WORKING MEMORY LOAD EFFECTS OF DIFFERENT TASKS DURING MULTITASKING MAY DIFFER: AN FMRI STUDY

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF

MASTER OF SCIENCE

IN

NEUROSCIENCE

By

Sümeyye Öztürk

July 2025

WORKING MEMORY LOAD EFFECTS OF DIFFERENT TASKS DURING MULTITASKING MAY DIFFER: AN FMRI STUDY
By Sümeyye Öztürk
July 2025

We certify that we have read this thesis and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Ausaf Ahmed Farooqui(Advisor)

Burcu Ayşen Ürgen

Buse Merve Ürgen

Approved for the Graduate School of Engineering and Science:

Orhan Arıkan
Director of the Graduate School

ABSTRACT

WORKING MEMORY LOAD EFFECTS OF DIFFERENT TASKS DURING MULTITASKING MAY DIFFER: AN FMRI STUDY

Sümeyye Öztürk
M.S. in Neuroscience
Advisor: Ausaf Ahmed Farooqui
July 2025

Multiple demand (MD) regions are well-known for their activation pattern in task-related events and are especially sensitive to working memory load (WM) conditions. Different WM loads deplete varying amounts of WM capacity and may lead to interference. However, whether MD regions are sensitive to precision in addition to cognitive load still remains unclear. This thesis investigates whether challenging multitasking induces interference and whether brain responses to demanding tasks are modulated by precision. Thirty healthy participants (M = 27.1) underwent two fMRI runs and completed a hierarchically nested visual WM task, where task A's items were maintained during task B under alternating high and low WM loads. Parametric modulation, whole-brain, and region-of-interest (ROI) analyses were performed. Results showed that MD regions were activated during the encoding phase of task B, whereas task A exhibited a deactivation pattern. However, tasks were not differentiated during their corresponding retrieval steps. Brain regions correlating with performance accuracy also did not differ. The Default Mode Network (DMN) showed sustained deactivation during recall, regardless of task type, indicating domain-general suppression mechanisms. These findings support hybrid WM models that allow flexible control mechanisms and discrete capacity limits. The MD network appears to adjust engagement depending on task context and phase, aligning with hierarchical models. Task-specific encoding but shared recall dynamics suggest that while storage may be selectively distributed, retrieval relies on generalized control. This enhances understanding of cognitive control–memory interactions and highlights the need to examine individual differences to inform interventions for WM-related disorders.

Keywords: Working memory, cognitive control, multiple demand regions, multitasking.

ÖZET

ÇOKLU GÖREV SIRASINDA FARKLI GÖREVLERİN ÇALIŞAN BELLEK YÜKÜNÜN ETKİLERİ FARKLILAŞABİLİR: BİR FMRG ÇALIŞMASI

Sümeyye Öztürk
Nörobilim, Yüksek Lisans
Tez Danışmanı: Ausaf Ahmed Farooqui
Temmuz 2025

Çoklu talep (MD) bölgeleri, görevle ilişkili olaylardaki aktivasyon örüntüleriyle bilinir ve özellikle çalışma belleği (ÇB) yükü koşullarına duyarlıdır. Farklı ÇB yükleri, değişen miktarlarda ÇB kapasitesini tüketebilir ve girişime yol açabilir. Ancak MD bölgelerinin bilişsel yüke ek olarak hassasiyete de duyarlı olup olmadığı hâlen belirsizdir. Bu tez, zorlu çoklu görevlerin girişime neden olup olmadığını ve beynin bilişsel açıdan zorlayıcı görevlere verdiği yanıtların hassasiyet tarafından modüle edilip edilmediğini araştırmaktadır. Otuz sağlıklı katılımcı ($M = 27.1$), iki fMRI oturumuna katılmış ve hiyerarşik olarak iç içe geçmiş bir görsel ÇB görevini tamamlamıştır. Bu görevde, Görev A'nın öğeleri, Görev B sırasında dönüşümlü olarak yüksek ve düşük ÇB yükleri altında korunmuştur. Parametrik modülasyon, tüm beyin ve ilgi bölgesi (ROI) analizleri gerçekleştirilmiştir. Sonuçlar, MD bölgelerinin Görev B'nin kodlama aşamasında aktive olduğunu, Görev A'nın ise bir deaktivasyon örüntüsü sergilediğini göstermiştir. Ancak görevler, ilgili geri çağırma adımlarında farklılaşmamıştır. Performans doğruluğu ile ilişkili beyin bölgelerinde de bir farklılık gözlenmemiştir. Varsayılan Mod Ağı (DMN), görev türünden bağımsız olarak geri çağırma sırasında sürdürülen deaktivasyon göstermiş ve alan-genel baskılama mekanizmalarına işaret etmiştir. Bulgular, esnek kontrol mekanizmalarına ve ayrık kapasite sınırlılıklarına izin veren hibrit ÇB modellerini desteklemektedir. MD ağı, görev bağlamı ve aşamasına göre katılımını ayarlıyor gibi görünmekte ve hiyerarşik modellerle uyum göstermektedir. Görev-özel kodlama ancak paylaşılan geri çağırma dinamikleri, depolamanın seçici biçimde dağıtılabileceğini, buna karşın geri çağırmanın genelleştirilmiş bir kontrol yapısına dayandığını düşündürmektedir. Bu durum, bilişsel control ve bellek etkileşimlerine dair anlayışı derinleştirmekte ve ÇB ile ilişkili bozukluklara yönelik müdahaleler için bireysel farklılıkların

araştırılması gereğini vurgulamaktadır.

Anahtar sözcükler: Çalışan bellek, bilişsel kontrol, çoklu-talep bölgeleri, çoklu görev.

Acknowledgement

First of all, I would like to thank my supervisor, Dr. Ausaf Farooqui. He was always patient and provided immense guidance. I have always found his way of thinking and knowing the way through science admirable. Thank you for sharing your wisdom with me. I also thank Dr. Burcu Ayşen Ürgen and Dr. Buse Merve Ürgen for sparing their time to evaluate my qualifications and thesis.

I would like to thank my labmates, Higher Cognitive Functions Lab members, as well. It was so fun to work with them, and I will never forget how we supported each other through this academic journey. Especially, my lab bestie Ayşe Betül Varol, thank you so much for making my time in the lab easier. In addition, I would like to thank my friends at the National Magnetic Resonance Research Center (UMRAM). Aside from the science, the most enjoyable part was when we had tea time together, discussing life and everything.

Last but not least, I would like to give a huge thanks to my family. My parents, Nuray Öztürk and Hasan Öztürk, who dedicated and sacrificed everything for my and my sister's education despite all the financial and other issues, I am trying my best to make the best use of their efforts. Thank you so much for being the best parents you could be. My sister, Betül Öztürk, who happened to be living abroad, still felt like she was here with me. You are the best sister I could ever have. Thank you so much for supporting all my decisions, dreams, and delusions.

Finally, I would like to finish my thesis with the verse that has been inspiring me throughout my academic journey:

"And Allah brought you out of the wombs of your mothers while you knew nothing, and gave you hearing, sight, and intellect so perhaps you would be thankful." (Nahl 16:78)

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Imagine a scenario where you are driving to find a place you have never been to. You go to a store and ask the person inside for directions. While returning to your car, you repeat the directions to avoid forgetting them. Now, this is the short-term memory (STM) you are using. Once you get back in your car, you actively try to apply the directions, matching them with the road names, some landmarks, and the right or left turns you should take. This is the working memory, where you retrieve the information and execute the relevant steps based on it [1]. In addition, all these turning points and matching landmarks can be considered as tasks of the main task of driving to a point. They are organized in a nested way to connect the clues and reach a destination one by one. This organization also requires hierarchical thinking and planning, leading the subject to arrange cognitive sources in a controlled way to prevent getting lost. We apply this goal-directed strategy to even the smalest actions in our daily lives. Consequently, it depletes our working memory capacity and gets help from certain areas of the brain.

1.1 Working Memory

Working memory (WM) temporarily maintains the information available when needed [2]. It differs from the classic short-term memory definitions and properties, even though they are closely related. Short-term memory is a passive and temporary information storage for 15-30 seconds [3] with a limited capacity of about 7±2 items [4]. It has an obvious temporal decay; the information can fade away over time [5]. Furthermore, while both systems are related to the prefrontal cortex, WM involves a larger neural network due to its executive functions [6].

This larger neural network also has a close relationship with long-term memory. Long-term memory (LTM) stores information from minutes to a lifetime [7]. It is associated with the structures of the medial temporal lobe and especially the hippocampus [8]. More importantly, the dynamic structure of WM might lead the information to be encoded in LTM through rehearsal and repetition [9].

As seen in the evidence example above, it can be considered a vital property of the human brain. It is directly related to a wide range of cognitive activities such as reasoning, planning, learning, and even executing simple tasks in daily life [10]. For example, patients with impaired working memory due to psychiatric and neurological disorders such as major depression, Parkinson's disease, or stroke showed significant deficits in these areas [11, 12, 13]. It was even considered a memory system for future action plans by Miller et al. (1960) [14]. Considering this widely perspective of WM, it is crucial to define and explore the limits and properties of this ability.

Computer scientists first used the term "working memory" to define the structures within a program that store and use information when required [15]. Miller et al. (1960) [14] proposed it in human research to better explain planned actions. He also raised the properties of working memory as a part of the execution, formulation, and revision of plans [14]. Planning can be considered to have two stages. The first stage is the plan production stage, where the plans and action sequences are generated, stored, and evaluated. The second stage is the execution stage of

the plan, which involves retrieving information from long-term memory, loading the necessary information into working memory, and executing the plan [16]. As in the example above, trying to remember the landmarks on the way, matching the necessary road names, and turning your car left or right accordingly gives the second stage. At the same time, deciding which

Baddeley & Hitch (1974) further improved it by stating that short-term memory (STM) actively controls the information rather than passively storing it [10]. They modeled it as a multicomponent system with two slave and temporary storage systems: *the visuospatial sketchpad* and the *phonological loop*, and a *central executive* to coordinate the activities of the slave systems. The central executive retrieves information from the slave components to conscious awareness, reflects on it, and adjusts it accordingly. In addition to short-term storage, the phonological loop holds speech-based information and plays a key role in long-term phonological learning, such as language acquisition. It is claimed to be associated with Brodmann areas 40 and 44. In addition, the visuospatial sketchpad holds visual, spatial, and kinesthetic information [10]. It is also suggested that it is represented within the right hemisphere [17]. This was the basic model of working memory. Later, they added a third storage component as an *episodic buffer* [17]. It was presented as an interface between two slave systems to enable information flow through episodic long-term memory. It was not modeled as an exact episodic memory because even though it stores episodes of information over time, since it is a temporary storage and an information integrator with limited capacity, it remains a buffer.

1.1.1 Working Memory Capacity

The limited capacity of working memory gives us insight into how the cognitive abilities of humans are also limited, under what conditions they can be altered, or how much they can change over time [18]. When we define capacity, we usually refer to the amount of information that can be remembered in a short time;

however, it can also refer to individual differences in the ability to hold information [18]. Some people can hold around two items, which is low WMC, while others might keep more than four items in mind, which is high WMC [19]. This capacity might be organized by three factors: WM *decay*, which allows only a certain amount of information to be rehearsed before it becomes irretrievable [20, 21]. The second factor is the *resource*, which must be divided between the items available and the processes simultaneously [22]. The resource can be continuous or discrete [23, 24]. Continuous resource hypothesis suggests that the working memory resources are shared equally to each item [24], while discrete-slot models suggest that a certain number of slots (e.g. 3-4) limits the number of items to be maintained in WM, leading to either recalling the stored item in the slots or guessing [25]. Finally, the last factor is the *interference* between these representations that do not decay independently and are not limited by resources [26].

1.1.1.1 Working Memory Decay

Decay theory was first presented by Broadbent (1958) [27], assuming that forgetting occurs as a function of time. It predicts that the amount of information recalled abides by the duration of the item [20]. Baddeley et al. (1975) [20] conducted an STM experiment to investigate whether the immediate memory span is constant or varies with the length of the recalled words. They found that memory span is linked to word length across a vast range of materials, short temporal duration is a better predictor of recalling the words, and the span could be predicted based on the number of words the participants can read in around 2 seconds. They also showed that fast readers were better memorizers. They showed that the ratio of reading rate to time was constant, as the participants could remember as much as they read in 1.8 seconds. Early studies also supported the existence of constant storage time for all types of items, suggesting a time-based capacity [28, 29, 30, 31, 32, 33].

Schweickert & Boruff (1986) [21] took this hypothesis further and conducted

a psychophysics study to investigate the trace threshold model, called the trace-decay hypothesis. This hypothesis assumed that verbal stimuli leave a verbal trace right after being presented and begin to decay. It will be accurate if the participant recalls the item before the trace completely decays. That is why the memory span equals the number of items that can be recalled with a % 50 probability. Moreover, the required time for pronouncing a list exactly one span long equals the time at which the probability of trace decay is one-half. So, they measured the recall probability of a list for different lengths and calculated the actual pronunciation times by reading the lists aloud. Their model assumed that the time it takes to recite the list aloud is less than the variable duration of the trace. As a result, their model accounted for % 95 of the variance in recall performance and the mean trace duration (1.88 seconds). Thus, they supported the phonological loop model of Baddeley & Hitch (1978) [10] and the idea that the memory span is linked to the pronunciation time rather than the number of items.

Although the trace-decay hypothesis could have explained both the origin of the span and why forgetting occurs when the span is exceeded, it has become unpopular [28]. Because research has shown some distortions in remembering, and the importance of the similarity factor in proactive and retroactive inhibition has shown a more dynamic relationship in forgetting, suggesting an interference from similar materials, not only decay (Bartlett, 1932; Koffka, 1935). Memory distortions might be caused by the fact that memory is constructive and inferential [34]. Recall errors and competition among memories are not distinct causes of forgetting but manifestations of deteriorated traces that force the brain to reconstruct memory. The "competition-in-recall" phenomenon can not be underestimated. One memory may inhibit the recall of another, or competing memories can be retrieved simultaneously. However, the subject is unable to differentiate the correct one, leading to a failure of discrimination. These distortions or wrong reconstructions actually imply that the brain actively rebuilds memories instead of simply forgetting or losing them, because recall errors support the idea that forgetting is more about how the memories are accessed or reconstructed. Since the decay hypothesis cannot explain these questions, it seems that the working

memory capacity cannot be explained thoroughly.

