets and distractors was based on an L-shaped item located in the centre of the object, but only the outer shape of the object and its orientation had to be remembered. After the display disappeared participants fixated a cross during a delay period of 8 s, which was followed by the presentation of a single test shape. Participants were then required to indicate whether the test shape matched in the form and orientation one of the target shapes. Attentional demand was manipulated by implementing two search conditions in which target items had either unique features and were highly discriminable from the distractors (pop-out = low attentional demand) or shared the features with the distractors and were thus difficult to discriminate (non pop-out = high attentional demand). Only in the latter case it was expected that the detection of targets would require attention-demanding serial search (Treisman and Gormican, 1988; Duncan and Humphreys, 1989). WM load was determined by the number of targets and varied from one to five. The individual presentation time of the search array needed for successful WM was determined by the participants via button-press and used as main dependent variable. This variable provided a direct index of the cognitive processes required for successful WM encoding and allowed isolating the processes that enabled participants to cope with the demands on attentional and WM processing.

The behavioural results revealed that presentation times increased with increased WM load and were considerably higher under high than under low attentional demand. Thus, the paradigm was suitable to induce task-dependent demands on WM encoding and attentional processing. WM performance at test was high and comparable across the different search conditions. Experiments 2 to 5 were aimed at investigating the processes underlying the costs on presentation time and thus the processes that enabled participants to successfully encode complex shapes into WM under concurrent demands on attention. Two possible strategies were tested. In a “search-and-encode strategy” participants would encode each shape as soon as they selected a relevant location, interleaving thus the search process with the WM encoding. In this case, presentation time should be simply divided between the two task components, and the presentation time that participants need in the non pop-out condition should be the sum of the presentation time in the pop-out condition and the time needed to select the relevant locations in the non pop-out condition. The other considered strategy was postulated to involve two separate steps of encoding (“two-step encoding strategy”). In the first step participants would select and memorise only the locations of all target items and only then would encode the associated shapes at a later step. It was expected that the additional process of memorising the target locations required additional processing time. For that case, I predicted a super-additive combination of the times for encoding and determination of target locations in the non pop-out condition. Experiment 2 and 3 tested the hypotheses of additivity vs. super-additivity of the times needed to encode and determinate the target locations. In Experiment 2, the time needed for simple visual search was measured. These times could not explain the increased presentation time produced by the lack of pop-out in Experiment 1. Therefore, Experiment 3 tested whether the slower processing in the non pop-out condition in Experiment 1 could be explained by repeated searches, owing to a putative lack of memory for visited target locations (Irwin, 1992; Peterson et al., 2001) and to the need to search the entire array. The need to search repeatedly was reduced by informing the participants at each trial about the upcoming number of targets. The time saved by this manipulation again could not explain the costs on presentation time produced by the lack of pop-out in Experiment 1. Therefore, the results from Experiments 1 to 3 indicated consistently super-additivity of the times for encoding and determination of the target locations, favouring the two-step encoding strategy. In Experiments 4 and 5 the two-step strategy was tested further. The times that participants needed to memorise the locations of the target items only were measured and it was investigated whether these times could explain quantitatively the difference between the pop-out and non pop-out conditions in Experiments 1 and 3. Indeed, in Experiments 4 and 5, the times needed to memorise the target locations accounted well for the presentation time offsets between pop-out and non pop-out conditions in Experiments 1 and 3, respectively. These results again favoured the two-step strategy that allowed participants to cope with the interference between WM and attention that would otherwise take place.

In the second part of this dissertation interference between visual attention and the encoding into visual WM was investigated on the level of neural activation using fMRI. The stimuli, procedure, and task design were the same as in the behavioural study with one important difference. To rule out differences in brain activation owed to differences in sensory stimulation the stimulus array was shown for a fixed amount of time. This time was determined based on the individual presentation times assessed in the behavioural experiments such that it was sufficient to allow successful encoding even in the most difficult condition.

In fMRI experiment 1 visual search was combined with delayed discrimination of complex objects. Participants were presented with a search array for 8 s and performed low or high attention-demanding visual search in order to encode one or three complex objects into visual WM. After an 8-s delay interval, a probe that consisted of a single object appeared and participants indicated whether the probe matched one of the memorised objects. The contrast analyses of fMRI data for the late encoding predictor revealed overlapping activation for attention-demanding visual search and object WM encoding in distributed posterior and frontal regions. In the right prefrontal cortex and bilateral insula BOLD activation additively increased with increased WM load and attentional demand. Conversely, the analysis revealed an interaction effect in several visual, parietal, and premotor areas. These regions showed overlapping activation for the two task components and were severely reduced in their WM load response under the condition with high attentional demand. Regions in the left prefrontal cortex were selectively responsive to WM load. Areas selectively responsive to high attentional demand were found within the right prefrontal and bilateral occipital cortex.

In fMRI experiment 2 visual search was combined with delayed discrimination of locations. Participants were presented with a search array for 5 s and performed low or high attention-demanding visual search in order to encode one, three, or five locations into spatial WM. After an 8-s delay interval, the original stimulus array was presented again with one of the shapes missing. Participants needed to indicate whether the location of the missing shape matched one of the target locations. The analysis focussed on the encoding phase. The results revealed overlapping activation for attention-demanding visual search and spatial WM encoding in distributed posterior and frontal regions. In the majority of these regions the overlap was associated with an additive increase in BOLD activation under high demands on attention and WM. Conversely, a subset of the overlap regions including the visual, parietal, and right premotor cortex, were severely reduced in their WM load response under the condition with high attentional demand as reflected in a significant interaction effect. Regions in the anterior prefrontal cortex were selectively responsive to increased WM load whereas posterior prefrontal regions and regions in the visual cortex in both hemispheres were selectively responsive to attentional demand.

The fMRI results provide convergent evidence that visual selective attention and the encoding of information into WM share, to a high degree, common neural resources but show also some degree of selectivity. Interference between attention and WM encoding-related activity appeared in distributed posterior regions. In contrast, regions in the PFC were selectively responsive to WM load and differed to some degree depending on the WM domain. Here, activation associated with increased WM load was delayed rather than reduced under high attentional demand which reflected the delay in encoding times estimated in the behavioural study. The findings indicate that competition for resources shared by visual attention and WM encoding can limit processing capabilities in distributed posterior brain regions and support the view that WM evolves from the recruitement of attentional mechanisms (Cowan, 2001; Wheeler und Treisman, 2002) the very same that act upon perceptual representations as well (Slotnick, 2004; Jonides et al., 2005; Pasternak and Greenlee, 2005; Postle, 2006; Ranganath, 2006). The similarity in the effects of interference between attention and the encoding of objects or locations into WM indicates that the attention-based model of WM encoding is valid across different WM domains.

The capacity of visual WM can be limited at various stages of processing. The behavioural and fMRI data presented in this dissertation illustrate that one major bottleneck of information processing arises from the common demands on neural and cognitive resources shared between visual WM and selective attention during the encoding stage.

Zusammenfassung in deutscher Sprache

Selektive visuelle Aufmerksamkeit und das visuelle Arbeitsgedächtnis stellen fundamentale kognitive Mechanismen an der Schnittstelle zwischen Wahrnehmen und Handeln dar.

Natürliche Szenen beinhalten in der Regel mehr Informationen, als das visuelle System zu einem Zeitpunkt parallel verarbeiten kann. Der Mechanismus der selektiven Aufmerksamkeit erlaubt es, irrelevante Aspekte einer Szene herauszufiltern und relevante Aspekte selektiv der Verarbeitung zugänglich zu machen. Die Rate, mit der visuelle Information aufmerksam fokussiert werden kann, sowie die Anzahl der Objekte, die gleichzeitig im Fokus der Aufmerksamkeit gehalten werden können (bis zu vier), sind begrenzt (Pylyshyn und Storm, 1988; Duncan et al., 1994).

Das psychologische Konstrukt des Arbeitsgedächtnisses beschreibt die kurzzeitige Speicherung und Manipulation von Informationen, die für höhere kognitive Funktionen wie Sprache, Problemlösen, Lernen oder Schlussfolgern benötigt werden (Baddeley, 1986). Nach dem Modell von Baddeley und Hitch (1974) umfasst das Arbeitsgedächtnis ein übergeordnetes System der Aufmerksamkeitskontrolle (zentrale Exekutive) und zwei Subsysteme, die der Speicherung und Manipulation von visuell-räumlicher Information (visuell-räumlicher Notizblock) und auditorischer bzw. sprachbasierter Information (phonologische Schleife) dienen. Ursprünglich wurde der visuell-räumliche Notizblock als einheitliches System dargestellt. Eine Vielzahl von Studien weist mittlerweile darauf hin, dass der visuell-räumliche Notizblock materialspezifisch in eine visuell-objektbasierte und eine visuell-räumliche Komponente unterteilt werden kann (Della Sala et al., 1999). Ein zentraler Aspekt des Arbeitsgedächtnisses stellt seine begrenzte Kapazität dar. Miller (1956) postulierte ursprünglich eine Kapazität von sieben plus minus zwei Chunks, definiert als Gedächtniseinheiten, die auf verschiedenen Gruppierungs- und Organisationsprozessen basierend, mehrere einzelne Informationen zusammenfassen. Neuere Untersuchen hingegen zeigen, dass die Speicherleistung des visuellen Arbeitsgedächtnisses auf ungefähr vier Chunks begrenzt zu sein scheint (Luck und Vogel, 1997; Cowan, 2001; Wheeler und Treisman, 2002).