1.1.1.2 Working Memory Resource

The resource hypothesis states that the capacity of working memory is limited because cognitive resources such as attention are not infinite and should be divided among all items to be kept in memory [35, 36]. Consequently, trying to store more items in WM leads to having fewer resources for each item. That is why each item is kept with less accuracy or is more susceptible to decay or interference by other items. This view supports the continuous views of WM and opposes the earlier models of a fixed number (four objects) of storage slots [4]. The visual WM studies strongly opposed the slot model. For example, Bays & Husain (2008) [37] tested the participants' propensity to recall the location and orientation of multiple visual items. They showed that recall errors increased as the number of items increased. Similarly, Alvarez & Cavanagh (2004) [38] also investigated the storage of different visual stimuli. They found that visual WM capacity differed for various types of stimuli. They also stated a relationship between the information load and the number of objects that can be stored, meaning that more capacity is allocated to more complex stimuli. These studies disproved the idea of a slot-wise WM capacity and presented a dynamic resource-based WM capacity.

However, more recent studies have criticized this hypothesis for its inconsistencies and limitations. First, the theory does not clearly explain whether these resources are neural activation, attention, or bandwidth. There is no exact measurable unit of resource, which makes falsifying or testing the hypothesis difficult [22]. Moreover, it tends to ignore the performance differences that come with different task demands or item types. For example, facial or verbal information tends to be remembered better even under high-load conditions. This suggests that WMC might not only have a general resource pool but also a modality-specific or content-dependent one [39]. Furthermore, under particular instructions, participants can balance the quantity of items and the quality of remembrance to hold fewer items more accurately [40], suggesting a strategy, flexibility, and a higher level of control rather than a simple resource division.

Furthermore, the hypothesis considers WM as an isolated system rather than fully emphasizing its interactions with other memory systems. However, previous research has shown that LTM influences WM in terms of recalling. For instance, familiar items are remembered better than unfamiliar items, and a resource allocation idea cannot explain this [41]. That is why the recall hypothesis of WMC also cannot be considered reliable, which led the researchers to the interference hypothesis.

1.1.1.3 Interfering with Working Memory

Nairne (1990) [42] presented the interference model of WMC to highlight the importance of possible errors in the immediate recall of similar items. It was initially hypothesized as a feature model to explain the modality-dependent features of WMC. Oberauer et al. (2012) [43] later extended it as a computational model for complex-span working memory tasks. They suggested that the simultaneous processing of information interferes with memory through distractors. More specifically, they assumed that the distractors are linked to the immediately prior item in order to replace and configure them by being encoded into working memory (WM). When the distractor interferes with recall, it tends to replace the last item rather than another item from the list [43]. This assumption was supported by other studies, which showed that when distractor words were spoken aloud, they were unintentionally encoded into working memory (WM) and even required attention, to the degree that they attracted attention [44, 45].

Another assumption is that its link to the last item is removed after the distractor is encoded. The classical view is that once the item is recalled, the brain suppresses it to avoid repeating [46]. As evidence, previous studies suggested that participants were less likely to make repetition errors [47, 48, 49]. The new model extends this idea and says that distractors are also suppressed once they are irrelevant. Hasher et al. (1999) [50] actually called this phenomenon 'deletion' and assumed that it is an inhibitory function of the human brain that decreases over aging, leading to an inefficient WM over time [51]. This inhibitory function is also evident in several studies in which participants encoded two sets of items,

some of which were irrelevant for the next task. Not surprisingly, the number of irrelevant items affected the accuracy. However, this effect faded within 1-2 seconds [52, 53, 54].

Interference can be divided into proactive interference (PI) and retroactive interference (RI). Proactive interference can be exemplified as follows. The participants hear a list of numbers 3, 9, 2, 6, 8... and then another list 1, 4, 7, 5, 9... and then need to decide if the number '2' is in the second list. The first thing they can do is to decide that '2' is not on the latest list, which means that other items interfere with memory, a common function of WM [55]. Meanwhile, retroactive interference is the disruption of the retrieval of prior items by newly acquired information. This kind of interference is suggested not to occur due to weakening of memory traces but due to competition at the time of retrieval [56]. It can also be explained as contextual overlapping cues that make it more difficult to distinguish between items from different lists [57].

In fact, before this computational model of interference, studies had already supported and suggested the interference hypothesis as a WMC function. Previous studies have revealed that capacity limitations are not observed when interference is minimized [58]. Moreover, when the objects were implanted in a naturalistic context, the participants had a greater recall capacity, most likely because the existence of a context decreased the interference by helping memorization [59, 60]. Furthermore, overcoming interference with WM tasks seems to predict intelligence measurements [61, 62]. All these arguments make the interference hypothesis a more reliable and tenable explanation of why the working memory capacity is limited. That is why the interference hypothesis seems to explain the limited capacity of WM more reliably and tenably. However, how the working memory resources are being controlled still needs to be investigated.

1.1.2 Controlling the Working Memory

The main element of controlling working memory resources and capacity passes through cognitive control. Cognitive control is the management of goal-directed

behaviors and mental processes to adapt to new, less familiar situations or decisions of competing actions [63]. It was first conceptualized by Hammond & Summers (1972) [64]. According to their theory, cognitive tasks include two processes: knowledge acquisition and cognitive control over acquired knowledge. They also proposed that feedback about the cognitive material is essential to maintaining control.

1.1.2.1 Attentional Control

Attentional control is the ability to inhibit interference and irrelevant responses while maintaining task-relevant information [65]. It encompasses the behavioral response inhibition, sustained attention, and cognitive inhibition functions [66]. This refers to WMC's ability to use attention to choose information relevant to the goal [67]. Research has been shifted to the perspective that WMC is not just about the amount of information held but how effectively attention can be allocated and protected from interference [68].

Engle et al. (1999) [69] first proposed it to demonstrate that high-WMC people perform better in tasks that require cognitive control because they can better manage interference and resist distraction. Therefore, the differences in WMC stem from variations in executive attention rather than storage limitations. Several empirical studies support this idea. For example, complex span experiments have shown that individuals with high working memory capacity (WMC) are better at multitasking due to their ability to allocate attention effectively. In contrast, individuals with low working memory capacity (WMC) made more errors in task performance and memory retention due to their poor handling of competing demands from the environment [67, 70]. This view also coincides with Baddeley's (1986) central executive model.

When it comes to the executive relation of attention with working memory, there are three aspects to be considered: attention as information storage and processing [71], perceptual attention with memory maintenance [72], and control of attention [73].

a) Attention as Storage and Processing

Attention is considered a limited resource for maintaining representations in an available state and is assumed to be divided between storage and processing [74]. Most of the assumptions of this view align with the idea that there is a single attention resource, and storage and processing compete for it. This assumption is supported by several studies that claim that WMC is limited by an attentional source [75, 76] and storage and processing compete for central processing capacity [77]. However, the dual-task cost is not observed when there is a gap of interval for a few seconds between the memory's encoding and the task's start [78]. This is an important element in terms of the validity of the hypothesis because their initial prediction was that having a dual task cost when the storage and processing demands of working memory (WM) are combined [79]. To further support this, Watanabe & Funahashi (2014) [80] recorded neural activity in the monkey lateral prefrontal cortex (LPFC) when they performed a spatial attention and WM task separately and simultaneously to reveal the neural basis of resource sharing. They found that the same LPFC neurons were active in both tasks, leading to spatial selectivity. When both tasks were performed together, those neurons lost their spatial selectivity and reduced activation to the attended and remembered locations. After the attention task was performed, the WM-evoked activity was observed again, and the same level of activation was observed under the single-task condition. However, examining the behavioral data revealed only a small dual-task cost.

b) Perceptual Attention and Memory

Another aspect is whether the attentional resource is shared between perceived and non-perceived objects. Most studies indicate that WM retention occurs without perceptual information [81]. However, studies have shown that the limited capacity of working memory (WM) also applies to the information in view. For example, Tsubomi et al. (2013) [82] found that the contralateral delay activity (CDA), a neural marker of the number of items a person stores in visual working memory, had the same amplitude up to 3 items, regardless of whether participants actively saw color patches or remembered them after the input disappeared. This

study demonstrates that the capacity limits for attending to perceptual stimuli and maintaining that stimulus in memory are similar.

To test the validity of this prediction, the researchers asked the following question: If perceptual attention and working memory (WM) share the same limited resource, when the load of one task increases, the performance of the other task should decrease. Fougnie & Marois (2006) [83] found that, although load-dependent interference was observed, the cost of dual tasks was smaller when two visual WM tasks were performed simultaneously compared to the combination of visual WM and visual attention tasks. On the other hand, Souza & Oberauer (2017) [84] found a minor interference when WM and perceptual attention tasks were performed simultaneously, suggesting separate resources. Moreover, interference was highest when the working memory task involved spatial locations [85]. That's why controlling the attention is essential for minimizing interference.

c) Control of Attention

Attention can be divided into two types: automatic attention and controlled attention. Automatic attention is triggered by reflexes and is considered not to consume resources; for example, reacting to a sudden noise. In contrast, controlled attention is driven by goals and is considered to consume resources, since it requires top-down processes, such as voluntarily searching for a target. The main question here is whether the act of attending consumes resources or the control of attention imposes a limit. Shiffrin & Schneider (1977) [86] explain why this debate exists by stating that controlling the processes that require continuous attention, and attention cannot be allocated to more than one process at the same time. The limitation lies not in the number of items that need to be attended to, but in how many different actions we can attend to simultaneously to control them and how attention is directed [87]. In this case, WM performance depends not only on storage capacity but also on one's ability to allocate attentional control.

1.1.2.2 Hierarchical Control

Cognitive control is known to have a hierarchical architecture, which refers to the selection of actions in an ordered way [88]. For example, a chess player who decides which piece to move requires abstract considerations, such as the rules of the game, the player's previous experience, and the current state of the game [89]. These goal-oriented planning and decision-making stages occur in a hierarchical manner.

Goal-directed behavior consists of extended actions and task entities such as task episodes, sub-episodes, or tasks. As exemplified above, moving a chess piece requires subsequent cognitive and motor action steps [90]. This organization of actions is called hierarchical behavior, in which subordinate components are formed to achieve superordinate goals. Another example is cooking a meal, which involves a higher-level or superordinate goal ("prepare dinner") and multiple subgoals ("boil water," "chop vegetables," etc.). When someone performs these steps, the overall goal is in the back of their mind. This temporally abstract goal facilitates internal monitoring of each subgoal to follow the sequential plan, ensuring that the time to match the relative stimulus and action is not missed. That is why episodic control is required for sequential tasks, so that a temporal epoch (episode) serves as a control signal [91]. Because the subgoal properties may also be represented as multi-level abstraction. For example, the subgoal of dividing the dishes normally would be considered as one task. However, plates would be classified on different boards according to their function, such as for special occasions and daily use. Likewise, cups would be categorized, including coffee cups, teacups, and others. This type of organization is referred to as a nested hierarchical organization [92]. These types of nested organizations will likely be presented across other networks for the stimulus components at each level of task control. These high abstractions refer to temporally extended task goals [91].

Schneider & Logan (2006) [93] examined the relationship between task-switching sequence and task-level processing with explicit and memorized task sequences. The experiments provided valuable insight into how WM manages

these task hierarchies. They showed that performance declines when the hierarchy becomes more complex or when the WM capacity is burdened. They also argued that working memory is crucial for representing and maintaining subgoals, ensuring that behavior is guided coherently and organized. Frankly, they suggested that WM keep the information of the current goal and the subgoals it is related to. For example, when an unexpected event occurs, WM helps to allocate attention from a lower-level task to a higher-level goal to reassess the plan. These results point to a flexible switch between different levels of the task hierarchy and the crucial role of WM load in task management.

Their results also support Logan & Gordon's (2001) [94] executive control theory. They assumed that "the task-level representation is hierarchical, with the higher level specifying the order in which the tasks occur and the lower level specifying each task separately" (p. 396). More openly, the theory represents two levels of a task set: 1) a task level and 2) a parameter level (a set of control parameters). These task sets are executed and modified by sending control parameters to lower-level or subordinate processing mechanisms. Frankly, this theory posits that WM stores goal information and regularly revises it in response to feedback and task demands. This enables individuals to manage complex, multi-step tasks without losing sight of the primary goal.

More importantly, previous studies have revealed that tasks are executed as a single entity to complete a task. For example, Dezfouli & Balleine (2012) [95] found that when two actions are combined to make an action sequence, the second action will be executed regardless of the result of the first action. These combined actions are referred to as 'task episodes,' as they are temporally extended to form a complete task. Farooqui & Manly (2017) [96] also proposed that task episodes are organized by cognitive entities as a sequence of behaviors for a coherent whole, facilitating control and monitoring of goal-directed behavior over an extended period of time.

Earlier investigations have revealed that hierarchical behavior is executed in the same manner as information is processed within working memory (WM) [93]. The behavioral and neural markers of working memory (WM) processes align

with the boundaries and switches of the task hierarchies. Reaction time and accuracy tend to fluctuate around the switches between task levels, and neuroimaging studies showed distinct neural signatures associated with the beginning and end of task episodes [90]. They found that when participants transitioned from one task sub-goal to another, they experienced a 'reconfiguration cost,' resulting in slower reaction times or temporary performance dips. Furthermore, this cost has changed by the task complexity WMC to buffer against the cognitive demands imposed by hierarchical transitions. It is also suggested that we divide extended tasks into shorter tasks, utilizing the WMC, to control them better and avoid overloading. However, a previous study [97] presented participants with demanding but independent trials, where they did not hold information in the WM, even though the trials were long sequences. Instead, the participants still divided the task into smaller units.

Several neuroimaging studies have also investigated these issues. Some studies [98] examined hierarchical cognition as the organization of tasks with multi-step sequences and nested goals. Their study showed that when participants performed multistep tasks with subgoals, several brain regions were activated to varying levels within the hierarchy. For example, the anterior prefrontal cortex (aPFC) and rostral lateral prefrontal cortex (rLPFC) appeared to encode higher-level goals, while the posterior regions were activated for lower-level steps. Badre (2008) [99] also detected a similar rostral-caudal organization. Thus, this literature implies the importance of a domain-general cognitive control network in the brain.

1.2 Multiple-Demand Regions

Over the functional neuroimaging history, researchers discovered a typical pattern of activity as a response to different types of cognitive challenges [100, 101]. Duncan & Owen (2000) [100], in their meta-analysis of fMRI studies, clustered activated voxels from fMRI studies that manipulate perceptual difficulty, working memory, attention, task novelty, and response inhibition and revealed a common

frontal lobe network, leading to this idea: Regardless of the notion of the task, during the demanding tasks, a specialized set of regions is activated [102]. This task-evoked activation pattern is called the Multiple-Demand (MD) system and encompasses specific areas in the prefrontal and parietal cortex, such as the inferior frontal sulcus (IFS), anterior insula, frontal operculum, presupplementary motor area (pre-SMA), anterior dorsal cingulate, and intraparietal sulcus (IPS) [102].

This question might arise: Why does cognitive demand consistently activate these regions? To answer this, we should inspect the executive processes during a cognitively demanding task. When a task has a high cognitive load, the demand for coordination, regulation, and management of mental processes such as attention, working memory, and decision-making also increases. These functions fall under the umbrella of *cognitive control*, which refers to the brain's ability to regulate and control lower-level sensory, memory, and motor operations according to goals and plans [103, 104]. It is especially crucial in new or non-routine situations, where we must inhibit familiar behavior and execute less familiar behavior [105]. In that sense, the MD network is the core substrate for these control processes, dynamically adapting to various task demands to support flexible and goal-directed cognition.

The adaptability of the MD system comes from its domain-general nature, which means that it is not specialized for a specific type of information. Instead, it responds to the act of "working" against challenges. Importantly, evidence from neuroimaging studies suggests that the activity of the MD regions scales with task difficulty or cognitive load. For example, Wen & Egner (2022) [106] manipulated difficulty to be low, medium, and high under easy and hard conditions. They matched the highest difficulty level in the easy condition with the lowest level in the hard condition. They found that MD activation was independent of context, meaning activation increased with the absolute difficulty of the task and did not rescale between easy and hard conditions. Activation during the highest level of the easy condition was equal to the lowest level of the hard condition. Similarly, strong MD activation during more difficult tasks reflects the increased demand to integrate the components of cognitive operations 4 to solve the task at hand

[107, 108]. These results highlight the notion that the MD network only alters its activation to meet the task's requirements.

Such management of cognitive sources and processes requires cognitive control as well. The MD network is primarily associated with hierarchical control, as it biases processing in other brain areas [109, 110]. Various studies have already revealed the association between hierarchical task control and the hierarchical organization of the prefrontal cortex (PFC) [88, 111]. When a task step is completed, we know it triggers the elimination of the previous control representation and the instantiation of the subsequent one [112]. In terms of nested hierarchy and task-task relationships, researchers [98] investigated how different levels of task and task completion are represented in the frontoparietal network (FPN). In the experiment, participants completed a visual target detection task and found that the activity in FPN regions varied at the hierarchical level of the task event. The completion status drove the activation. It was at its highest when the entire trial was completed, intermediate when a task was completed, and lowest during the task itself. This leads to the idea that the completion of the task serves as feedback and updates the local control networks, whereas the completion of the full task activates the whole control network for the task episode.

In addition, Farooqui & Manly (2018) [113] found that the frontoparietal regions were deactivated even though the participants viewed and attended to meaningful stimuli. These deactivations were strongest when the stimulus was presented coherently and did not demand cognitive control. These results led us to the idea that deactivation does not necessarily mean a lack of attention, and frontoparietal activation requires complete cognitive control, not just perception or attention, suggesting that deactivation points to reduced control needs. Furthermore, they [114] also demonstrated that the cognitive structure of task episodes is associated with neural markers. At the beginning of the task episodes, there are deactivations in frontoparietal networks and the Default Mode Network. They are assumed to point to the initial instantiation of a cognitive entity of the incoming task episode as a consequence of hierarchical cognition.

As mentioned above, the MD network was responsive to various types of tasks,

including working memory. However, their relationship is not based only on cognitive control. For example, as part of the MD regions, the dorsolateral prefrontal cortex (dlPFC) plays a crucial role in maintaining and manipulating information, whereas the anterior cingulate cortex (ACC) facilitates the monitoring and allocation of cognitive resources in response to task demands [115]. Additionally, PPC has a role in the allocation of attention [116]. Functional connectivity studies also highlight the relationship between the increased co-activation of the dlPFC and PCC with higher working memory (WM) capacity, as well as the co-activation of the dlPFC and ACC under high cognitive load of WM [117, 118]. Therefore, every sub-region of the MD network has a distinct connection and function with each other to maintain control as a whole.

Furthermore, clinical neuroimaging studies support the link between these areas and WM. They indicated that lesions in the dlPFC, its connected areas, such as the medial parietal cortex, and associated white matter tracts were related to specific deficits in working memory functions [119, 120]. Another example is that disruption of frontoparietal connectivity results in schizophrenia and ADHD, where WM deficits are well known [121].

In this regard, we can say that the MD network facilitates working memory. Murray et al. (2017) [122] proposed a mathematical model combined with neural recordings, suggesting that the sustained activity of the MD network and its subnetworks plays a key role in maintaining information in working memory. In addition, increasing MD network connectivity through cognitive training also improves WM in patients with ADHD and schizophrenia [123]. Furthermore, individual differences in working memory are also related to MD network activations [124]. Assem et al. (2020) [124] found that increased activity in these regions is parallel to increased performance of WM tasks.

In general, the MD network is known to be sensitive to cognitive demand. In contrast, another domain-general brain network was found to be mostly negatively correlated with MD regions, as is known to be responsive to the resting state [125].

1.3 Default Mode Network (DMN)

Another common activity pattern discovered in the brain is the Default Mode Network (DMN). This network was first introduced by Raichle et al. (2001) [125] through their resting-state MRI studies. She identified that a specific set of brain regions showed increased activity during the resting state and decreased activity during goal-oriented tasks. These regions involve the medial prefrontal cortex (mPFC), posterior cingulate cortex (PCC), inferior parietal lobule (IPL), lateral temporal cortex, and hippocampal formation [126]. According to common knowledge, we assume that the brain is generally more active when it is working. However, the DMN is more active when participants focus on their internal world, such as daydreaming, mind-wandering, autobiographical memory, and the future [127]. These groundbreaking studies revealed a key idea: There is a brain network that enables the brain to rest by activating more. Since then, DMN research has been used to understand typical and atypical brain functioning, including disorders such as Alzheimer's disease, depression, and schizophrenia [128, 129].

So, it is safe to say that the DMN plays a crucial role in various internally directed cognitive functions. For example, it has a key function in autobiographical memory retrieval to reflect on past experiences [130]; future thinking to simulate future scenarios in mind to help create a coherent sense of self over time [131]. Another function of DMN is the theory of mind processes to understand others' perspectives and mental states, constructing a social-cognitive framework in complex human interaction [132]. Furthermore, it also contributes to semantic processing, moral reasoning, and emotional introspection [133]. These results showed that the DMN is not just a passive background state but a system that supports high-level mental simulations.

More importantly, it is also related to cognitive control and cognitive load. Neuroimaging studies have indicated that during sustained attention or executive function tasks, such as working memory, inhibition, or task switching, the DMN is consistently deactivated [134]. This task-induced deactivation suggests that the DMN response competes with goal-directed control systems, and its deactivation

is essential for task performance to the point that insufficient suppression of DMN during demanding tasks can negatively affect performance and lead to distraction, such as mind wandering and attention deficits [135]. Furthermore, an increase in cognitive load is associated with a greater suppression of DMN. For example, higher demands on working memory lead to more substantial deactivation of the DMN, supporting the idea of distributing neural resources to focused control networks such as the frontoparietal network (FPN) [136].

The DMN and MD systems are typically negatively correlated; when the MD system exhibits task-evoked activation, the DMN displays task-induced deactivation, and such [137]. In that way, they balance internal management and external task engagement. However, recent studies also suggest the possibility of cooperation between these networks in complex planning, creative problem solving, and moral decision-making processes [138].

However, Piccoli et al. (2015) [139] demonstrated that the DMN and working memory networks are not uniformly anti-correlated across all task phases; instead, their interaction can vary dynamically depending on the phases of memory, such as encoding, maintenance, and retrieval. It may also be reactivated during the retrieval phase to aid in reconstructing information [140]. Some studies have also shown that the DMN can exhibit task-positive activation under specific working memory (WM) conditions. For instance, the DMN may support working memory (WM) when maintaining personally relevant information or when integrating long-term memory with WM demands [141]. Moreover, it has been shown that DMN can also co-activate with frontoparietal control networks to demanding tasks as well [142]. Even better, this co-activation may predict working memory (WM) performance more accurately when the task requires internally guided retrieval [143]. For instance, Seoane et al. (2024) [144] emphasized the importance of the DMN in integrating self-referential processing and constructing mental scenes, which are essential for effective memory encoding and retrieval. This view suggests that the DMN interacts with working memory networks by supporting mechanisms essential for organizing and maintaining contextual information.

1.4 Current Thesis

Building on the existing literature, this study aims to clarify how the Multiple Demand (MD) network supports working memory (WM) under varying load conditions during multitasking, and whether it tracks the precision of recall. Although traditional fMRI studies have emphasized load-dependent increases in MD activation, it remains unclear whether these regions are sensitive not only to task demands but also to recall fidelity, measured as trial-by-trial angular deviation. This distinction bears directly on theoretical models of WM capacity, particularly the contrast between continuous-resource and discrete-slot accounts.

Two hypotheses were tested:

1. If the WM capacity is unitary, the working memory load increases in one task should interfere with performance in the other. However, if WM operates as a hierarchical system, then tasks may be prioritized or strategically separated and lead to task-specific effects.

2. If MD regions represent cognitive load in discrete steps (e.g., high vs. low), rather than tracking the fidelity of recall, then BOLD activity should vary categorically with load but not scale with recall precision (angular deviation).

To test these hypotheses, the current study compared encoding and recall activity across two tasks and load conditions, and applied a parametric modulation analysis to assess whether accuracy-related variance explains MD activation.

Chapter 2

Methods

2.1 Participants

The study consisted of 30 healthy and right-handed participants (20 F, M = 20.6, SD = 1.1, range = 18-23). 29 of them had an undergraduate degree, and one of them had a PhD degree. They had normal or corrected-to-normal eyesight.

Ethical approval was obtained from the Research Ethics Committee of Bilkent University, and all participants gave their informed consent prior to participation.

2.2 MRI Acquisition

MR images were collected at the National Magnetic Resonance Research Center (UMRAM), Bilkent University, using a 3 Tesla MR scanner (Magnetom Trio, Siemens AG, Erlangen, Germany) with a 32-channel array head coil. Functional images were acquired with a T2*-weighted gradient-recalled echo-planar imaging (EPI) sequence. The parameters are as follows: repetition time (TR) of 2 s, echo time (TE) of 30 ms, 32 oblique slices (slice thickness: 3 mm, interslice gap: 0.75 mm), in-plane resolution of 3.0 $\times$3.0 mm^2, matrix dimensions of 64 $\times$ 64, a

field of view measuring 192 mm, and a flip angle of 78 °. A T1 *-weighted 3D anatomical sequence collected structural data with the following parameters: 1 mm slice thickness, resolution of 1.0 ×1.0 × 1.0 mm^3 isometric voxels, a field of view spanning 256 mm, and 176 slices. GRAPPA and phase oversampling were used at a % 20 with an acceleration factor of 2.

2.3 Procedure and Experimental Design

The experiment (see Figure 2.1) was programmed with PsychoPy (version 2021.2.3 [145]). The stimuli were shown on a gray background. The experiment consisted of two main tasks, Task A and Task B, and was carried out under low and high-load conditions. Throughout the study, participants completed 40 episodes: 20 low-load and 20 high-load. These were interleaved, and their order was reset after every four episodes. The participants underwent two fMRI runs, each consisting of 10 low-load and 10 high-load episodes, lasting between 15 and 25 minutes. The stimuli were presented on an MR-safe 32-inch LCD monitor (refresh rate: 60 Hz, resolution: 1600 × 1200 pixels; T-32, Troyka Med A.S., Ankara, Türkiye). The monitor was placed near the rear of the scanner bore and viewed from a distance of 165 cm using a mirror mounted on the head coil. The responses were recorded using an MR-safe button box.

The main stimuli were colored lines. Each line was approximately 0.15 units long, and the width was 2.5 units (visual angle determined by screen and MRI setup). Orientations were chosen pseudo-randomly with the rule that a line and its opposite (for example, 45° and 135°) were not shown in the same trial. Lines could appear in up to four different quadrants on the screen: right, top, left, and bottom, centered at offsets of ±0.3. In task A, four main colors were used: red, green, brown, and yellow, and for task B, the lines were cream, yellow, blue, and white. Throughout the experiment, the colors for each task never mixed. Fixation crosses were used to signal incoming recall of the task: green crosses meant Task A and blue crosses meant Task B. Fixation durations were jittered (0.5–6.0 s) to decorrelate events in the fMRI signal. After each stimulus display,

a mask was placed on the screen to prevent after-images.

Each episode followed a four-step extended oblique lines design. In Step 1 (Task A), participants were presented with one line in low-load episodes or two lines in high-load episodes at the center of the screen. When two lines were presented, they were always in different colors. Participants were asked to memorize their orientations. The lines remained visible until the participants pressed a button, after which a mask was shown for 500 milliseconds. In Step 2 (Task B), two additional lines were presented, and the participants had to memorize their orientations again. The lines remained on the screen until the participants responded, and a mask followed for 500 milliseconds. The number of lines in task B did not change and was always set to two. In Step 3, participants were probed to recall the orientation of one of the task B lines. The probe was presented in the exact location and with the same color as the original stimulus. The participants rotated the line using the button box. The index finger turned clockwise and the middle finger anticlockwise, with each press changing the angle by 1°. A separate button confirmed their response. Afterward, they saw feedback showing the angular deviation between their response and the correct orientation. In Step 4, participants were asked to recall the orientation of one of the task A lines from Step 1. The probe matched the original in color and location, and participants adjusted the orientation using the same controls as in Step 3. After confirming their answer, the feedback was again shown, including the angular deviation.

Between each step, a jittered interval of 0.5 to 6.0 seconds was presented, calculated from the offset of the previous step to the onset of the next. These jitters helped ensure that brain activity from one step did not overlap too much with the next in the fMRI signal and accounted for the slice timing variability. After the feedback from Step 4, the participants went to the start screen for the next episode.