Während in zahlreichen Experimenten Aufmerksamkeits- und Arbeitsgedächtnisprozesse separat untersucht wurden, ist die Beziehung zwischen diesen beiden kognitiven Systemen selten Gegenstand der Betrachtung gewesen und bisher nur unzureichend geklärt. Das Ziel der vorliegenden Arbeit war es, Interaktionen zwischen den Prozessen der selektiven Aufmerksamkeit und der Enkodierung von Information in das visuelle Arbeitsgedächtnis im Kontext des gemeinsamen Merkmals der begrenzten Kapazität zu untersuchen. Dazu wurde eine Aufgabe entwickelt, in der die Anforderungen an beide Mechanismen unabhängig von einander manipuliert wurden. Der erste Teil der Arbeit umfasst eine Verhaltensstudie, in der die kognitiven Prozesse, die es den Probanden erlaubten, die Aufgabe trotz begrenzter Verarbeitungsressourcen zu bewältigen, isoliert wurden. Im zweiten Teil der Arbeit wurde die kombinierte Aufgabe in zwei Studien mit der Methode der funktionellen Magnetresonanztomographie (fMRT) untersucht mit dem Ziel, die gemeinsamen und spezifischen neuronalen Korrelate von Aufmerksamkeits- und Arbeitsgedächtnisprozessen zu beschreiben und Interaktionen zwischen diesen Prozessen auf neuronaler Ebene zu charakterisieren.

Nach traditionellen Modellen der Informationsverarbeitung können Gedächtnis und Aufmerksamkeit als getrennte kognitive Systeme betrachtet werden, die mit unterschiedlichen Funktionen assoziiert sind (Broadbent, 1958; Atkinson und Shiffrin, 1968). Aufmerksamkeits- und Arbeitsgedächtnisprozesse operieren demnach an unterschiedlichen Stufen der Informationsverarbeitung. Dem Mechanisums der selektiven Aufmerksamkeit wurde dabei eine Torfunktion zugeschrieben, wonach dieser den Transfer der sensorischen Information in das Arbeitsgedächtnis kontrolliert. Diesen Modellannahmen zufolge sollten Aufmerksamkeits- und Arbeitsgedächtnisprozesse anhand qualitativ unterschiedlicher neuronaler Aktivitätsmuster repräsentiert sein. Dem gegenüber stehen neuere Arbeitsgedächtnismodelle, in denen vorschlagen wird, dass Aufmerksamkeit und Arbeitsgedächtnis auf einem gemeinsamen kognitiven Mechanismus beruhen. So definiert z.B. Cowan (1988) das Arbeitsgedächtnis als

Repräsentationen des Langzeitgedächtnisses, welche momentan im Fokus der Aufmerksamkeit stehen. Von Bedeutung ist insbesondere die Annahme, dass Prozesse der selektiven Aufmerksamkeit einen limitierenden Faktor der Kapazität des Arbeitsgedächtnisses darstellen (Cowan, 2001; Wheeler und Treisman, 2002). Nach diesen Modellen ist anzunehmen, dass Aufmerksamkeits- und Arbeitsgedächtnisprozessen ein gemeinsames neuronales Korrelat zugrunde liegt.

Separate Untersuchungen mit bildgebenden Verfahren weisen darauf hin, dass Arbeitsgedächtnis- und Aufmerksamkeitsprozesse mit neuronaler Aktivität insbesondere in frontalen und parietalen Regionen korrelieren (Pessoa und Ungerleider, 2004). Überlappende Aktivitätsmuster konnten auch in vergleichenden Studien aufgezeigt werden (LaBar et al., 1999; Pollmann und von Cramon, 2000; Corbetta et al., 2002). Der Befund überlappender neuronaler Korrelate für Aufmerksamkeits- und Arbeitsgedächtnisprozessen spricht jedoch nicht notwendigerweise für eine funktionale Beziehung zwischen diesen kognitiven Systemen. Dass Aufmerksamkeitsprozesse funktional bedeutsam sind für das Halten von Informationen im Arbeitsgedächtnis konnte in Verhaltensstudien nachgewiesen werden, in denen Arbeitsgedächtnis- und Aufmerksamkeitsaufgaben kombiniert wurden (Smyth und Scholey, 1994; Awh et al., 1998; Oh und Kim, 2004; Woodman und Luck, 2004). Die räumliche Arbeitsgedächtnisleistung war z.B. signifikant geringer, wenn die Probanden während der Haltephase ihre Aufmerksamkeit auf eine andere Diskriminationsaufgabe richten mussten (Awh et al., 1998).

Um die Frage beantworten zu können, ob die Mechanismen der visuellen selektiven Aufmerksamkeit und des visuellen Arbeitsgedächtnisses auf gemeinsamen oder getrennten neuronalen Ressourcen beruhen, wurde daher ein Paradigma entwickelt, in dem die Anforderungen an beide Prozesse unabhängig von einander manipuliert wurden. Falls beide Prozesse gemeinsame begrenzte neuronale Ressourcen beanspruchen, sollten diese Ressourcen erschöpft sein, wenn die Anforderungen an beide Prozesse hoch sind. Auf Ebene der BOLD-Antwort sollte dies in einem Interaktionseffekt zwischen den beiden Aufgabenmanipulationen sichtbar sein. Dies wäre dann als Hinweis auf eine gemeinsame Limitierung der neuronalen Ressourcen zu werten, die für

Arbeitsgedächtnis- und Aufmerksamkeitsprozesse zur Verfügung stehen. Es wurden zwei fMRT-Experimente durchgeführt, in denen jeweils eine visuelle Suchaufgabe mit einer Aufgabe zur Enkodierung von Information in das visuelle Arbeitsgedächtnis kombiniert wurde (Experiment 1: Objekte; Experiment 2: Positionen). Die Verwendung unterschiedlicher Materialarten erlaubte die Untersuchung domänenspezifischer vs. domänenübergreifender Interferenzeffekte zwischen Aufmerksamkeits- und Arbeitsgedächtnisprozessen.

In einer Verhaltensstudie wurde zunächst überprüft, inwieweit die Manipulation der Aufgabenbedingungen dazu geeignet war, unterschiedliche Gedächtnis- und Aufmerksamkeitsanforderungen zu operationalisieren. Darüber hinaus konnten die kognitiven Prozesse isoliert werden, die es den Probanden erlaubten, die Aufgabe trotz begrenzter Verarbeitungsressourcen zu bewältigen. Es wurden insgesamt fünf Experimente durchgeführt mit jeweils identischem Stimulusmaterial. Die Aufgaben unterschieden sich bezüglich der Instruktionen und der präsentierten Testreize. Im Hauptexperiment wurde den Probanden ein visuelles Feld, bestehend aus neun Figuren präsentiert. Jede Figur war in der Mitte mit einem farbigen, L-förmigen Winkel markiert. Die Probanden waren instruiert, nach zuvor definierten Zielwinkeln zu suchen und sich die so markierten Figuren zu merken. Nach einer Haltephase von 8 sec. wurde den Probanden ein Testreiz dargeboten und sie gaben per Tastendruck an, ob die Figur des Testreizes mit einer der zuvor gespeicherten Figuren übereinstimmte. Die Aufmerksamkeitsanforderung wurde über den Schwierigkeitsgrad der visuellen Suche manipuliert. In der leichten Suchbedingung unterschieden sich die Ziel- und Distraktorwinkel in der Farbe und waren somit leicht zu diskriminieren. In der schwierigen Suchbedingung wurden Ziel- und Distraktorwinkel in den gleichen Farben dargeboten und konnten nur anhand der Orientierung diskriminiert werden. Die Detektion der Zielreize erforderte somit einen aufmerksamkeitsbeanspruchenden Suchprozess (Treisman und Gormican, 1988; Duncan und Humphreys, 1989). Die Arbeitsgedächtnislast wurde über die Anzahl der Zielwinkel, welche die Anzahl der zu enkodierenden Figuren bestimmte, festgelegt. Diese variierte zwischen 1 und 5. Als wichtigste abhängige Variable wurde die Zeit, die die Probanden für die Suche und erfolgreiche Enkodierung der Figuren in das Arbeitsgedächtnis benötigten, erhoben. Die individuelle Präsentationszeit des Suchfelds gaben die Probanden per Tastendruck an. Diese Variable stellte ein direktes Maß für die enkodierungsbezogenen Prozesse dar und erlaubte es, die kognitiven Prozesse zu isolieren, die es den Probanden ermöglichten, komplexe Figuren bei gleichzeitiger Bearbeitung einer aufmerksamkeitsbeanspruchenden visuellen Suche in das Arbeitsgedächtnis zu enkodieren.