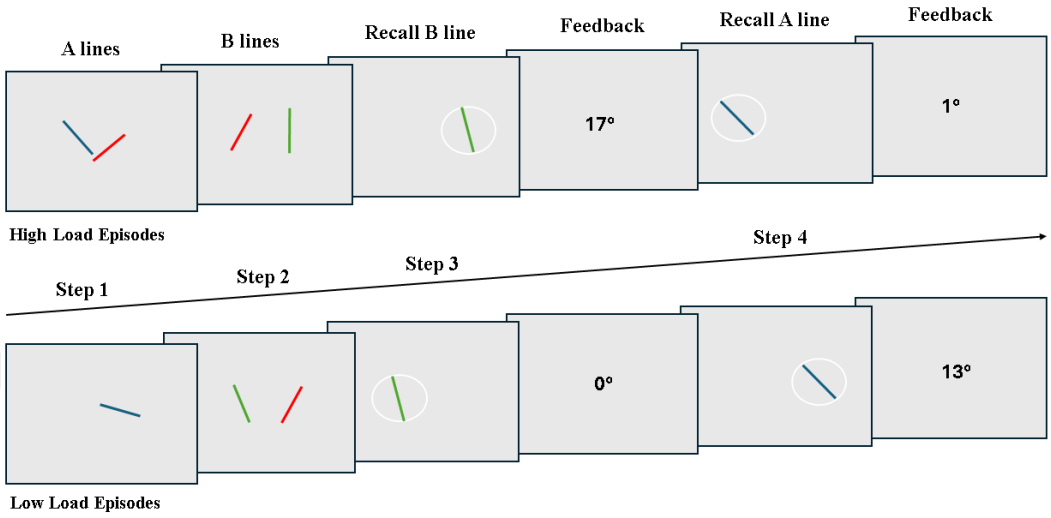

Figure 2.1: Experimental Design. Participants performed an interleaved dual-task WM paradigm with high- and low-load conditions. In task A (Step 1), they encoded 1 (low-load) or 2 (high-load) colored oblique lines. After a mask, task B (Step 2) required encoding 2 additional lines. Recall was probed for task B (Step 3) and then task A (Step 4), with responses made via button presses (1° adjustments per press). Feedback showed angular deviation. Each trial included jittered inter-step intervals (500–600 ms). The experiment comprised 2 fMRI runs, with stimuli displayed on an MR-safe monitor and responses recorded via button box.

Behavioral data included stimulus orientations, participant responses, reaction times, and deviation scores.

2.4 Analyses

Prior to whole brain analyses, preprocessing was performed using Statistical Parameter Mapping (SPM12) with an automatic analysis pipeline (aa, version 5

[146]) in MATLAB (version R2022b, [147]). As part of the preprocessing, functional images were slice-time corrected, using the middle slice as a reference, and co-registered with the structural T1-weighted images. The realignment, including the motion correction during the scanning and normalization steps, was applied according to the Montreal Neurological Institute (MNI) template. EPI images were resampled at a resolution of 2 x 2 x 2 mm^3 and smoothed using an 8 mm full-width at half-maximum Gaussian kernel. The dataset was subjected to high-pass filtering with a cut-off frequency of 128 seconds.

Whole-brain analysis was modeled by the general linear model (GLM). Different stages of the trial, such as the onset of every regressor, intervals, and their durations, were modelled using epoch regressors that corresponded to their duration. Ten regressors were modelled: task A & task B in the recall and presentation steps under high and low load, in addition to parametric values of the two recall regressors under high load. The duration of the presentation was calculated from the onset of the first line's presentation until the end of the last trial that participants finished to their satisfaction. Likewise, the duration of the recall steps was calculated from the onset of the first probed lines until the end of the last answer to the probed lines (see Figure 2.1). These regressors were convolved with a basis function representing the canonical hemodynamic response.

The model included six contrasts:

1. Recall of task A under high load >recall of task A under low load (High A Recall >Low A Recall)
2. Recall of task B under high load >recall of task B under low load (High B Recall >Low B Recall)
3. Presentation of task A under high load >presentation of task A under low load (High A Presentation >Low A Presentation)
4. Presentation of task B under high load >presentation of task B under low load (High B Presentation >Low B Presentation)

5. Parametric recall of task A under high load >all other regressors (High A Recall Parametric >All)

6. Parametric recall of task B under high load >all other regressors (High B Recall >All)

A parametric modulation analysis was performed with the general linear model to examine how trial-by-trial variations in performance accuracy modulate brain activation. This approach validated whether accuracy-related variance systematically influenced BOLD responses during the recall phases of Task A and Task B separately, under high and low load conditions. The objective of setting the performance accuracy as a parametric modulator was to identify brain regions within the MD network that scale their activation by working memory load and retrieval success. This analysis provided a more detailed understanding beyond categorical effects, revealing whether regions involved in cognitive control and working memory show linear sensitivity to accuracy. For this reason, the angular deviation of the participants was set as a parametric modulator. The angle of deviation was decided on the recall steps for each task, where the participants probed the angle of the lines presented previously for the corresponding tasks. They demonstrated how the participants deviated from the original angle (if correct, 0°) of the presented lines, considering the precision of performance. Furthermore, the values of the related regressors were mean-centered, making the parametric modulator orthogonal to the regressors they modulate. This prevents the main regressors from being correlated with the parametric modulator.

Later, to display the activity pattern in the MD regions, ROI analyses were performed. Well-known MD regions were selected to create ROIs for both hemispheres separately: anterior insula, anterior middle frontal gyrus (aMFG), frontal eye fields (FEF), intraparietal sulcus (IPS), medial middle frontal gyrus (mMFG), posterior middle frontal gyrus (pMFG), and pre-supplementary motor area (preSMA). ROIs were obtained using the JuBrain Anatomy Toolbox [148]. The results were used to perform Bayesian repeated measures ANOVA on JASP [149].

Chapter 3

Results

This thesis aimed to investigate the neural correlates of working memory load effects during multitasking in terms of WM capacity and MD response patterns. To test for them, activity across two tasks and load conditions was compared for the encoding and retrieval steps of the interleaving visual working memory task. Data were analyzed through a whole-brain analysis in addition to parametric modulation analyses, ROI analyses, and Bayesian repeated measures ANOVA analyses. The results are presented below.

3.1 Behavioral Results

Participants took longer to press the button under high load, except for retrieving task B (see Figure 3.1). The difference between high-load and low-load conditions appears to be greater for task A line encoding, most likely due to the effect of adjusting the task. Additionally, participants recalled the orientations of task B more accurately than those of task A.

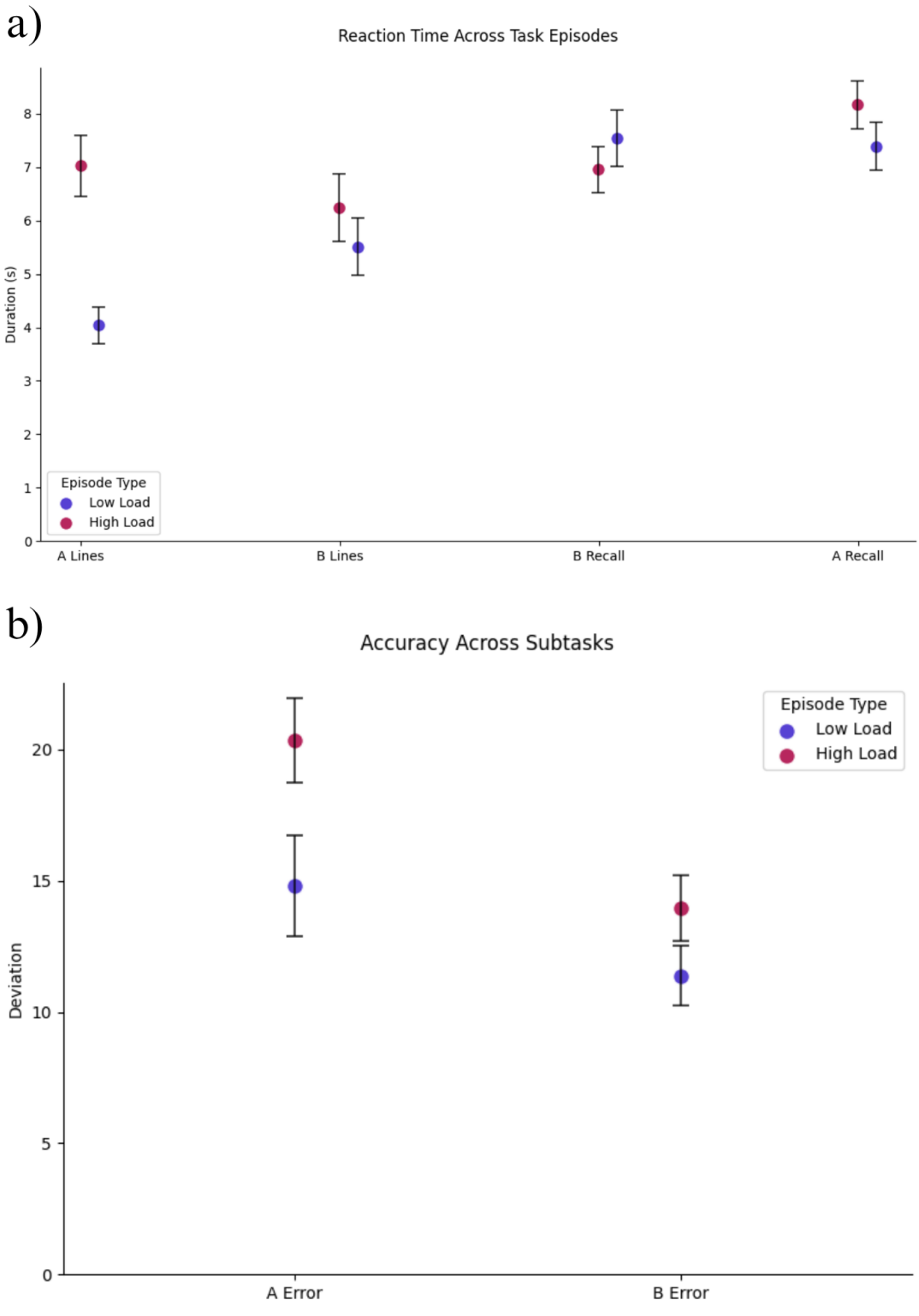

Figure 3.1: Participants' reaction time and accuracy. a) Mean duration of time until participants pressed the button to proceed to the next trial, b) mean angular deviation of line orientation, which participants probed to recall for both tasks.

3.2 Whole Brain Analyses

To investigate the neural correlates of the effect of working memory load on the type of recall and presentation, whole-brain voxel-wise analyses were conducted. The results were filtered at an uncorrected voxel-wise threshold of p <0.001 without a cluster extent threshold ($T = 3.39$, $k = 0$). No clusters survived the FEW or FDR correction. The contrasts were always set as high load >low load for tasks A and B.

3.2.1 Task Presentation (Encoding)

For task A (High Load >Low Load), group-level analysis revealed significant activation across bilateral parietal regions. Specifically, peak activation was observed in the left superior parietal lobule (SPL) and the left intraparietal sulcus (IPS). Additional activation was observed through the right superior parietal cortex and left posterior parietal areas. Furthermore, significant deactivation was observed in the right post-cingulate gyrus (post CG), the bilateral anterior cingulate cortex (ACC), and the right cuneus (see Figure 3.2.a).

On the contrary, for task B (High Load >Low Load), the largest cluster and peak activation were in the bilateral inferior parietal lobule (IPL) and the superior parietal lobule (SPL). The deactivation cluster was very small and localized in the anterior cingulate cortex (ACC) (see Figure 3.2.b).

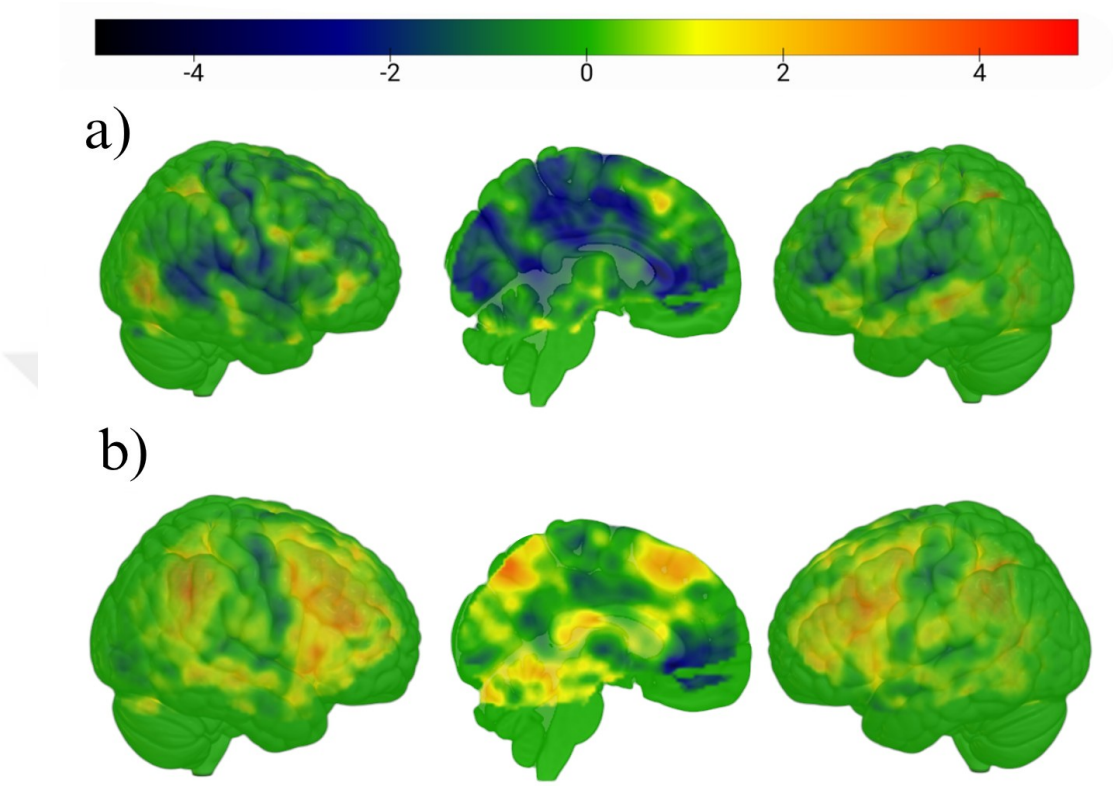

Figure 3.2: Activation and deactivation pattern of working memory load on *a)* task A presentation and *b)* task B presentation masked by the contrast of High Load >Low Load. Positive values indicate regions that were more active, whereas negative values indicate deactivated regions. Significant regions were marked in accordance with the threshold ($T = 3.39$, $p < 0.001$).

Both contrasts were compared to each other to examine the interaction between working memory load (high versus low) and task type (task A versus task B), using the contrast (High Load A - Low Load A) minus (High Load B - Low Load B). The results revealed significant differences ($F = 13.39$ (df = [1, 29]), $p < 0.001$) in the intraparietal sulcus (IPS), superior frontal sulcus (SFS), cuneus, and lateral occipital cortex (LOC) (see Figure 3.3).