Die Resultate zeigten, dass die individuellen Präsentationszeiten mit zunehmender Gedächtnislast anstiegen und in der schwierigen Suchbedingung deutlich länger waren im Vergleich zur einfachen Suchbedingung. Diese Ergebnisse bestätigten eine geeignete Operationalisierung von Enkodierungs- und Aufmerksamkeitsprozessen anhand der gewählten Aufgabenmanipulationen. Wie aber waren diese Kosten genau zu erklären? Zur Beantwortung dieser Frage wurden zwei Strategien getestet. Eine mögliche Strategie beinhaltete das abwechselnde Suchen der Zielwinkel und Enkodieren der Objekte, d.h. es wurde angenommen, dass die Probanden nach der Detektion eines Zielwinkels die entsprechende Figur direkt enkodierten. In diesem Fall sollte sich die Präsentationszeit in der schwierigen Suchbedingung aus der Summe der Zeit, die für die Enkodierung der Figuren benötigt wurde (Präsentationszeit in der leichten Suchbedigung) und der Zeit, die für die Detektion der Zielwinkel benötigt wurde, ergeben. Die andere Strategie umfasste zwei separate Stufen des Enkodierens. Es wurde angenommen, dass die Probanden zunächst alle Zielwinkel detektierten und deren Positionen memorierten und sich erst in einem zweiten Schritt die relevanten Figuren merkten. Der Prozess der Enkodierung der relevanten Positionen sollte zusätzlich Zeit beanspruchen. Daher wurde in der schwierigen Suchbedingung ein super-additiver Zusammenhang zwischen Enkodierungs- und Detektionszeiten erwartet. In Experiment 2 wurde die Zeit, die die Probanden für die reine Detektion der Zielwinkel benötigten, erhoben. Die reine Detektionszeit konnte den Unterschied in den Präsentationszeiten zwischen der schwierigen und einfachen visuellen Suche in Experiment 1 nicht vollständig erklären. Daher wurde in Experiment 3 untersucht, ob ein verlängertes oder wiederholtes Absuchen der Positionen im Suchfeld die längeren Präsentationszeiten in der schwierigen Suchbedingung in Experiment 1 bedingten. Diese Erklärung konnte ebenfalls ausgeschlossen werden. Obwohl die Probanden vor jedem Durchgang über die Anzahl der Zielwinkel informiert wurden und somit die Notwendigkeit des wiederholten Absuchens des Feldes reduziert wurde, war die Präsentationszeit in der schwierigen Suchbedingung höher als die Summe der Zeiten, die für die reine Detektion der Zielwinkel und die Enkodierung der Figuren zu erwarten gewesen wäre. Die Ergebnisse zeigten also einen super-additiven Zusammenhang zwischen Detektions- und Enkodierungszeit und favorisierten die zweistufige Enkodierungsstrategie. Die Verwendung dieser Strategie wurde in den Experimenten 4 und 5 weiter überprüft. Die Probanden wurden instruiert, nur die Positionen der Zielwinkel zu enkodieren, und es wurde untersucht, ob die benötigte Präsentationszeit den Unterschied zwischen der einfachen und schwierigen Suche in den Experimenten 1 und 3 quantitativ erklären konnte. Dies war tatsächlich der Fall. Die Ergebnisse zeigten, dass die in der schwierigen Suchbedingung zusätzlich benötigte Präsentationszeit ähnlich ausgeprägt war, wenn die Probanden die Figuren (Experimente 1 und 3) oder nur deren Positionen (Experiment 4 und 5) enkodieren mussten. Diese Ergebnisse standen im Einklang mit der von den Probanden berichteten zweistufigen Enkodierungsstrategie, die es ihnen ermöglichte, komplexe Figuren trotz interferierender Aufmerksamkeitsanforderungen in das Arbeitsgedächtnis zu enkodieren.

In der fMRT-Studie wurde die Hypothese überprüft, dass die Kapazitätslimitierung des visuellen Arbeitsgedächtnisses auf begrenzte neuronale Ressourcen zurückzuführen sei, die gemeinsam von Arbeitsgedächtnis- und Aufmerksamkeitsprozessen beansprucht werden. Dazu wurde die in der Verhaltensstudie vorgestellte Aufgabe in leicht abgeänderter Form verwendet. Um zu verhindern, dass unterschiedliche kortikale Aktivität auf unterschiedliche sensorische Stimulation zurückgeführt werden konnte, wurde die Darbietungszeit des Suchfeldes konstant gehalten. Diese Zeit wurde, basierend auf den Ergebnissen der Verhaltensstudie, so festgelegt, dass ausreichend Zeit bestand, die Figuren auch in der schwierigsten Bedingung erfolgreich zu enkodieren.

Im fMRT-Experiment 1 wurde die visuelle Suchaufgabe (leichte vs. schwierige Suche) mit der Enkodierung von Figuren kombiniert. Das Suchfeld wurde für 8 sec. präsentiert. Die Probanden waren instruiert, die Zielwinkel zu suchen und sich die damit assoziierten Figuren (ein oder drei Figuren) zu merken. Der Testreiz erschien nach weiteren 8 sec. und die Probanden gaben per Tastendruck an, ob die Figur des Testreizes mit einer der zuvor gespeicherten Figuren übereinstimmte. Die Analyse fokussierte auf die Enkodierungsphase. Die Ergebnisse zeigten für die Enkodierung von Figuren und die aufmerksamkeitsbeanspruchende visuelle Suche stark überlappende Aktivierungsmuster in verteilten posterioren und frontalen Arealen. Im rechten präfrontalen Kortex und beidseitig in der Inselregion stieg das fMRT-Signal additiv mit zunehmender Arbeitsgedächtnislast und zunehmender Aufmerksamkeitsanforderung an. Im Gegensatz dazu konnte in mehreren visuellen, parietalen und prämotorischen Arealen ein Interaktionseffekt aufgezeigt werden. In diesen Arealen war der Effekt der Arbeitsgedächtnismanipulation in Kombination mit der schwierigen visuellen Suche geringer ausgeprägt als in Kombination mit der einfachen visuellen Suche. Aktivierungen, die selektiv mit der Manipulation der Gedächtnislast korrelierten, waren im linken präfrontalen Kortex zu beobachten. Regionen im rechten präfrontalen Kortex und bilateral im visuellen Kortex waren selektiv mit der Aufmerksamkeitsmanipulation assoziiert.

Im fMRT-Experiment 2 wurde die visuelle Suchaufgabe (leichte vs. schwierige Suche) mit der Enkodierung von Positionen (ein, drei oder fünf) kombiniert. Das Suchfeld wurde für 5 sec. gezeigt. Nach einer Haltephase von 8 sec. wurde den Probanden das Suchfeld, diesmal mit einer ausgelassenen Position, erneut präsentiert. Sie gaben per Tastendruck an, ob die Position der ausgelassenen Figur mit einer der zuvor gespeicherten Positionen übereinstimmte. Die Ergebnisse zeigten für die Enkodierung von Positionen und die aufmerksamkeitsbeanspruchende visuelle Suche stark überlappende Aktivierungsmuster in verteilten posterioren und frontalen Arealen. In der Mehrheit dieser Areale war ein additiver Anstieg des fMRT-Signals unter zunehmender Arbeitsgedächtnislast und zunehmender Aufmerksamkeitsanforderung zu beobachten. In mehreren visuellen und parietalen Arealen und im rechten prämotorischen Kortex hingegen war der Effekt der Arbeitsgedächtnismanipulation in Kombination mit der schwierigen visuellen Suche geringer ausgeprägt als in

Kombination mit der einfachen visuellen Suche. Enkodierungsspezifische Aktivierungen waren in anterioren präfrontalen Regionen zu beobachten, während posterior gelegene präfrontale Areale und einige visuelle Areale selektiv mit der Aufmerksamkeitsmanipulation assoziiert waren.

Die Ergebnisse der fMRT-Experimente konnten übereinstimmend zeigen, dass die Enkodierung von Information in das visuelle Arbeitsgedächtnis (Figuren oder Positionen) und visuelle selektive Aufmerksamkeit größtenteils auf gemeinsamen neuronalen Ressourcen beruhen. Solche gemeinsamen neuronalen Korrelate konnten in posterioren Regionen lokalisiert werden. Hier war der Effekt der Gedächtnislast in Kombination mit der schwierigen visuellen Suche geringer ausgeprägt als in Kombination mit der einfachen visuellen Suche. Im Gegensatz dazu waren Regionen im Präfrontalkortex selektiv mit z. T. materialspezifischen Enkodierungsprozessen assoziiert. In diesen Regionen trat der Effekt der Gedächtnislast unter der schwierigen visuellen Suche nicht reduziert sondern zeitlich verzögert auf. Diese zeitliche Verzögerung stand im Einklang mit der in der Verhaltensstudie beobachteten Zunahme der Präsentationszeit in der schwierigen Suchbedingung. Die Ergebnisse demonstrieren, dass der Wettbewerb um begrenzte Ressourcen, die von Enkodierungs- und Aufmerksamkeitsprozessen beansprucht werden, die neuronale Verarbeitungskapazität in posterioren Arealen limitiert. Die Befunde stehen im Einklang mit kognitiven Modellen, die Aufmerksamkeitsprozessen eine funktionale Bedeutung für das kurzfristige Halten von Informationen zuschreiben (Cowan, 2001; Wheeler und Treisman, 2002) und unterstützen die Annahme, dass dem Arbeitsgedächtnis die gleichen kognitiven und neuronalen Ressourcen zugrunde liegen, die auch für die perzeptuelle Verarbeitung herangezogen werden (Slotnick, 2004; Jonides et al., 2005; Pasternak und Greenlee, 2005; Postle, 2006; Ranganath, 2006). Die Lokalisation von Interferenzeffekten zwischen Aufmerksamkeitsprozessen auf der einen Seite und der Enkodierung von Figuren oder Positionen auf der anderen Seite in ähnlichen posterioren Arelaen gibt darüber hinaus Hinweise auf eine materialunabhängige Gültigkeit dieser Modelle.

Kapazitätsbegrenzungen des visuellen Arbeitsgedächtnisses können an verschiedenen Stufen der Verarbeitung zum Tragen kommen. Die Befunde der in dieser Dissertation vorgestellten Verhaltensexperimente und fMRT-Experimente verdeutlichen, dass begrenzte neuronale und kognitive Ressourcen, die gemeinsam von den Mechanismen der visuellen Aufmerksamkeit und des visuellen Arbeitsgedächtnisses beansprucht werden, zu einem Flaschenhals in der Informationsverarbeitung während der Phase der Enkodierung führen können.

References

Aguirre GK, Zarahn E, D'Esposito M (1998). The variability of human, BOLD hemodynamic responses. Neuroimage 8:360-369.

Alvarez GA, Cavanagh P (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. Psychol Sci 15:106-111.

Amunts K, Schleier A, Bürgel U, Mohlberg H, Uylings HBM, Zille K (1999). Broca's region revisited: Ciytoarchitecture and intersubject variability. J Comp Neurol 412: 319-341.

Anderson EJ, Mannan SK, Husain M, Rees G, Sumner P, Mort DJ, McRobbie D, Kennard C (2007). Involvement of prefrontal cortex in visual search. Exp Brain Res 180:289-302.

Atkinson RC, Shiffrin RM (1968). Human memory: A proposed system and its control processes. In: The psychology of learning and motivation: Advances in research and theory (Spence KW, ed), pp 89-145. New York: Academic Press.

Awh E, Jonides J (2001). Overlapping mechanisms of attention and spatial working memory. Trends Cogn Sci 5:119-126.

Awh E, Jonides J, Reuter-Lorenz PA (1998). Rehearsal in spatial working memory. J Exp Psychol Hum Percept Perform 24:780-790.

Awh E, Vogel EK, Oh SH (2006). Interactions between attention and working memory. Neurosci 139:201-208.

Baddeley AD (1986). Working memory. New York: Oxford University Press.

Baddeley AD (1993). Working memory or working attention? In: Attention: Selection, awareness, and control. A tribute to Donald Broadbent. (Baddeley AD, Weiskrantz L, eds), pp 152-170. Oxford: Oxford University Press.

Baddeley AD (2000). The episodic buffer: a new component of working memory? Trends Cogn Sci 4:417-423.

Baddeley AD (2003). Working memory: looking back and looking forward. Nat Rev Neurosci 4:829-839.

Baddeley AD, Hitch GJ (1974). Working memory. In: The Psychology of learning and motivation (Bower GH, ed), pp 47-89. New York: Academic Press.

Barrouillet P, Bernardin S, Portrat S, Vergauwe E, Camos V (2007). Time and cognitive load in working memory. J Exp Psychol Learn Mem Cogn 33:570-585.

Bate WJ, Bullitt J, Powell LF (1963). The idler and the adventurer. New Haven: Yale UP.

Bauer RH, Fuster JM (1976). Delayed-matching and delayed-response deficit from cooling dorsolateral prefrontal cortex in monkeys. J Comp Physiol Psychol 90:293-302.

Besner D, Davies J, Daniels S (1981). Reading for meaning: The effects of concurrent articulation. Q J Exp Psychol 33:415-437.

Bledowski C, Cohen Kadosh K, Wibral M, Rahm B, Bittner RA, Hoechstetter K, Scherg M, Maurer K, Goebel R, Linden DE (2006). Mental chronometry of working memory retrieval: a combined functional magnetic resonance imaging and event-related potentials approach. J Neurosci 26:821-829.

Bledowski C, Prvulovic D, Hoechstetter K, Scherg M, Wibral M, Goebel R, Linden DEJ (2004). Localizing P300 generators in visual target and distractor processing: A combined event-related potential and functional magnetic resonance imaging study. J Neurosci 24:9353-9360.

Bor D, Duncan J, Wiseman RJ, Owen AM (2003). Encoding strategies dissociate prefrontal activity from working memory demand. Neuron 37:361-367.

Bortz J (1993). Statistik. Berlin: Springer.

Boynton GM, Engel SA, Glover GH, Heeger DJ (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. J Neurosci 16:4207-4221.

Broadbent DE (1957). A mechanical model for human attention and immediate memory. Psychol Rev 64:205-215.

Broadbent DE (1958). Perception and Communication. London: Pergamon Press.

Buckner RL (1998). Event-related fMRI and the hemodynamic response. Hum Brain Map 6:373-377.

Buckner RL, Bandettini PA, O'Craven KM, Savoy RL, Petersen SE, Raichle ME, Rosen BR (1996). Detection of cortical activation during averaged single trials of a cognitive task using functional magnetic resonance imaging. PNAS 93:14878-14883.

Bundesen C (1990). A theory of visual attention. Psychol Rev 97:523-547.

Buxton RB (2001). The elusive initial dip. Neuroimage 13:953-958.

Callicott JH, Mattay VS, Bertolino A, Finn K, Coppola R, Frank JA, Goldberg TE, Weinberger DR (1999). Physiological characteristics of capacity constraints in working memory as revealed by functional MRI. Cereb Cortex 9:20-26.

Cavanagh P, Alvarez GA (2005). Tracking multiple targets with multifocal attention. Trends Cogn Sci 9:349-354.

Chafee MV, Goldman-Rakic PS (1998). Matching patterns of activity in primate prefrontal area 8a and parietal area 7ip neurons during a spatial working memory task. J Neurophysiol 79:2919-2940.

Champod AS, Petrides M (2007). Dissociable roles of the posterior parietal and the prefrontal cortex in manipulation and monitoring processes. PNAS 104:14837-14842.

Chase WG, Ericsson KA (1981). Skilled memory. In: Cognitive skills and their acquisition (Anderson JR, ed), pp 141-189. Hillsdale, NJ: Erlbaum.

Chase WG, Simon HA (1973). The mind's eye in chess. In: Visual information processing (Chase WG, ed), pp 215-281. New York: Academic Press.

Chelazzi L (1999). Serial attention mechanisms in visual search: a critical look at the evidence. Psychol Res 62:195-219.

Chun MM, Wolfe JM (2001). Visual attention. In: Blackwell’s handbook of perception (Goldstein EB, ed), pp 272-310. Oxford: Blackwell.

Chun MM, Turk-Browne NB (2007). Interactions between attention and memory. Curr Op Neurobiol 17:177-184.

Cohen MS (1996). Rapid MRI and functional applications. In: Brain mapping. The Methods (Toga AW, Mazziotta JC, eds), pp 223-255. San Diego: Academic Press.

Coltheart M (2006). What has functional neuroimaging told us about the mind (so far)? Cortex 42:323-331.

Corbetta M, Akbudak E, Conturo TE, Snyder AZ, Ollinger JM, Drury HA, Linenweber MR, Petersen SE, Raichle ME, Van Essen DC, Shulman GL (1998). A common network of functional areas for attention and eye movements. Neuron 21:761-773

Corbetta M, Kincade JM, Shulman GL (2002). Neural systems for visual orienting and their relationships to spatial working memory. J Cogn Neurosci 14:508-523.

Corbetta M, Miezin FM, Shulman GL, Petersen SE (1993). A PET study of visuospatial attention. J Neurosci 13:1202-1226.

Corbetta M, Shulman GL (1998). Human cortical mechanisms of visual attention during orienting and search. Philos Trans R Soc Lond B Biol Sci 353:1353-1362.

Corbetta M, Shulman GL (2002). Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci 3:201-215.

Corbetta M, Shulman GL, Miezin FM, Petersen SE (1995). Superior parietal cortex activation during spatial attention shifts and visual feature conjunction. Science 270:802-805.

Coull JT, Frith CD, Frackowiak RS, Grasby PM (1996). A fronto-parietal network for rapid visual information processing: a PET study of sustained attention and working memory. Neuropsychologia 34:1085-1095.

Courtney SM (2004). Attention and cognitive control as emergent properties of information representation in working memory. Cogn Affect Behav Neurosci 4:501-516.

Courtney SM, Petit L, Maisog JM, Ungerleider LG, Haxby JV (1998). An area specialized for spatial working memory in human frontal cortex. Science 279:1347-1351.

Courtney SM, Ungerleider LG, Keil K, Haxby JV (1996). Object and spatial visual working memory activate separate neural systems in human cortex. Cereb Cortex 6:39-49.

Courtney SM, Ungerleider LG, Keil K, Haxby JV (1997). Transient and sustained activity in a distributed neural system for human working memory. Nature 386:608-611.

Cowan N (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. Psychol Bull 104:163-191.

Cowan N (1993). Activation, attention, and short-term memory. Mem Cogn 21:162-167.

Cowan N (1998). Visual and auditory working memory capacity. Trends Cogn Sci 2:77-78.

Cowan N (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. Behav Brain Sci 24:87-114.

Cowan N (2005). Working memory capacity. Hove: Psychology Press.

Culham JC, Cavanagh P, Kanwisher NG (2001). Attention response functions: characterizing brain areas using fMRI activation during parametric variations of attentional load. Neuron 32:737-745.