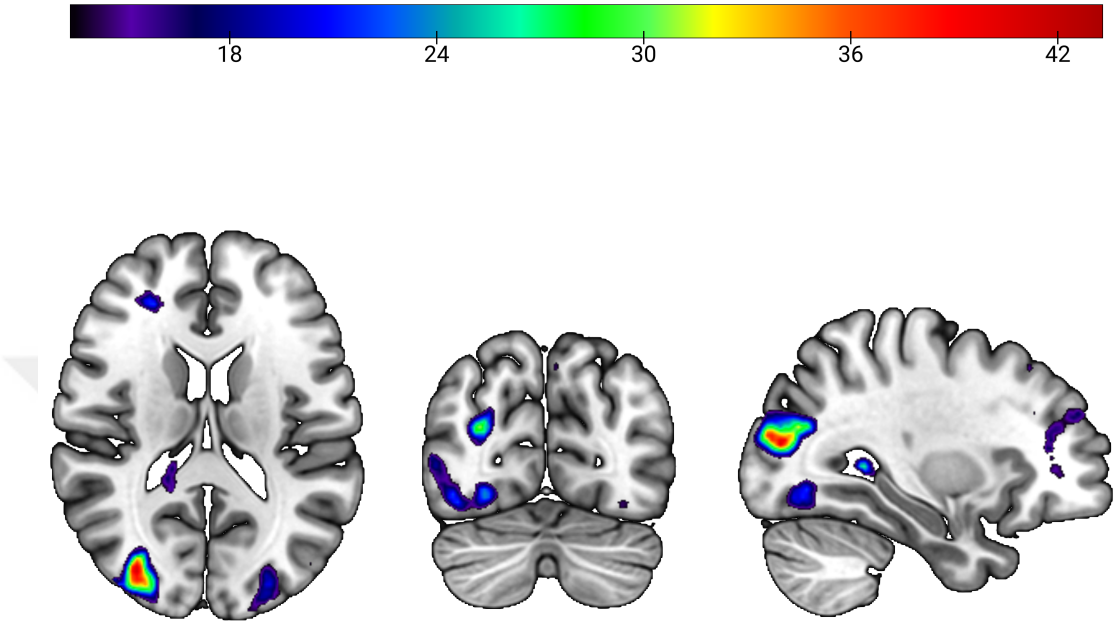

Figure 3.3: Differentiation pattern of load effect on presentation of task A and task B, showing the comparison of task A's high >low load contrast and task B's high >low load contrast (task A >task B). Hot colors indicate stronger activity. Significant regions exhibit the activation above the threshold ($F = 13.39$, $p < 0.001$).

To specify the response pattern in the MD regions with the high-low load contrast, region-of-interest analyses were conducted for both tasks separately. As shown in Figure 3.4, task B exhibited consistently positive beta values across nearly all MD regions, indicating increased activation under high load conditions. In contrast, task A showed a more diverse pattern, with several ROIs displaying negative values. Notably, regions such as the left and right middle frontal gyrus (L-mMFG, R-mMFG) and the left anterior MFG (L-aMFG) exhibited pronounced differences between tasks. task B shows strong load-related activation, and task A shows reduced or negative effects. These results suggest that the MD network is more robustly engaged by high load demands during task B, meaning task-dependent recruitment of cognitive control resources (see Figure 3.4).

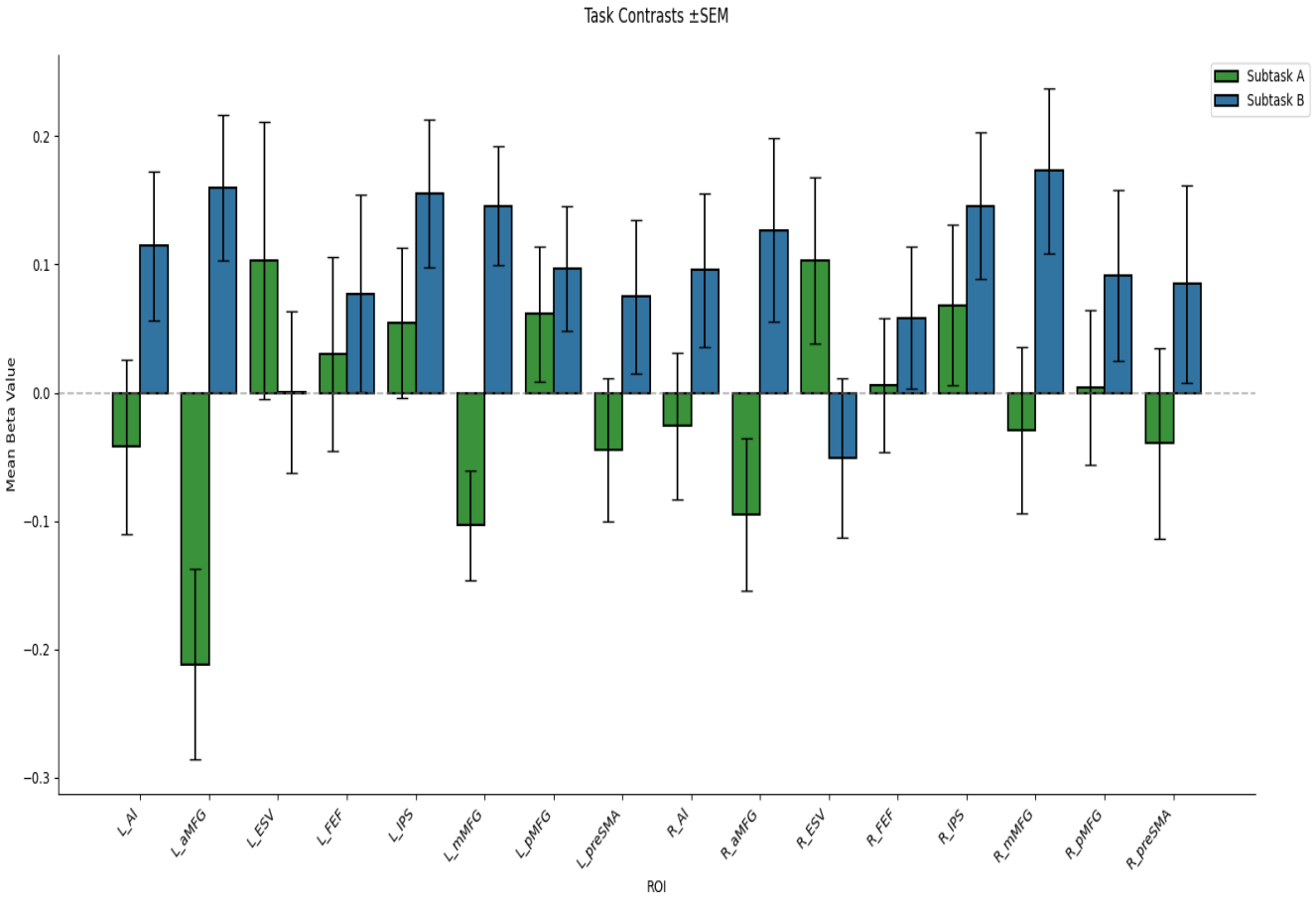

Figure 3.4: Mean beta values across ROIs of the Multiple Demand (MD) network, indicating the contrast of High Load >Low Load during task A and task B presentation steps. Each bar represents the average activation in a given region of interest (ROI), with error bars indicating the standard error of the mean (SEM) for within-subject variability. (L: Left, R: Right, AI: Anterior Insula, preSMA: Pre-supplementary Motor Area, FEF: Frontal Eye Field, pMFG: Posterior Middle Frontal Gyrus, mMFG: Middle Middle Frontal Gyrus, aMFG: Anterior Middle Frontal Gyrus, IPS: Intraparietal Sulcus)

3.2.2 Task Recall (Retrieval)

For the task A recall (High Load >Low Load), small clusters were observed in the precuneus, intraparietal sulcus (IPS), and superior parietal lobule (SPL), whereas significant deactivations were observed in the insula, inferior and superior parietal lobules (IPL/SPL), cingulate gyrus and sulcus, supplementary and presupplementary motor areas (SMA/preSMA), anterior cingulate cortex (ACC), and lateral occipital cortex (LOC) bilaterally (see Figure 3.5.a).

For the task B recall (High Load >Low Load), no clusters survived except a small deactivation in the left V1 and around the post-cingulate gyrus (see Figure 3.5.b).

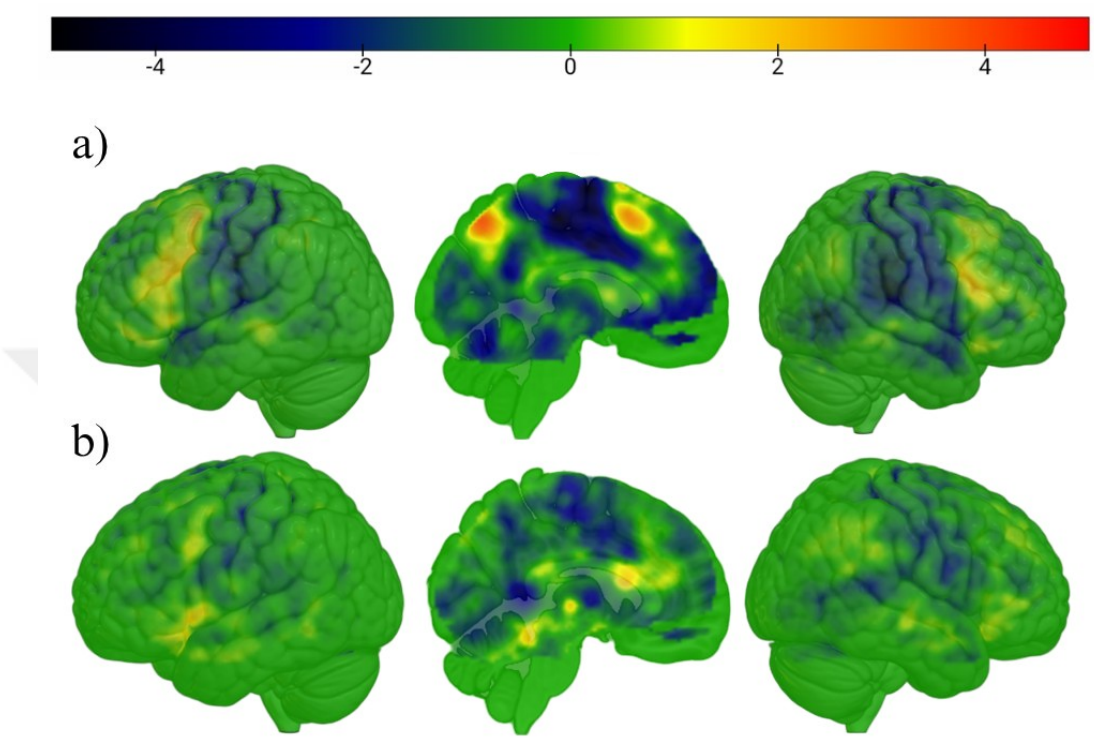

Figure 3.5: Activation and deactivation pattern of working memory load on *a)* task A recall and *b)* task B recall masked by the contrast of High Load >Low Load. Positive values indicate regions that were more active, whereas negative values indicate deactivated regions. Significant regions were marked in accordance with the threshold ($T = 3.39$, $p < 0.001$).

Both contrasts were compared to each other to examine the interaction between working memory load (high versus low) and task type (task A versus task B), using the contrast (High Load A - Low Load A) minus (High Load B - Low Load B). The results revealed significant differences ($F = 13.39$ (df = [1, 29]), $p < 0.001$) in the insula, anterior cingulate cortex (ACC), early visual cortices (V1, V2), supplementary and presupplementary motor areas (SMA/preSMA), superior temporal sulcus (STS), and precentral, central, and postcentral sulci (see Figure 3.6).

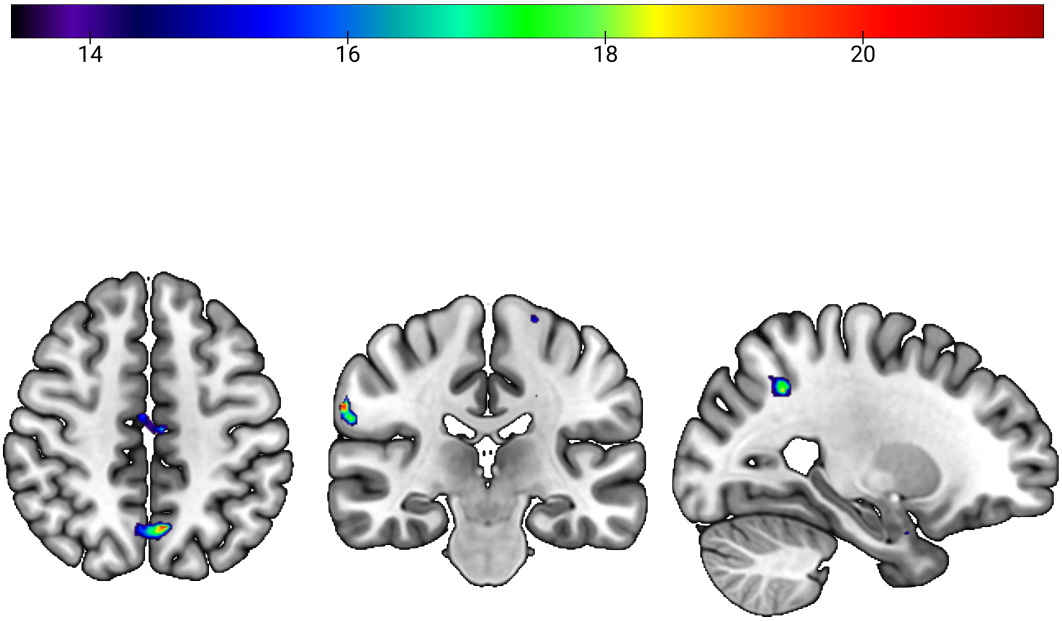

Figure 3.6: Differentiation pattern of load effect on retrieval of task A and task B, showing the comparison of task A's high >low load contrast and task B's high >low load contrast (task A >task B). Hot colors indicate stronger activity. Significant regions exhibit the activation above the threshold ($F = 13.39$, $p < 0.001$).

The contrast between high and low load conditions revealed a different pattern from the encoding phase. As shown in Figure 3.7, beta values across MD regions were generally closer to zero, with both task A and task B exhibiting minimal differential activation. Most ROIs showed small negative values for both tasks, and the directionality of effects was largely inconsistent across regions. For example, the left middle frontal gyrus (L-mMFG) and left posterior MFG (L-pMFG) exhibited slight positive contrasts in task A, whereas the same regions showed negative or near-zero values in task B. Across the board, the error bars were relatively large, overlapping zero in nearly all ROIs, indicating no clear or consistent load-related modulation during recall. These findings suggest that, unlike encoding, the MD network does not robustly differentiate high from low load conditions during recall, and this lack of differentiation is consistent across both tasks.