Curtis CE, D'Esposito M (2003). Persistent activity in the prefrontal cortex during working memory. Trends Cogn Sci 7:415-423.

Dale AM, Buckner RL (1997). Selective averaging of rapidly presented individual trials using fMRI. Hum Brain Map 5:329-340.

D'Esposito M, Aguirre GK, Zarahn E, Ballard D, Shin RK, Lease J (1998). Functional MRI studies of spatial and nonspatial working memory. Brain Res Cogn Brain Res 7:1-13.

D'Esposito M, Postle BR, Ballard D, Lease J (1999). Maintenance versus manipulation of information held in working memory: an event-related fMRI study. Brain Cogn 41:66-86.

De Fockert JW, Rees G, Frith CD, Lavie N. (2001). The role of working memory in visual selective attention. Science 291:1803-1806.

Della Sala S, Gray C, Baddeley AD, Allamano N, Wilson L (1999). Pattern span: a tool for unwelding visuo-spatial memory. Neuropsychologia 37:1189-1199.

Della Sala S, Logie RH (1993). When working memory does not work: The role of working memory in neuropsychology. In: Handbook of neuropsychology (Boller F, Spinnler H, eds), pp 1-63. Amsterdam: Elsevier.

Della Sala S, Logie RH (2002). Neuropsychological impairments of visual and spatial working memory. In: Handbook of Memory Disorders (Baddeley AD, Kopelman MD, Wilson BA, eds), pp 271-292. Chichester: Wiley.

Desimone R (1998). Visual attention mediated by biased competition in extrastriate visual cortex. Philos Trans R Soc Lond B Biol Sci 353:1245-1255.

Desimone R, Duncan J (1995). Neural mechanisms of selective visual attention. Annu Rev Neurosci 18:193-222.

Desmond JE, Glover GH (2002). Estimating sample size in functional MRI (fMRI) neuroimaging studies: statistical power analyses. J Neurosci Methods 118:115-128.

Deutsch JA, Deutsch D (1963). Attention: some theoretical considerations. Psychol Rev 70:80-90.

Devor A, Ulbert I, Dunn AK, Narayanan SN, Jones SR, Andermann ML, Boas DA, Dale AM (2005). Coupling of the cortical hemodynamic response to cortical and thalamic neuronal activity. PNAS 102:3822-3827.

Donders FC (1969). On the speed of mental processes. Acta Psychol (Amst) 30:412-431.

Donner TH, Kettermann A, Diesch E, Ostendorf F, Villringer A, Brandt SA (2002). Visual feature and conjunction searches of equal difficulty engage only partially overlapping frontoparietal networks. Neuroimage 15:16-25.

Downing PE (2000). Interactions between visual working memory and selective attention. Psychol Sci 11:467-473.

Downing PE, Dodds C (2004). Competition in visual working memory for control of search. Vis Cogn 11:689-703.

Druzgal TJ, D'Esposito M (2003). Dissecting contributions of prefrontal cortex and fusiform face area to face working memory. J Cogn Neurosci 15:771-784.

Duncan J (1980). The locus of interference in the perception of simultaneous stimuli. Psychol Rev 87:272-300.

Duncan J (1984). Selective attention and the organization of visual information. J Exp Psychol Gen 113:501-517.

Duncan J, Humphreys GW (1989). Visual search and stimulus similarity. Psychol Rev 96:433-458.

Duncan J, Ward R, Shapiro K (1994). Direct measurement of attentional dwell time in human vision. Nature 369:313-315.

Ericsson KA, Kintsch W (1995). Long-term working memory. Psychol Rev 102:211-245.

Fechner GT (1860). Elemente der Psychophysik. Leipzig: Breitkof und Härtel.

Fletcher PC, Henson RN (2001). Frontal lobes and human memory: insights from functional neuroimaging. Brain 124:849-881.

Foo P, Warren WH, Duchon A, Tarr MJ (2005). Do humans integrate routes into a cognitive map? Map- versus landmark-based navigation of novel shortcuts. J Exp Psychol Learn Mem Cogn 31:195-215.

Fougnie D, Marois R (2006). Distinct capacity limits for attention and working memory: Evidence from attentive tracking and visual working memory paradigms. Psychol Sci 17:526-534.

Fox MD, Snyder AZ, Vincent JL, Corbetta M, van Essen DC, Raichle ME (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. PNAS 102:9673-9678.

Friston KJ, Fletcher P, Josephs O, Holmes A, Rugg MD, Turner R (1998). Event-related fMRI: characterizing differential responses. Neuroimage 7:30-40.

Friston KJ, Holmes AP, Poline JB, Grasby PJ, Williams SC, Frackowiak RS, Turner R (1995). Analysis of fMRI time-series revisited. Neuroimage 2:45-53.

Funahashi S, Bruce CJ, Goldman-Rakic PS (1993). Dorsolateral prefrontal lesions and oculomotor delayed-response performance: evidence for mnemonic "scotomas". J Neurosci 13:1479-1497.

Fuster JM, Alexander GE (1971). Neuron activity related to short-term memory. Science 173:652-654.

Garden S, Cornoldi C, Logie RH (2002). Visuo-spatial working memory in navigation. Appl Cogn Psychol 16:35-50.

Genovese CR, Lazar NA, Nichols T (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. Neuroimage 15:870-878.

Giesbrecht B, Woldorff MG, Song AW, Mangun GR (2003). Neural mechanisms of top-down control during spatial and feature attention. Neuroimage 19:496-512.

Gitelman DR, Parrish TB, Friston KJ, Mesulam MM (2002). Functional anatomy of visual search: regional segregations within the frontal eye fields and effective connectivity of the superior colliculus. Neuroimage 15:970-982.

Glahn DC, Kim J, Cohen MS, Poutanen VP, Therman S, Bava S, Van Erp TG, Manninen M, Huttunen M, Lonnqvist J, Standertskjold-Nordenstam CG, Cannon TD (2002). Maintenance and manipulation in spatial working memory: dissociations in the prefrontal cortex. Neuroimage 17:201-213.

Gobet F, Lane PC, Croker S, Cheng PC, Jones G, Oliver I, Pine JM (2001). Chunking mechanisms in human learning. Trends Cogn Sci 5:236-243.

Gold JM, Wilk CM, McMahon RP, Buchanan RW, Luck SJ (2003). Working memory for visual features and conjunctions in schizophrenia. J Abnorm Psychol 112:61-71.

Goldberg ME, Bisley J, Powell KD, Gottlieb J, Kusunoki M (2002). The role of the lateral intraparietal area of the monkey in the generation of saccades and visuospatial attention. Ann N Y Acad Sci 956:205-215.

Goldman PS, Rosvold HE (1970). Localization of function within the dorsolateral prefrontal cortex of the rhesus monkey. Exp Neurol 27:291-304.

Goldman-Rakic PS (1987). Circuitry of the prefrontal cortex and the regulation of behavior by representational memory. In: Handbook of neurobiology (Mountcastle VB, ed), pp 373-417. Bethesda: American Physiologigal Society.

Gusnard DA, Raichle ME (2001). Searching for a baseline: functional imaging and the resting human brain. Nat Rev Neurosci 2:685-694.

Hampson M, Driesen NR, Skudlarski P, Gore JC, Constable RT (2006). Brain connectivity related to working memory performance. J Neurosci 26:13338-13343.

Hasher L, Zacks, R. T. (1979). Automatic and effortful processes in memory. J Exp Psychol Gen 108:356-388.

Heeger DJ, Ress D (2002). What does fMRI tell us about neuronal activity? Nat Rev Neurosci 3:142-151.

Henson R (2003). Analysis of fMRI timeseries: Linear time-invariant models, event-related fMRI and optimal experimental design. In: Human Brain Function (Frackowiak RSJ, Frith C, Dolan R, Friston KJ, Price CJ, Zeki S, Ashburner J, Penny WD, eds), pp 1-12. New York: Academic Press.

Henson R (2005). What can functional neuroimaging tell the experimental psychologist? Q J Exp Psychol A 58:193-233.

Hitch G (1990). Developmental fractionation of working memory. In: Neuropsychological impairments of short-term memory (Vallar G, Shallice T, eds), pp 221-246. Cambridge, UK: Cambradige University Press.

Hitch GJ, Halliday S, Schaafstal AM, Schraagen JM (1988). Visual working memory in young children. Mem Cognit 16:120-132.

Honey GD, Fletcher, PC (2006). Investigating principles of human brain function underlying working memory: What insights from schizophrenia? Neurosci 139:59-71.

Horowitz TS, Wolfe JM (2001). Search for multiple targets: remember the targets, forget the search. Percept Psychophys 63:272-285.

Howseman AM, Bowtell RW (1999). Functional magnetic resonance imaging: imaging techniques and contrast mechanisms. Philos Trans R Soc Lond B Biol Sci 354:1179-1194.

Husain M, Nachev P (2007). Space and the parietal cortex. Trends Cogn Sci 11:30-36.

Irwin DE (1992). Memory for position and identity across eye movements. J Exp Psychol Learn Mem Cogn 18:307-317.

Jack AI, Sylvester CM, Corbetta M (2006). Losing our brainless minds: how neuroimaging informs cognition. Cortex 42:418-421.

James W (1890). Principles of psychology. New York: Holt.

Jha AP (2002). Tracking the time-course of attentional involvement in spatial working memory: an event-related potential investigation. Cogn Brain Res 15:61-69.

Jha AP, McCarthy G (2000). The influence of memory load upon delay-interval activity in a working-memory task: an event-related functional MRI study. J Cogn Neurosci 12 Suppl 2:90-105.

Jolicœur P, Dell'Acqua R (1998). The demonstration of short-term consolidation. Cogn Psychol 36:138-202.

Jolicœur P, Dell’Acqua R. (1999). Attentional and structural constraints on visual encoding. Psychol Res 62:154-164.

Jonides J, Lacey SC, Nee DE (2005). Processes of working memory in mind and brain. Curr Direct Psychol Sci 14:2-5.

Jonides J, Nee DE, Berman MG (2006). What has functional neuroimaging told us about the mind? So many examples, so little space. Cortex 42:414-417.

Josephs O, Henson RN (1999). Event-related functional magnetic resonance imaging: modelling, inference and optimization. Philos Trans R Soc Lond B Biol Sci 354:1215-1228.

Kahneman D (1973). Attention and effort. New York: Prentice-Hall.

Kanwisher N, Wojciulik E (2000). Visual attention: insights from brain imaging. Nat Rev Neurosci 1:91-100.

Kastner S, Pinsk MA, De Weerd P, Desimone R, Ungerleider LG (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. Neuron 22:751–761

Kastner S, Ungerleider LG (2000). Mechanisms of visual attention in the human cortex. Annu Rev Neurosci 23:315-341.

Kiebel SJ, Holmes AP (2003). The general linear model. In: Human Brain Function (Frackowiak RSJ, Frith C, Dolan R, Friston KJ, Price CJ, Zeki S, Ashburner J, Penny WD, eds), pp 1-36. New York: Academic Press.

Kuperberg G, Heckers S (2000). Schizophrenia and cognitive function. Curr Opin Neurobiol 10:205-210.

Kwong KK (1995). Functional magnetic resonance imaging with echo planar imaging. Magn Reson Q 11:1-20.

LaBar KS, Gitelman DR, Parrish TB, Mesulam M (1999). Neuroanatomic overlap of working memory and spatial attention networks: a functional MRI comparison within subjects. Neuroimage 10:695-704.

Laurienti PJ, Burdette JH, Wallace MT, Yen YF, Field AS, Stein BE (2002). Deactivation of sensory-specific cortex by cross-modal stimuli. J Cogn Neurosci 14:420-429.

Lavie N, De Fockert J. (2005). The role of working memory in attentional capture. Psychon Bull Rev 12:669-674.

Lavie N, Hirst A, de Fockert JW, Viding E. (2004) Load theory of selective attention and cognitive control. J Exp Psychol Gen 133: 339-254.

Leonards U, Sunaert S, Van Hecke P, Orban GA (2000). Attention mechanisms in visual search - an fMRI study. J Cogn Neurosci 12:61-75.

Lepsien J, Griffin IC, Devlin JT, Nobre AC (2005). Directing spatial attention in mental representations: Interactions between attentional orienting and working-memory load. Neuroimage 26:733-743.

Lepsien J, Nobre AC (2007). Attentional modulation of object representations in working memory. Cereb Cortex 17:2072-2083.

Leung HC, Gore JC, Goldman-Rakic PS (2002). Sustained mnemonic response in the human middle frontal gyrus during on-line storage of spatial memoranda. J Cogn Neurosci 14:659-671.

Leung HC, Oh H, Ferri J, Yi Y (2007). Load response functions in the human spatial working memory circuit during location memory updating. Neuroimage 35:368-377.

Leung HC, Seelig D, Gore JC (2004). The effect of memory load on cortical activity in the spatial working memory circuit. Cogn Affect Behav Neurosci 4:553-563.

Linden DEJ, Bittner RA, Muckli L, Waltz JA, Kriegeskorte N, Goebel R, Singer W, Munk MH (2003). Cortical capacity constraints for visual working memory: dissociation of fMRI load effects in a fronto-parietal network. Neuroimage 20:1518-1530.

Linden DEJ, Prvulovic D, Formisano E, Vollinger M, Zanella FE, Goebel R, Dierks T (1999). The functional neuroanatomy of target detection: an fMRI study of visual and auditory oddball tasks. Cereb Cortex 9:815-823.

Logie RH (1995). Visuo-spatial working memory. Hove, UK: Erlbaum.

Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A (2001). Neurophysiological investigation of the basis of the fMRI signal. Nature 412:150-157.

Luck SJ, Chelazzi L, Hillyard SA, Desimone R (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. J Neurophysiol 77:24-42.

Luck SJ, Vogel EK (1997). The capacity of visual working memory for features and conjunctions. Nature 390:279-281.

Manoach DS, White NS, Lindgren KA, Heckers S, Coleman MJ, Dubal S, Holzman PS (2004). Hemispheric specialization of the lateral prefrontal cortex for strategic processing during spatial and shape working memory. Neuroimage 21:894-903.

Mansfield P (1977). Multi-planar image formation using NMR spin echoes. J Phys C:L55-L58.

Marois R, Ivanoff J (2005). Capacity limits of information processing in the brain. Trends Cogn Sci 9:296-305.

Marr D (1982). Vision. San Francisco: Freeman.

Martinez A, Anllo-Vento L, Sereno MI, Frank LR, Buxton RB, Dubowitz DJ, Wong EC, Hinrichs H, Heinze HJ, Hillyard SA (1999). Involvement of striate and extrastriate visual cortical areas in spatial attention. Nat Neurosci 2:364-369.

Mason MF, Norton MI, Van Horn JD, Wegner DM, Grafton ST, Macrae CN (2007). Wandering minds: the default network and stimulus-independent thought. Science 315: 393-395.

Mazoyer B, Zago L, Mellet E, Bricogne S, Etard O, Houde O, Crivello F, Joliot M, Petit L, Tzourio-Mazoyer N (2001). Cortical networks for working memory and executive functions sustain the conscious resting state in man. Brain Res Bull 54:287-298.

McKiernan KA, Kaufman JN, Kucera-Thompson J, Binder JR (2003). A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. J Cogn Neurosci 15:394-408.

Miezin FM, Maccotta L, Ollinger JM, Petersen SE, Buckner RL (2000). Characterizing the hemodynamic response: effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. Neuroimage 11:735-759.

Miller GA (1956). The magical number seven plus or minus two: some limits on our capacity for processing information. Psychol Rev 63:81-97.

Miller EK, Cohen JD (2001). An integrative theory of prefrontal cortex function. Annu Rev Neurosci 24:167-202.

Miller EK, Desimone R (1994). Parallel neuronal mechanisms for short-term memory. Science 263:520-522.

Miller GA, Galanter E, Pribram KH (1960). Plans and the structure of behavior. New York: Holt, Rinehart and Winston.

Mishkin M, Ungerleider LG, Macko KA (1983). Object vision and spatial vision: two cortical pathways. Trends Neurosci 6:414-417.

Mitchell RL, Elliott R, Woodruff PW (2001). fMRI and cognitive dysfunction in schizophrenia. Trends Cogn Sci 5:71-81.

Mohr HM, Goebel R, Linden DE (2006). Content- and task-specific dissociations of frontal activity during maintenance and manipulation in visual working memory. J Neurosci 26:4465-4471.

Mohr HM, Linden DE (2005). Separation of the systems for color and spatial manipulation in working memory revealed by a dual-task procedure. J Cogn Neurosci 17:355-366.

Motter BC (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. J Neurophysiol 70:909-919.

Munk MH, Linden DE, Muckli L, Lanfermann H, Zanella FE, Singer W, Goebel R (2002). Distributed cortical systems in visual short-term memory revealed by event-related functional magnetic resonance imaging. Cereb Cortex 12:866-876.

Murray DJ (1968). Articulation and acoustic confusability in short-term memory. J Exp Psychol Gen 78:679-684.

Neisser U (1974). Kognitive Psychologie. Stuttgart: Klett.

Nieder A (2004). The number domain- can we count on parietal cortex? Neuron 44:407-409.

Niessing J, Ebisch B, Schmidt KE, Niessing M, Singer W, Galuske RAW (2005). Hemodynamic signals correlate tightly with synchronized gamma oscillations. Science 309:948-951.

Nikolić D, Singer W. (2007). Creation of visual long-term memory. Percept Psychophys 69:904-912.

Nobre AC, Coull JT, Maquet P, Frith CD, Vandenberghe R, Mesulam MM (2004). Orienting attention to locations in perceptual versus mental representations. J Cogn Neurosci 16:363-373.

Nobre AC, Coull JT, Walsh V, Frith CD (2003). Brain activations during visual search: contributions of search efficiency versus feature binding. Neuroimage 18:91-103.

Nobre AC, Sebestyen GN, Gitelman DR, Mesulam MM, Frackowiak RS, Frith CD (1997). Functional localization of the system for visuospatial attention using positron emission tomography. Brain 120:515-533.

Norman DA, Shallice T (1986). Attention to action: Willed and automatic control of behaviour. In: Consciousness and self-regulation: Advances in research and theory (Davidson RJ, Schwartz GE, Shapiro DE, eds), pp 1-18. New York: Plenum.