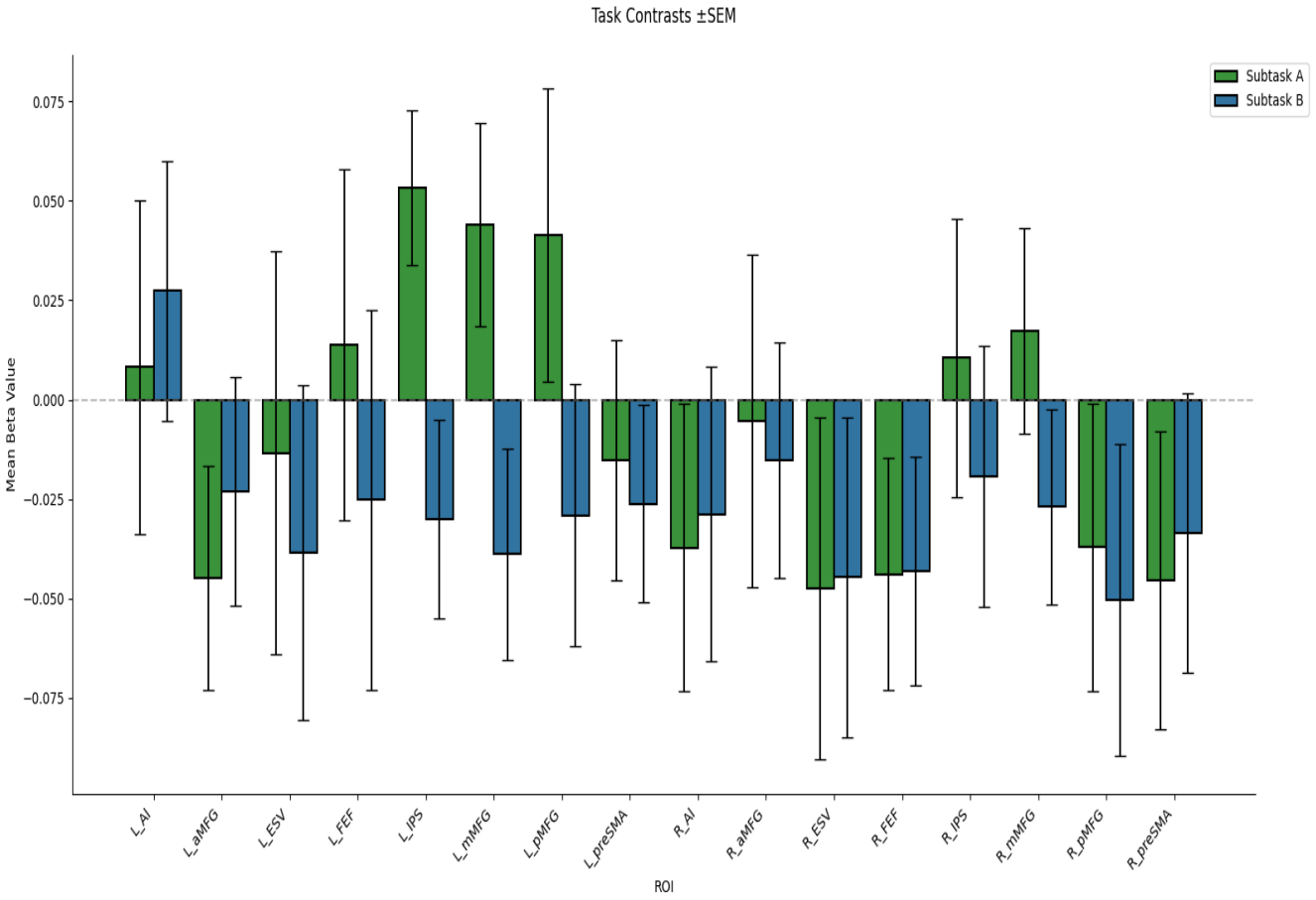

Figure 3.7: Mean beta values across ROIs of the Multiple Demand (MD) network, indicating the contrast of High Load >Low Load during task A and task B retrieval steps. Each bar represents the average activation in a given region of interest (ROI), with error bars indicating the standard error of the mean (SEM) for within-subject variability. (L: Left, R: Right, AI: Anterior Insula, preSMA: Pre-supplementary Motor Area, FEF: Frontal Eye Field, pMFG: Posterior Middle Frontal Gyrus, mMFG: Middle Middle Frontal Gyrus, aMFG: Anterior Middle Frontal Gyrus, IPS: Intraparietal Sulcus)

To test the significance of these responses, A Bayesian repeated measures ANOVA was conducted to examine the effects of task type, working memory load types, and ROI on brain response. The results provided overwhelming evidence for the encoding phase model that includes all main effects and all two- and

three-way interactions of ROI, Task, and Load ($ROI + Task + Load + ROI \times Task + ROI \times Load + Task \times Load + ROI \times Task \times Load$) as the best-fitting model compared to the null model ($P(M \mid \text{data}) = 0.982$, $\text{BF}_{10} = 2.06 \times 10^{55}$). However, for the retrieval phase, the interaction between task and load effect was absent ($BF\_incl = 6.282\times\ 10^{-8}$), and the main effects of the working memory load ($BF\_incl = 2.369\times\ 10^{-11}$) and the type of task (BF_incl = 0.000) were not significant.

Overall, consistent load-related increases in MD activation were observed during encoding for task B, particularly in bilateral frontal regions, while task A showed more variable or even negative responses. However, during recall, this pattern reversed. task A showed slightly more consistent load-related effects across certain regions, particularly in the left hemisphere, whereas task B exhibited minimal and inconsistent modulation. These differences between tasks were more pronounced during the recall phase, suggesting that task A may have imposed greater demands on memory retrieval. Additionally, some right-hemisphere regions, such as pMFG and IPS, showed load-related deactivations, especially during encoding, which may reflect suppression of default-mode network activity under increased task demands.

3.2.3 Do different MD regions predict accuracy for different tasks?

As shown in Figure 3.8, a comparison of the accuracy-correlated brain responses to task A and task B under high load (High Load task A - High Load task B) revealed a significant cluster in the orbitofrontal cortex (OFC) ($F = 13.39$ (df = [1, 29]), $p < 0.001$) along with some small activation in the frontal cortex.

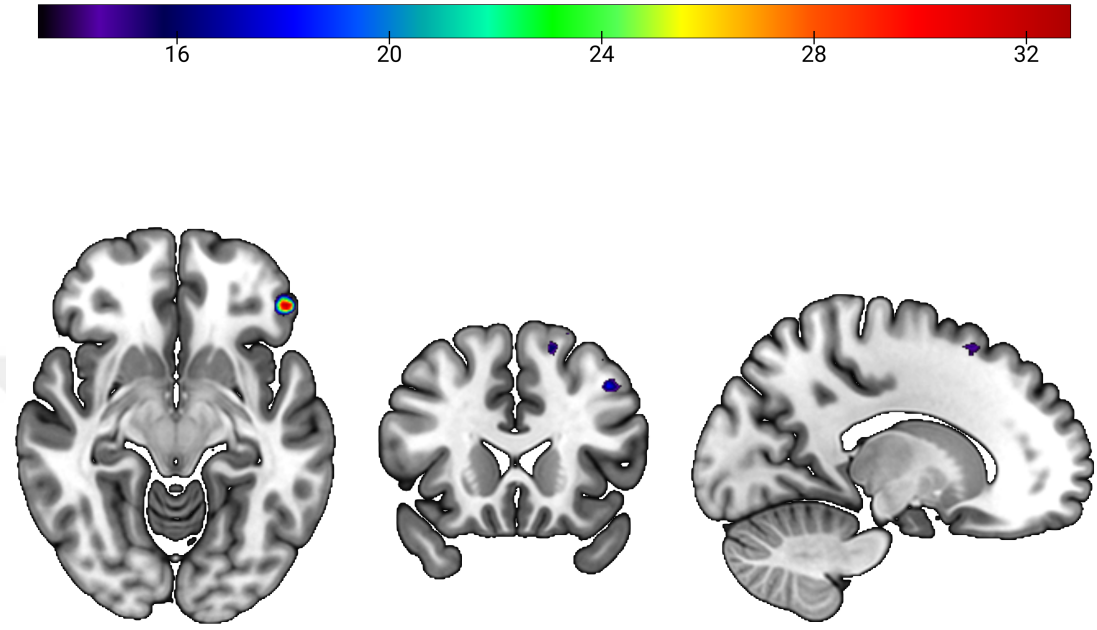

Figure 3.8: Differentiation pattern of load effect on parametric retrieval of task A and task B, showing the comparison of task A's high >low load contrast and task B's high >low load contrast (task A >task B). Hot colors indicate stronger activity. Significant regions exhibit the activation above the threshold ($F = 13.39$, p <0.001).

However, a Bayesian repeated measures ANOVA investigating the effects of task type and ROI revealed no effect. The task-only model showed modest evidence (posterior probability = 0.361, BF_{10}= 0.874), while models including ROI or the ROI × Task interaction had minimal support (all posterior probabilities <0.12). The analysis of individual effects indicated weak evidence for including task ($P(M \mid \text{data}) = 0.470$, BF_incl = 0.590) and even weaker evidence for ROI ($P(M \mid \text{data}) = 0.227$) or their interaction ($P(M \mid \text{data}) = 0.002$). However, a post hoc comparison between two task levels (A vs. B) revealed strong evidence for a difference (BF = 3997.82), suggesting a localized effect despite the overall null findings.

In summary, these results indicate that MD regions respond differently to increased load during encoding depending on whether participants perform task

A or B. This was reflected in the high–low contrast of beta values, where task A showed widespread positive effects (greater activation under high load). In contrast, task B exhibited a more mixed and regionally variable pattern (Figure 3.4). On the other hand, during the recall phase, no significant interaction was found between load and task. WM load effects did not differ across tasks. The high–low load contrast during recall (Figure 3.7) showed reduced activation under high load, particularly in task B. However, the load-related differences were more uniform across ROIs. Finally, regions that showed modulation by performance accuracy during recall did not differ across tasks as well, suggesting that the neural correlates of recall success are shared between task A and task B, even though encoding demands vary.

Chapter 4

Discussion

This thesis investigated how the MD network and the DMN respond when WM is challenged by multitasking and working memory load. An interleaved task design was used to distinguish encoding and retrieval demands and task-specific neural responses under different load levels. This provided insights into whether maintaining information on one task while engaging in another causes interference or whether resources are flexibly allocated. More importantly, this study reveals how WM prioritizes and retrieves information under pressure.

The findings revealed a key dissociation between the encoding and retrieval phases. During encoding, the effects of the WM load differed between tasks. However, they did not differ between tasks during retrieval and in the neural correlates of performance accuracy. This suggests that encoding relies on task-specific resource allocation, but retrieval utilizes a domain-general mechanism. These findings challenge purely unitary or shared-resource models but instead support dynamic and phase-sensitive models of WM, such as embedded-process and hierarchical frameworks.

Task-Dependent Encoding Reflects Flexible Control

The behavioral data confirmed that our WM load manipulation was successful because the participants made more errors under high load conditions in task

A. The neural response during encoding also exhibited task differences. Specifically, task A was found to be associated with load-dependent deactivation in MD regions, whereas task B triggered activation. This asymmetry implies that the MD system arranges its response pattern based on the structure and demands of each task. This finding is consistent with hierarchical WM models [150, 99], which propose that WM has a top-down control system that allocates cognitive resources by task relevance. In this framework, task B may have been treated as more relevant or time-sensitive, leading to greater MD engagement, even though it was presented later. The anticipatory suppression of task A before task B even started supports the idea that WM dynamically prioritizes the upcoming task demands.

This also aligns with dual mechanisms of control theory [151]. It distinguishes proactive control from reactive control, where the former is involved before task execution, and the latter is employed in response to current demands. Our results suggest that the MD network works proactively during encoding by suppressing irrelevant or interfering information to support efficient incoming retrieval.

Uniform Retrieval Suggests Domain-General Access

In contrast to encoding, our recall results did not show differences between tasks, either in WM load effects or in performance-related brain activity. This null result was consistent across the MD regions and suggests that the brain retrieves information using a domain-general retrieval system once the information is encoded. Regardless of how or under which task context the information was encoded, recall mechanisms appear to access it uniformly and non-selectively.

This supports Cowan's embedded process model (2017) [152], which posits a shared 'focus of attention' that holds a small number of activated memory representations. According to this model, once the items are retrieved, the brain has independent access to them regardless of their encoding type. Similarly, Bundesen's (1990) [153] visual attention theory proposes that the prior encoding structure no longer affects access once memory items are prioritized.

The results align with previous research proposing that even if tasks are encoded differently, those differences can be reversed or eliminated during retrieval [154]. Frankly, the brain may use a common mechanism to retrieve information and process it in a similar way under pressure and cognitive load [118]. Therefore, the brain might use a more general strategy that makes the operations more robust and efficient instead of following the exact structure of each task during recall. This helps the brain cope better with cognitive demands.

From Segregation to Integration of MD Network

The MD system responded to each task during encoding but not during retrieval. This shift suggests a phase-dependent role for MD regions, which allows them to organize and distinguish task demands during encoding. However, it switches to a more unified and non-task-specific retrieval type once the task requires access to stored information.

This pattern supports the idea that the MD system works flexibly and in a context-sensitive way [155, 102]. During encoding, different parts of the MD system become active depending on the structure and goals of the task. However, shared areas of the brain are used during recall, mainly helping to provide attention, monitoring, and choosing the correct information. This dynamic activity aligns with neurocomputational models of WM control[156], which argue that the brain can direct and update task information as needed.

The fact that MD activation was associated with working memory load but not with performance accuracy reinforces this idea further. It shows that these brain regions are more involved in control demands and priorities than in stimulus fidelity. This is consistent with slot-based working memory models [40], which suggest that memory is made of separate pieces, not one smooth precision scale. So, the MD system seems to manage how the information is organized and retrieved rather than how precisely it is encoded.

Selective Encoding Suppression and Global Retrieval Disengagement of DMN

Similarly to the MD system, the DMN also dynamically shifted based on the phase. Encoding of items from task A under high working memory load led to strong deactivation, while items from task B did not. This fits the dual-competition framework [157], which proposes that task-relevant and task-irrelevant processes compete for limited neural resources. Because DMN suppression during the demanding encoding phase helps reduce internal distractions and improves task performance [134].

The different brain patterns in MD and DMN systems during encoding show that the brain uses separate control and filtering processes. However, during recall, both tasks use the same general brain systems, showing shared selection mechanisms. MD areas follow task demands rather than how clearly something is remembered, and DMN suppression changes depending on the task phase and importance. These results support the idea that working memory is flexible and adapts over time, instead of using one fixed pool of brain resources.

However, DMN was suppressed during recall regardless of the task. This supports the idea that the brain uses a general process to inhibit internal processes during recall [137, 158]. The absence of task differences implies that the brain does not selectively filter internal distractions at that point. Instead, it uses a general inhibitory mechanism during retrieval to better focus.

In general, the findings challenge the unitary WM models assuming that a shared resource is distributed between different tasks [37]. Instead, they support more diverse models, such as Cowan's embedded process model [152] and Baddeley's multi-component framework [17]. These models separate the storage and control mechanisms [152], and allow the brain to be flexible depending on the task phase [17]. The different brain patterns in MD and DMN systems during encoding show that the brain utilizes different control and filtering processes. However, both tasks use the same general brain systems during recall, indicating unified selection mechanisms. MD regions track task demands rather than precision, and DMN suppression adapts with the task phase and priority. These observations support the idea that working memory is dynamic and adapts over time instead of using a fixed pool of resources.

4.0.1 Conclusion

This thesis reveals a dynamic interplay between WM load effects during multitasking. Neural responses differ according to the task and the load of working memory during encoding. This supports hierarchical and control-based models. However, they shift to a domain-general access strategy during recall that minimizes task differences and supports a unified retrieval mechanism. The results show that working memory is sensitive to context and dynamic. It uses different control methods depending on the task phase, sometimes inhibiting distractions before they happen (proactive suppression) and sometimes selecting the correct information as needed (reactive selection). These findings support hybrid models rather than classical unitary ones. Additionally, they contribute to understanding how memory and cognitive control interact. Future research should investigate individual differences, as working memory capacity varies from person to person. In this way, the results highlight the dynamic and adaptive nature of working memory, emphasizing that retrieval mechanisms operate differently from encoding and are less sensitive to task-specific demands.

Bibliography

[1] T. P. Alloway and E. Copello, "Working memory: The what, the why, and the how," *The Educational and Developmental Psychologist*, vol. 30, no. 2, pp. 105–118, 2013.

[2] E. J. Adams, A. T. Nguyen, and N. Cowan, "Theories of working memory: Differences in definition, degree of modularity, role of attention, and purpose," *Language, speech, and hearing services in schools*, vol. 49, no. 3, pp. 340–355, 2018.

[3] R. C. Atkinson and R. M. Shiffrin, "Human memory: A proposed system and its control processes," in *Psychology of learning and motivation*, vol. 2, pp. 89–195, Elsevier, 1968.

[4] G. A. Miller *et al.*, "The magical number seven, plus or minus two," *Psychological review*, vol. 63, no. 2, pp. 81–97, 1956.

[5] N. Cowan, "What are the differences between long-term, short-term, and working memory?," *Progress in brain research*, vol. 169, pp. 323–338, 2008.

[6] M. D'Esposito and B. R. Postle, "The cognitive neuroscience of working memory," *Annual review of psychology*, vol. 66, no. 1, pp. 115–142, 2015.

[7] L. R. Squire, "Memory systems of the brain: a brief history and current perspective," *Neurobiology of learning and memory*, vol. 82, no. 3, pp. 171–177, 2004.

[8] L. R. Squire and J. T. Wixted, "The cognitive neuroscience of human memory since hm," *Annual review of neuroscience*, vol. 34, no. 1, pp. 259–288, 2011.

[9] R. W. Engle, "Working memory capacity as executive attention," *Current directions in psychological science*, vol. 11, no. 1, pp. 19–23, 2002.

[10] A. D. Baddeley and G. Hitch, "Working memory," *Psychology of Learning and Motivation*, vol. 8, pp. 47–89, 1974.

[11] J. G. Goldman, C. Williams-Gray, R. A. Barker, J. E. Duda, and J. E. Galvin, "The spectrum of cognitive impairment in lewy body diseases," *Movement Disorders*, vol. 29, no. 5, pp. 608–621, 2014.

[12] S. Lugtmeijer, N. A. Lammers, E. H. de Haan, F.-E. de Leeuw, and R. P. Kessels, "Post-stroke working memory dysfunction: A meta-analysis and systematic review," *Neuropsychology review*, vol. 31, pp. 202–219, 2021.

[13] E. Rose and K. Ebmeier, "Pattern of impaired working memory during major depression," *Journal of affective disorders*, vol. 90, no. 2-3, pp. 149–161, 2006.

[14] G. A. Miller, E. Galanter, and K. H. Pribram, "Plans and the structure of behavior.," *New York*, 1960.

[15] A. Newell and H. Simon, "The logic theory machine–a complex information processing system," *IRE Transactions on information theory*, vol. 2, no. 3, pp. 61–79, 1956.

[16] K. Gilhooly, "Working memory and planning," in *The cognitive psychology of planning*, pp. 81–87, Psychology Press, 2004.

[17] A. Baddeley, "The episodic buffer: a new component of working memory?," *Trends in cognitive sciences*, vol. 4, no. 11, pp. 417–423, 2000.

[18] K. Oberauer, S. Farrell, C. Jarrold, and S. Lewandowsky, "What limits working memory capacity?," *Psychological bulletin*, vol. 142, no. 7, p. 758, 2016.

[19] S. J. Luck and E. K. Vogel, "Visual working memory capacity: from psychophysics and neurobiology to individual differences," *Trends in cognitive sciences*, vol. 17, no. 8, pp. 391–400, 2013.

[20] A. D. Baddeley, N. Thomson, and M. Buchanan, "Word length and the structure of short-term memory," *Journal of verbal learning and verbal behavior*, vol. 14, no. 6, pp. 575–589, 1975.

[21] R. Schweickert and B. Boruff, "Short-term memory capacity: Magic number or magic spell?," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 12, no. 3, p. 419, 1986.

[22] W. J. Ma, M. Husain, and P. M. Bays, "Changing concepts of working memory," *Nature neuroscience*, vol. 17, no. 3, pp. 347–356, 2014.

[23] N. Cowan, J. N. Rouder, C. L. Blume, and J. S. Saults, "Models of verbal working memory capacity: What does it take to make them work?," *Psychological review*, vol. 119, no. 3, p. 480, 2012.

[24] W. J. Ma, R. van den Berg, and H. Shin, "Continuous resources and variable precision in working memory," *Journal of Vision*, vol. 13, no. 9, pp. 1365–1365, 2013.

[25] C. Donkin, R. M. Nosofsky, J. M. Gold, and R. M. Shiffrin, "Discrete-slots models of visual working-memory response times.," *Psychological Review*, vol. 120, no. 4, p. 873, 2013.

[26] K. Oberauer and R. Kliegl, "A formal model of capacity limits in working memory," *Journal of memory and language*, vol. 55, no. 4, pp. 601–626, 2006.

[27] D. E. Broadbent, *Perception and communication*. Elsevier, 2013.

[28] J. Brown, "Some tests of the decay theory of immediate memory," *Quarterly journal of experimental psychology*, vol. 10, no. 1, pp. 12–21, 1958.

[29] L. Peterson and M. J. Peterson, "Short-term retention of individual verbal items.," *Journal of experimental psychology*, vol. 58, no. 3, p. 193, 1959.

[30] J. S. Reitman, "Mechanisms of forgetting in short-term memory," *Cognitive Psychology*, vol. 2, no. 2, pp. 185–195, 1971.

[31] J. S. Reitman, "Without surreptitious rehearsal, information in short-term memory decay," *Journal of verbal learning and verbal behavior*, vol. 13, no. 4, pp. 365–377, 1974.

[32] H. L. Roediger, J. L. Knight, and B. H. Kantowitz, "Inferring decay in short-term memory: The issue of capacity," *Memory & Cognition*, vol. 5, pp. 167–176, 1977.

[33] R. M. Shiffrin, "Information persistence in short-term memory.," *Journal of Experimental Psychology*, vol. 100, no. 1, p. 39, 1973.

[34] J. Brown, "Distortions in immediate memory," *Quarterly journal of experimental psychology*, vol. 8, no. 3, pp. 134–139, 1956.

[35] D. A. Norman and D. G. Bobrow, "On data-limited and resource-limited processes," *Cognitive psychology*, vol. 7, no. 1, pp. 44–64, 1975.

[36] D. Kahneman, *Attention and effort*, vol. 1063. Citeseer, 1973.

[37] P. M. Bays and M. Husain, "Dynamic shifts of limited working memory resources in human vision," *Science*, vol. 321, no. 5890, pp. 851–854, 2008.

[38] G. A. Alvarez and P. Cavanagh, "The capacity of visual short-term memory is set both by visual information load and by number of objects," *Psychological science*, vol. 15, no. 2, pp. 106–111, 2004.

[39] N. Cowan, "The magical mystery four: How is working memory capacity limited, and why?," *Current directions in psychological science*, vol. 19, no. 1, pp. 51–57, 2010.

[40] W. Zhang and S. J. Luck, "Discrete fixed-resolution representations in visual working memory," *Nature*, vol. 453, no. 7192, pp. 233–235, 2008.

[41] K. A. Ericsson and W. Kintsch, "Long-term working memory.," *Psychological review*, vol. 102, no. 2, p. 211, 1995.

[42] J. S. Nairne, "A feature model of immediate memory," *Memory & cognition*, vol. 18, pp. 251–269, 1990.

[43] K. Oberauer, S. Lewandowsky, S. Farrell, C. Jarrold, and M. Greaves, "Modeling working memory: An interference model of complex span," *Psychonomic bulletin & review*, vol. 19, pp. 779–819, 2012.

[44] J. W. Aldridge, H. R. Garcia, and G. Mena, "Habituation as a necessary condition for maintenance rehearsal," *Journal of Memory and Language*, vol. 26, no. 6, pp. 632–637, 1987.

[45] R. H. Phaf and G. Wolters, "Attentional shifts in maintenance rehearsal," *The American journal of psychology*, pp. 353–382, 1993.

[46] L. J. Lohnas, "A retrieved context model of serial recall and free recall," *Computational Brain & Behavior*, vol. 8, no. 1, pp. 1–35, 2025.

[47] M. Duncan and S. Lewandowsky, "The time course of response suppression: No evidence for a gradual release from inhibition," *Memory*, vol. 13, no. 3-4, pp. 236–246, 2005.

[48] R. N. Henson, "Item repetition in short-term memory: Ranschburg repeated.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 24, no. 5, p. 1162, 1998.

[49] J. C. Jahnke, "Output interference and the ranschburg effect," *Journal of Verbal Learning and Verbal Behavior*, vol. 8, no. 5, pp. 614–621, 1969.

[50] L. Hasher, R. T. Zacks, and C. P. May, *Inhibitory control, circadian arousal, and age.*, pp. 653–675. The MIT Press, 1999.

[51] L. Hasher and R. T. Zacks, "Working memory, comprehension, and aging: A review and a new view," *Psychology of learning and motivation*, vol. 22, pp. 193–225, 1988.

[52] K. Oberauer, "Removing irrelevant information from working memory: a cognitive aging study with the modified sternberg task.," *Journal of experimental Psychology: learning, Memory, and cognition*, vol. 27, no. 4, p. 948, 2001.

[53] K. Oberauer, "Access to information in working memory: exploring the focus of attention.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 28, no. 3, p. 411, 2002.

[54] K. Oberauer, "Control of the contents of working memory–a comparison of two paradigms and two age groups.," *Journal of experimental Psychology: learning, Memory, and cognition*, vol. 31, no. 4, p. 714, 2005.

[55] A. D. Endress and S. Szabó, "Interference and memory capacity limitations.," *Psychological review*, vol. 124, no. 5, p. 551, 2017.

[56] M. C. Anderson, "Rethinking interference theory: Executive control and the mechanisms of forgetting," *Journal of memory and language*, vol. 49, no. 4, pp. 415–445, 2003.

[57] M. W. Howard and M. J. Kahana, "A distributed representation of temporal context," *Journal of mathematical psychology*, vol. 46, no. 3, pp. 269–299, 2002.

[58] A. D. Endress and M. C. Potter, "Large capacity temporary visual memory.," *Journal of Experimental Psychology: General*, vol. 143, no. 2, p. 548, 2014.

[59] A. Hollingworth, "Constructing visual representations of natural scenes: the roles of short-and long-term visual memory.," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 30, no. 3, p. 519, 2004.

[60] D. Melcher, "Persistence of visual memory for scenes," *Nature*, vol. 412, no. 6845, pp. 401–401, 2001.

[61] T. S. Braver, J. R. Gray, G. C. Burgess, *et al.*, "Explaining the many varieties of working memory variation: Dual mechanisms of cognitive control," *Variation in working memory*, vol. 75, no. 106, 2007.

[62] G. C. Burgess, J. R. Gray, A. R. Conway, and T. S. Braver, "Neural mechanisms of interference control underlie the relationship between fluid intelligence and working memory span.," *Journal of experimental psychology: general*, vol. 140, no. 4, p. 674, 2011.

[63] D. Badre and A. D. Wagner, "Selection, integration, and conflict monitoring: assessing the nature and generality of prefrontal cognitive control mechanisms," *Neuron*, vol. 41, no. 3, pp. 473–487, 2004.

[64] K. R. Hammond and D. A. Summers, "Cognitive control.," *Psychological review*, vol. 79, no. 1, p. 58, 1972.

[65] F. A. Conners, "Attentional control and the simple view of reading," *Reading and Writing*, vol. 22, pp. 591–613, 2009.

[66] C. N. Arrington, P. A. Kulesz, D. J. Francis, J. M. Fletcher, and M. A. Barnes, "The contribution of attentional control and working memory to reading comprehension and decoding," *Scientific Studies of Reading*, vol. 18, no. 5, pp. 325–346, 2014.

[67] M. J. Kane, A. R. Conway, D. Z. Hambrick, and R. W. Engle, "Variation in working memory capacity as variation in executive attention and control.," *American Psychological Association*, 2007.

[68] N. Unsworth and R. W. Engle, "The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory.," *Psychological review*, vol. 114, no. 1, p. 104, 2007.

[69] R. W. Engle, M. J. Kane, S. W. Tuholski, *et al.*, "Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex," *Models of working memory: Mechanisms of active maintenance and executive control*, vol. 4, pp. 102–134, 1999.

[70] M. J. Kane, A. R. Conway, T. K. Miura, and G. J. Colflesh, "Working memory, attention control, and the n-back task: a question of construct

validity.," *Journal of Experimental psychology: learning, memory, and cognition*, vol. 33, no. 3, p. 615, 2007.

[71] M. A. Just and P. A. Carpenter, "A capacity theory of comprehension: individual differences in working memory.," *Psychological review*, vol. 99, no. 1, p. 122, 1992.

[72] A. Kiyonaga and T. Egner, "Resource-sharing between internal maintenance and external selection modulates attentional capture by working memory content," *Frontiers in human neuroscience*, vol. 8, p. 670, 2014.

[73] R. J. Allen, A. D. Baddeley, and G. J. Hitch, "Is the binding of visual features in working memory resource-demanding?," *Journal of experimental psychology: General*, vol. 135, no. 2, p. 298, 2006.