Ogawa S, Lee TM, Kay AR, Tank DW (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. PNAS 87:9868-9872.

Oh SH, Kim MS (2004). The role of spatial working memory in visual search efficiency. Psychon Bull Rev 11:275-281.

Ollinger JM, Corbetta M, Shulman GL (2001). Separating processes within a trial in event-related functional MRI. Neuroimage 13:218-229.

Olsson H, Poom L (2005). Visual memory needs categories. PNAS 102:8776-8780.

Owen AM, Herrod NJ, Menon DK, Clark JC, Downey SP, Carpenter TA, Minhas PS, Turkheimer FE, Williams EJ, Robbins TW, Sahakian BJ, Petrides M, Pickard JD (1999). Redefining the functional organization of working memory processes within human lateral prefrontal cortex. Eur J Neurosci 11:567-574.

Owen AM, Iddon JL, Hodges JR, Summers BA, Robbins TW (1997). Spatial and non-spatial working memory at different stages of Parkinson's disease. Neuropsychologia 35:519-532.

Pashler, H (1988). Familiarity and visual change detection. Percept Psychophys 44:369-378.

Pashler H (1994). Dual-task interference in simple tasks: data and theory. Psychol Bull 116:220-244.

Pashler HE (1998). The psychology of attention. Cambridge: MIT Press.

Passingham D, Sakai K (2004). The prefrontal cortex and working memory: physiology and brain imaging. Curr Opin Neurobiol 14:163-168.

Pasternak T, Greenlee MW (2005). Working memory in primate sensory systems. Nat Rev Neurosci: 97-107.

Penny WD, Holmes AJ (2003). Random-effects analysis. In: Human Brain Function (Frackowiak RSJ, Friston KJ, Frith C, Dolan R, Price CJ, Zeki S, Ashburner J, Penny WD, eds), pp 1-12. New York: Academic Press.

Pessoa L, Gutierrez E, Bandettini P, Ungerleider L (2002). Neural correlates of visual working memory: fMRI amplitude predicts task performance. Neuron 35:975-987.

Pessoa L, Kastner S, Ungerleider LG (2003). Neuroimaging studies of attention: from modulation of sensory processing to top-down control. J Neurosci 23:3990-3998.

Pessoa L, Ungerleider LG (2004). Top-down mechanisms for working memory and attentional processes. In: The new cognitive neurosciences (Gazzaniga MS, ed), pp 919-930. Cambridge: MIT Press.

Peterson MS, Kramer AF, Wang RF, Irwin DE, McCarley JS (2001). Visual search has memory. Psychol Sci 12:287-292.

Petrides M (1994). Frontal lobes and working memory: evidence from investigations of the effects of cortical excisions in nonhuman primates. In: Handbook of neuropsychology (Boller F, Grafman J, eds), pp 59-82. Amsterdam: Elsevier.

Petrides M (2000). The role of the mid-dorsolateral prefrontal cortex in working memory. Exp Brain Res 133:44-54.

Phillips WA (1974). On the distinction between sensory storage and short-term visual memory. Percept Psychophys 16:283-290.

Phillips WA, Christie DFM (1977). Components of visual memory. Q J Exp Psychol 29:117-133.

Pollmann S, von Cramon DY (2000). Object working memory and visuospatial processing: functional neuroanatomy analyzed by event-related fMRI. Exp Brain Res 133:12-22.

Posner MI, Petersen SE (1990). The attention system of the human brain. Annu Rev Neurosci 13:25-42.

Postle BR (2005). Delay-period activity in the prefrontal cortex: one function is sensory gating. J Cogn Neurosci 17:1679-1690.

Postle BR (2006). Working memory as an emergent property of the mind and brain. Neurosci 139:23-38.

Postle BR, Awh E, Jonides J, Smith EE, D'Esposito M (2004). The where and how of attention-based rehearsal in spatial working memory. Brain Res Cogn Brain Res 20:194-205.

Postle BR, D'Esposito M (1999). "What"-Then-Where" in visual working memory: an event-related fMRI study. J Cogn Neurosci 11:585-597.

Postle BR, Druzgal TJ, D'Esposito M (2003). Seeking the neural substrates of visual working memory storage. Cortex 39:927-946.

Postle BR, Jonides J, Smith EE, Corkin S, Growdon JH (1997). Spatial, but not object, delayed response is impaired in early Parkinson's disease. Neuropsychol 11:171-179.

Pylyshyn ZW, Storm RW (1988). Tracking multiple independent targets: evidence for a parallel tracking mechanism. Spat Vis 3:179-197.

Ranganath C (2006). Working memory for visual objects: complementary roles of inferior temporal, medial temporal, and prefrontal cortex. Neurosci 139:277-289.

Ranganath C, DeGutis J, D'Esposito M (2004). Category-specific modulation of inferior temporal activity during working memory encoding and maintenance. Brain Res Cogn Brain Res 20:37-45.

Rao SC, Rainer G, Miller EK (1997). Integration of what and where in the primate prefrontal cortex. Science 276:821-824.

Rees G, Howseman A, Josephs O, Frith CD, Friston KJ, Frackowiak RSJ, Turner R (1997). Characterizing the relationship between BOLD contrast and regional cerebral blood flow measurements by varying the stimulus presentation rate. Neuroimage 6:270-278

Rensink RA (2000). Seeing, sensing, and scrutinizing. Vision Res 40:1469-1487.

Rensink RA (2002). Change detection. Annu Rev Psychol 53:245-277.

Reynolds JH, Chelazzi L, Desimone R (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. J Neurosci 19:1736-1753.

Robertson LC (2003). Binding, spatial attention and perceptual awareness. Nat Rev Neurosci 4:93-102.

Robertson I, Treisman A, Grabowecky M (1997). The Interaction of spatial and object pathways: Evidence from Balint's syndrome. J Cogn Neurosci 9:295-317.

Rosen AC, Rao SM, Caffarra P, Scaglioni A, Bobholz JA, Woodley SJ, Hammeke TA, Cunningham JM, Prieto TE, Binder JR (1999). Neural basis of endogenous and exogenous spatial orienting. A functional MRI study. J Cogn Neurosci 11:135-152.

Rowe JB, Toni I, Josephs O, Frackowiak RS, Passingham RE (2000). The prefrontal cortex: response selection or maintenance within working memory? Science 288:1656-1660.

Sack AT, Camprodon JA, Pascual-Leone A, Goebel R (2005). The dynamics of interhemispheric compensatory processes in mental imagery. Science 208: 702-704.

Sakai K, Passingham RE (2003). Prefrontal interactions reflect future task operations. Nat Neurosci 6:75-81.

Sala JB, Rama P, Courtney SM (2003). Functional topography of a distributed neural system for spatial and nonspatial information maintenance in working memory. Neuropsychologia 41:341-356.

Savoy RL, Bandettini PA, O'Craven KM, Kwong KK, Davis TL, Baker JR, Weiskoff RM, Rosen BR (1995). Pushing the temporal resolution of fMRI: Studies of very brief visual stimuli, onset variability and asynchrony, and stimulus-correlated changes in noise. Proc Soc Magn Res Third Sci Meet Exhi 2:450.

Schmidt BK, Vogel EK, Woodman GF, Luck SJ (2002). Voluntary and automatic attentional control of visual working memory. Percept Psychophys 64:754-763.

Schneider W (1999). Working memory in a multilevel hybrid connectionist control architecture (CAP2). In: Models of working memory (Miyake A, Shah P, eds), pp 340-374. Cambridge, UK: Cambridge University Press.

Scholl BJ (2001). Objects and attention: the state of the art. Cogn 80:1-46.

Serences JT, Schwarzbach J, Courtney SM, Golay X, Yantis S (2004). Control of object-based attention in human cortex. Cereb Cortex 14:1346-1357.

Shafritz KM, Gore JC, Marois R (2002). The role of the parietal cortex in visual feature binding. PNAS 99:10917-10922.

Shiffrin RM, Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. Psychol Rev 84:127-190.

Shmuel A, Augath M, Oeltermann A, Logothetis NK (2006). Negative functional MRI response correlates with decreases in neuronal activity in monkey visual area V1. Nat Neurosci 9:569-577.

Shmuel A, Yacoub E, Pfeuffer J, Van de Moortele PF, Adriany G, Hu X, Ugurbil K (2002). Sustained negative BOLD, blood flow and oxygen consumption response and its coupling to the positive response in the human brain. Neuron 36: 1195-1210.

Shomstein S, Behrmann M (2006). Cortical systems mediating visual attention to both objects and spatial locations. PNAS 103:11387-11392.

Shulman GL, Astafiev SV, McAvoy MP, d'Avossa G, Corbetta M (2007). Right TPJ deactivation during visual search: functional significance and support for a filter hypothesis. Cereb Cortex 17:2625-2633.

Shulman GL, Fiez JA, Corbetta M, Buckner RL, Miezin FM, Raichle ME, Pertersen SE (1997). Common blood flow changes across visual tasks: II. Decreases in cerebral cortex. J Cogn Neurosci 9:648-663.

Sigman M, Dehaene S (2005). Parsing a cognitive task: a characterization of the mind's bottleneck. Plos Biol 3:334-349.

Simons DJ, Rensink RA (2005). Change blindness: past, present, and future. Trends Cogn Sci 9:16-20.

Slotnick SD (2004). Visual working memory and visual perception recruit common neural substrates. Behav Cogn Neurosci Rev:207-221.

Smith EE JJ, Koeppe RA, Awh E, Schumacher EH, Minoshima S (1995). Spatial versus object working memory: PET investigations. J Cogn Neurosci 7:337-356.

Smith EE, Jonides J (1999). Storage and executive processes in the frontal lobes. Science 283:1657-1661.

Smyth MM, Scholey KA (1994). Interference in immediate spatial memory. Mem Cogn 22:1-13.

Sperling G (1960). The information available in brief visual presentations. Psychol Monogr Gen Appl 74:1-29.

Sternberg S (1966). High-speed scanning in human memory. Science 153:652-654.

Sternberg S (1969). The discovery of processing stages: extension of Donders' method. In: Attention and performance II (Koster WG, ed), pp 276-315. Amsterdam: North-Holland.

Sternberg S (2001). Separate modifiability, mental modules, and the use of pure and composite measures to reveal them. Acta Psychol (Amst) 106:147-246.

Thorndyke PW, Hayes-Roth B (1982). Differences in spatial knowledge acquired from maps and navigation. Cogn Psychol 14:560-589.

Todd JJ, Marois R (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. Nature 428:751-754.

Tomasi D, Ernst T, Caparelli EC, Chang L (2006). Common deactivation patterns during working memory and visual attention tasks: an intra-subject fMRI study at 4 Tesla. Hum Brain Map 27:694-705.

Treisman A (1960). Contextuel cues in selective listening. Q J Exp Psychol 12:242-248.

Treisman A (1998). Feature binding, attention and object perception. Philos Trans R Soc Lond B Biol Sci 353:1295-1306.

Treisman A, Gelade G (1980). A feature-integration theory of attention. Cogn Psychol 12:97-136.

Treisman A, Gormican S (1988). Feature analysis in early vision: evidence from search asymmetries. Psychol Rev 95:15-48.

Treisman A, Sato S (1990). Conjunction search revisited. J Exp Psychol Hum Percept Perform 16:459-478.

Tresch MC, Sinnamon HM, Seamon JG (1993). Double dissociation of spatial and object visual memory: evidence from selective interference in intact human subjects. Neuropsychologia 31:211-219.

Uludag K, Dubowitz DJ, Yoder EJ, Restom K, Liu TT, Buxton RB (2004). Coupling of cerebral blood flow and oxygen consumption during physiological activation and deactivation measured with fMRI. Neuroimage 23:148-155.

Ungerleider MG, Mishkin M (1982). Two cortical visual systems. In: Analysis of visual behavior (Ingle DJ, Goodale MA, Mansfield RJW, eds), pp 549–586. Cambridge: MIT Press.

Van Asselen M, Fritschy E, Postma A (2006). The influence of intentional and incidental learning on acquiring spatial knowledge during navigation. Psychol Res 70:151-156.

Vazquez AL, Noll DC (1997). Nonlinear aspects of the BOLD response in functional MRI. Neuroimage 7:108-118.

Vecera SP, Farah MJ (1994). Does visual attention select objects or locations? J Exp Psychol Gen 123:146-160.

Vogel EK, Machizawa MG (2004). Neural activity predicts individual differences in visual working memory capacity. Nature 428:748-751.

Vogel EK, Woodman GF, Luck SJ (2001). Storage of features, conjunctions and objects in visual working memory. J Exp Psychol Hum Percept Perform 27:92-114.

Wager TD, Smith EE (2003). Neuroimaging studies of working memory: a meta-analysis. Cogn Affect Behav Neurosci 3:255-274.

Wesenick M-B (2003). Limitations of human visual working memory. Dissertation, München: Ludwig-Maximilians Universität.

Wheeler ME, Treisman AM (2002). Binding in short-term visual memory. J Exp Psychol Gen 131:48-64.

Wilkinson D, Halligan P (2004). The relevance of behavioural measures for functional-imaging studies of cognition. Nat Rev Neurosci 5:67-73.

Wilson FA, Scalaidhe SP, Goldman-Rakic PS (1993). Dissociation of object and spatial processing domains in primate prefrontal cortex. Science 260:1955-1958.

Wolfe JM (1998a). Visual search. In: Attention (Pashler H, ed), pp 13-74. London: University College London Press.

Wolfe JM (1998b). What can 1 million trials tell us about visual search? Psychol Sci 9:33-39.

Woodman GF, Luck SJ (2004). Visual search is slowed when visuospatial working memory is occupied. Psychon Bull Rev 11:269-274.

Woodman GF, Vogel EK, Luck SJ (2001). Visual search remains efficient when visual working memory is full. Psychol Sci 12:219-224.

Xu Y, Chun MM (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. Nature 440:91-95.

Yantis S, Schwarzbach J, Serences JT, Carlson RL, Steinmetz MA, Pekar JJ, Courtney SM (2002). Transient neural activity in human parietal cortex during spatial attention shifts. Nat Neurosci 5:995-1002.

Yantis S, Serences JT (2003). Cortical mechanisms of space-based and object-based attentional control. Curr Opin Neurobiol 13:187-193.

Zarahn E, Aguirre G, D'Esposito M (1997). A trial-based experimental design for fMRI. Neuroimage 6:122-138.

CURRICULUM VITAE

Jutta Susanne Mayer
geboren am 29.05.1977 in Mainz, ledig

	Berufliche Tätigkeit
seit Okt. 2002	Doktorandin im Labor für Klinische Neurophysiologie und Neuroimaging an der Klinik für Psychiatrie, Psychosomatik und Psychotherapie der Johann Wolfgang Goethe-Universität, Frankfurt am Main Betreuer: Prof. R. Sireteanu und Prof. D. Linden Thema: Interactions between visual working memory and attention: Evidence from fMRI
Jan. 2004-März 2006	Promotionsstipendium der Studienstiftung des deutschen Volkes
	Psychiatrisches Jahr im Rahmen des Weiterbildungsstudiengangs Psychologische Psychotherapie:
Dez. 2003 - Febr. 2004	Praxengemeinschaft für Neuropsychologie, Psychotherapie und Logopädie, Frankfurt am Main
März 2004 - Mai 2005	Klinik für Psychiatrie, Psychosomatik und Psychotherapie der Johann Wolfgang Goethe-Universität, Frankfurt am Main
seit April 2006	Psychotherapeutische Tätigkeit im Rahmen des Weiterbildungsstudiengangs Psychologische Psychotherapie in der Ambulanz des Ausbildungsinstituts an der Johann Wolfgang Goethe-Universität, Frankfurt am Main
seit Mai 2006	Klinische Tätigkeit (Demenzdiagnostik) in der Klinik für Psychiatrie, Psychosomatik und Psychotherapie der Johann Wolfgang Goethe-Universität, Frankfurt am Main
	Publikationen
2007	Mayer, J. S., Bittner, R. B., Nikolić, D., Bledowski, C., Goebel, R., Linden, D. E. J. (2007). Common neural substrates for visual working memory and attention, *NeuroImage*, *36*, 441-453.
	Mayer, J. S., Bittner, R. A., Linden, D. E. J., Nikolić, D. Attentional demand influences strategies for encoding into visual working memory, *Advances in Cognitive Psychology*, *3*, 429-448.
	Michal, M., Röder, C., Mayer, J. S., Lengler, U., Krakow, K. (2007). Spontaneous dissociation during functional MRI experiments. *Journal of Psychiatric Research*, *41*, 69-73.

	Studium
Okt. 1996 – April 2002	Psychologie an der Johannes Gutenberg-Universität in Mainz
Okt. 1998	Vordiplom in Psychologie
April 2002	Hauptdiplom in Psychologie (Gesamtnote: 1,1)
	Diplomarbeit: Konditionierung von zusammengesetzten Stimuli bei Honigbienen: Diskrimination zwischen Stimuluspaaren und ihren Komponenten
Okt. 2003- Okt. 2008	Weiterbildungsstudiengang Psychologische Psychotherapie (Schwerpunkt: Verhaltenstherapie) an der Johann Wolfgang Goethe-Universität, Frankfurt am Main
Nov. 2008	Approbation Psychologische Psychotherapeutin
	Tätigkeiten neben dem Studium
Jan. 1998 – Juli 1998	Nachtbereitschaftstätigkeit in einem sozialpsychiatrischen Wohnheim des Evangelischen Vereins für Innere Mission in Nassau in Wiesbaden
Aug. 1998 – Juli 1999 und März 2000 – Sept. 2000	Nachtbereitschaftstätigkeit in einem sozialpsychiatrischen Wohnheim der Paritätischen Psychiatrischen Dienste in Mainz
	Praktika
Aug. 1999 – Feb. 2000	Forschungspraktikum am Békésy Laboratory of Neurobiology an der University of Hawaii in Honolulu (USA)
März 2000 – Mai 2000	Praktikum am DRK-Schmerzzentrum in Mainz
Juli 2001 – Aug. 2001	Praktikum in der Abteilung für Neuropsychologie am Städtischen Krankenhaus München-Bogenhausen
	Schulausbildung
1983 – 1987	Grundschule Mainz-Laubenheim
1987 – 1996	Gymnasium am Kurfürstlichen Schloss, Mainz Abschluss: Allgemeine Hochschulreife