[74] N. Cowan, E. M. Elliott, J. S. Saults, C. C. Morey, S. Mattox, A. Hismjatullina, and A. R. Conway, "On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes," *Cognitive psychology*, vol. 51, no. 1, pp. 42–100, 2005.

[75] J. M. Chein, A. B. Moore, and A. R. Conway, "Domain-general mechanisms of complex working memory span," *Neuroimage*, vol. 54, no. 1, pp. 550–559, 2011.

[76] C. C. Morey and M. Bieler, "Visual short-term memory always requires general attention," *Psychonomic bulletin & review*, vol. 20, pp. 163–170, 2013.

[77] P. Barrouillet, S. Bernardin, S. Portrat, E. Vergauwe, and V. Camos, "Time and cognitive load in working memory.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 33, no. 3, p. 570, 2007.

[78] E. Hazeltine and T. Wifall, "Searching working memory for the source of dual-task costs," *Psychological Research*, vol. 75, no. 6, pp. 466–475, 2011.

[79] E. Vergauwe, P. Barrouillet, and V. Camos, "Do mental processes share a domain-general resource?," *Psychological science*, vol. 21, no. 3, pp. 384–390, 2010.

[80] K. Watanabe and S. Funahashi, "Neural mechanisms of dual-task interference and cognitive capacity limitation in the prefrontal cortex," *Nature neuroscience*, vol. 17, no. 4, pp. 601–611, 2014.

[81] K. Oberauer, "Working memory and attention–a conceptual analysis and review," *Journal of cognition*, vol. 2, no. 1, p. 36, 2019.

[82] H. Tsubomi, K. Fukuda, K. Watanabe, and E. K. Vogel, "Neural limits to representing objects still within view," *Journal of Neuroscience*, vol. 33, no. 19, pp. 8257–8263, 2013.

[83] D. Fougnie and R. Marois, "Distinct capacity limits for attention and working memory: Evidence from attentive tracking and visual working memory paradigms," *Psychological science*, vol. 17, no. 6, pp. 526–534, 2006.

[84] A. S. Souza and K. Oberauer, "The contributions of visual and central attention to visual working memory," *Attention, Perception, & Psychophysics*, vol. 79, pp. 1897–1916, 2017.

[85] G. F. Woodman and M. M. Chun, "The role of working memory and long-term memory in visual search," *Visual Cognition*, vol. 14, no. 4-8, pp. 808–830, 2006.

[86] R. M. Shiffrin and W. Schneider, "Controlled and automatic human information processing: Ii. perceptual learning, automatic attending and a general theory.," *Psychological review*, vol. 84, no. 2, p. 127, 1977.

[87] K. Oberauer and H.-Y. Lin, "An interference model of visual working memory.," *Psychological review*, vol. 124, no. 1, p. 21, 2017.

[88] D. Badre and M. D'Esposito, "Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex," *Journal of cognitive neuroscience*, vol. 19, no. 12, pp. 2082–2099, 2007.

[89] C. Ranti, C. H. Chatham, and D. Badre, "Parallel temporal dynamics in hierarchical cognitive control," *Cognition*, vol. 142, pp. 205–229, 2015.

[90] A. A. Farooqui and T. Manly, "We do as we construe: extended behavior construed as one task is executed as one cognitive entity," *Psychological Research*, vol. 83, no. 1, pp. 84–103, 2019.

[91] D. Badre and D. E. Nee, "Frontal cortex and the hierarchical control of behavior," *Trends in cognitive sciences*, vol. 22, no. 2, pp. 170–188, 2018.

[92] A. Kiyonaga and M. D'Esposito, *Competition and control during working memory*. Cambridge University Press, 2020.

[93] D. W. Schneider and G. D. Logan, "Hierarchical control of cognitive processes: switching tasks in sequences.," *Journal of Experimental Psychology: General*, vol. 135, no. 4, p. 623, 2006.

[94] G. D. Logan and R. D. Gordon, "Executive control of visual attention in dual-task situations.," *Psychological review*, vol. 108, no. 2, p. 393, 2001.

[95] A. Dezfouli and B. W. Balleine, "Habits, action sequences and reinforcement learning," *European Journal of Neuroscience*, vol. 35, no. 7, pp. 1036–1051, 2012.

[96] A. A. Farooqui and T. Manly, "Cognitive entities underpinning task episodes," *bioRxiv*, p. 216796, 2017.

[97] A. A. Farooqui, T. Gezici, and T. Manly, "Chunking of control: An unrecognized aspect of cognitive resource limits," *Journal of Cognition*, vol. 6, no. 1, p. 25, 2023.

[98] A. A. Farooqui, D. Mitchell, R. Thompson, and J. Duncan, "Hierarchical organization of cognition reflected in distributed frontoparietal activity," *Journal of Neuroscience*, vol. 32, no. 48, pp. 17373–17381, 2012.

[99] D. Badre, "Cognitive control, hierarchy, and the rostro–caudal organization of the frontal lobes," *Trends in cognitive sciences*, vol. 12, no. 5, pp. 193–200, 2008.

[100] J. Duncan and A. M. Owen, "Common regions of the human frontal lobe recruited by diverse cognitive demands," *Trends in neurosciences*, vol. 23, no. 10, pp. 475–483, 2000.

[101] J. Duncan, "Eps mid-career award 2004: brain mechanisms of attention," *Quarterly Journal of Experimental Psychology*, vol. 59, no. 1, pp. 2–27, 2006.

[102] E. Fedorenko, J. Duncan, and N. Kanwisher, "Broad domain generality in focal regions of frontal and parietal cortex," *Proceedings of the National Academy of Sciences*, vol. 110, no. 41, pp. 16616–16621, 2013.

[103] V. Menon and M. D'Esposito, "The role of pfc networks in cognitive control and executive function," *Neuropsychopharmacology*, vol. 47, no. 1, pp. 90–103, 2022.

[104] E. K. Miller, "The prefontral cortex and cognitive control," *Nature reviews neuroscience*, vol. 1, no. 1, pp. 59–65, 2000.

[105] K. Matsumoto and K. Tanaka, "Conflict and cognitive control," *Science*, vol. 303, no. 5660, pp. 969–970, 2004.

[106] T. Wen and T. Egner, "Context-independent scaling of neural responses to task difficulty in the multiple-demand network," *Cerebral Cortex*, vol. 33, no. 10, pp. 6013–6027, 2023.

[107] J. Duncan, "The structure of cognition: attentional episodes in mind and brain," *Neuron*, vol. 80, no. 1, pp. 35–50, 2013.

[108] J. Duncan, M. Assem, and S. Shashidhara, "Integrated intelligence from distributed brain activity," *Trends in cognitive sciences*, vol. 24, no. 10, pp. 838–852, 2020.

[109] R. Desimone, J. Duncan, *et al.*, "Neural mechanisms of selective visual attention," *Annual review of neuroscience*, vol. 18, no. 1, pp. 193–222, 1995.

[110] E. K. Miller and J. D. Cohen, "An integrative theory of prefrontal cortex function," *Annual review of neuroscience*, vol. 24, no. 1, pp. 167–202, 2001.

[111] K. Christoff, K. Keramatian, A. M. Gordon, R. Smith, and B. Mädler, "Prefrontal organization of cognitive control according to levels of abstraction," *Brain research*, vol. 1286, pp. 94–105, 2009.

[112] N. Sigala, M. Kusunoki, I. Nimmo-Smith, D. Gaffan, and J. Duncan, "Hierarchical coding for sequential task events in the monkey prefrontal cortex," *Proceedings of the National Academy of Sciences*, vol. 105, no. 33, pp. 11969–11974, 2008.

[113] A. A. Farooqui and T. Manly, "When attended and conscious perception deactivates fronto-parietal regions," *Cortex*, vol. 107, pp. 166–179, 2018.

[114] A. A. Farooqui and T. Manly, "Hierarchical cognition causes task-related deactivations but not just in default mode regions," *eneuro*, vol. 5, no. 6, 2018.

[115] A. W. MacDonald, J. D. Cohen, V. A. Stenger, and C. S. Carter, "Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control," *Science*, vol. 288, no. 5472, pp. 1835–1838, 2000.

[116] J. B. Hutchinson, M. R. Uncapher, and A. D. Wagner, "Posterior parietal cortex and episodic retrieval: convergent and divergent effects of attention and memory," *Learning & memory*, vol. 16, no. 6, pp. 343–356, 2009.

[117] M. M. Botvinick, T. S. Braver, D. M. Barch, C. S. Carter, and J. D. Cohen, "Conflict monitoring and cognitive control.," *Psychological review*, vol. 108, no. 3, p. 624, 2001.

[118] A. Gazzaley and A. C. Nobre, "Top-down modulation: bridging selective attention and working memory," *Trends in cognitive sciences*, vol. 16, no. 2, pp. 129–135, 2012.

[119] A. K. Barbey, M. Koenigs, and J. Grafman, "Dorsolateral prefrontal contributions to human working memory," *cortex*, vol. 49, no. 5, pp. 1195–1205, 2013.

[120] N. G. Müller and R. T. Knight, "The functional neuroanatomy of working memory: contributions of human brain lesion studies," *Neuroscience*, vol. 139, no. 1, pp. 51–58, 2006.

[121] A. Anticevic, G. Repovs, and D. M. Barch, "Working memory encoding and maintenance deficits in schizophrenia: neural evidence for activation and

deactivation abnormalities," *Schizophrenia bulletin*, vol. 39, no. 1, pp. 168–178, 2013.

[122] J. Murray, J. Jaramillo, and X. Wang, "Working memory and decision-making in a frontoparietal circuit model," *Journal of Neuroscience*, vol. 37, pp. 12167–12186, 2017.

[123] J. Karbach and P. Verhaeghen, "Making working memory work: a meta-analysis of executive-control and working memory training in older adults," *Psychological science*, vol. 25, no. 11, pp. 2027–2037, 2014.

[124] M. Assem, I. Blank, Z. Mineroff, A. Ademoğlu, and E. Fedorenko, "Activity in the fronto-parietal multiple-demand network is robustly associated with individual differences in working memory and fluid intelligence," *Cortex*, vol. 131, pp. 1–16, 2020.

[125] M. E. Raichle, A. M. MacLeod, A. Z. Snyder, W. J. Powers, D. A. Gusnard, and G. L. Shulman, "A default mode of brain function," *Proceedings of the national academy of sciences*, vol. 98, no. 2, pp. 676–682, 2001.

[126] R. L. Buckner, J. R. Andrews-Hanna, and D. L. Schacter, "The brain's default network: anatomy, function, and relevance to disease," *Annals of the new York Academy of Sciences*, vol. 1124, no. 1, pp. 1–38, 2008.

[127] M. F. Mason, M. I. Norton, J. D. Van Horn, D. M. Wegner, S. T. Grafton, and C. N. Macrae, "Wandering minds: the default network and stimulus-independent thought," *science*, vol. 315, no. 5810, pp. 393–395, 2007.

[128] F. R. Azarias, G. H. D. R. Almeida, L. F. de Melo, R. E. G. Rici, and D. A. Maria, "The journey of the default mode network: Development, function, and impact on mental health," *Biology*, vol. 14, no. 4, p. 395, 2025.

[129] M. S. Sendi, E. Zendehrouh, C. A. Ellis, Z. Liang, Z. Fu, D. H. Mathalon, J. M. Ford, A. Preda, T. G. van Erp, R. L. Miller, *et al.*, "Aberrant dynamic functional connectivity of default mode network in schizophrenia and links to symptom severity," *Frontiers in neural circuits*, vol. 15, p. 649417, 2021.

[130] R. N. Spreng and C. L. Grady, "Patterns of brain activity supporting autobiographical memory, prospection, and theory of mind, and their relationship to the default mode network," *Journal of cognitive neuroscience*, vol. 22, no. 6, pp. 1112–1123, 2010.

[131] D. R. Addis, A. T. Wong, and D. L. Schacter, "Remembering the past and imagining the future: common and distinct neural substrates during event construction and elaboration," *Neuropsychologia*, vol. 45, no. 7, pp. 1363–1377, 2007.

[132] L. Schilbach, S. B. Eickhoff, A. Rotarska-Jagiela, G. R. Fink, and K. Vogeley, "Minds at rest? social cognition as the default mode of cognizing and its putative relationship to the "default system" of the brain," *Consciousness and cognition*, vol. 17, no. 2, pp. 457–467, 2008.

[133] J. R. Andrews-Hanna, J. Smallwood, and R. N. Spreng, "The default network and self-generated thought: Component processes, dynamic control, and clinical relevance," *Annals of the new York Academy of Sciences*, vol. 1316, no. 1, pp. 29–52, 2014.

[134] A. Anticevic, M. W. Cole, J. D. Murray, P. R. Corlett, X.-J. Wang, and J. H. Krystal, "The role of default network deactivation in cognition and disease," *Trends in cognitive sciences*, vol. 16, no. 12, pp. 584–592, 2012.

[135] D. H. Weissman, K. Roberts, K. Visscher, and M. Woldorff, "The neural bases of momentary lapses in attention," *Nature neuroscience*, vol. 9, no. 7, pp. 971–978, 2006.

[136] K. A. McKiernan, J. N. Kaufman, J. Kucera-Thompson, and J. R. Binder, "A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging," *Journal of cognitive neuroscience*, vol. 15, no. 3, pp. 394–408, 2003.

[137] M. D. Fox, A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. Van Essen, and M. E. Raichle, "The human brain is intrinsically organized into dynamic, anticorrelated functional networks," *Proceedings of the National Academy of Sciences*, vol. 102, no. 27, pp. 9673–9678, 2005.

[138] R. N. Spreng, W. D. Stevens, J. P. Chamberlain, A. W. Gilmore, and D. L. Schacter, "Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition," *Neuroimage*, vol. 53, no. 1, pp. 303–317, 2010.

[139] T. Piccoli, G. Valente, D. Linden, M. Re, F. Esposito, A. Sack, and F. Salle, "the default mode network and the working memory network are not anti-correlated during all phases of a working memory task," *Plos One*, vol. 10, p. e0123354, 2015.

[140] U. Hasson, H. Nusbaum, and S. Small, "task-dependent organization of brain regions active during rest," *Proceedings of the National Academy of Sciences*, vol. 106, pp. 10841–10846, 2009.

[141] J. Smallwood, K. Brown, B. Baird, and J. Schooler, "cooperation between the default mode network and the frontal–parietal network in the production of an internal train of thought," *Brain Research*, vol. 1428, pp. 60–70, 2012.

[142] R. Spreng and D. Schacter, "default network modulation and large-scale network interactivity in healthy young and old adults," *Cerebral Cortex*, vol. 22, pp. 2610–2621, 2011.

[143] R. Bluhm, C. Clark, A. McFarlane, K. Moores, M. Shaw, and R. Lanius, "default network connectivity during a working memory task," *Human Brain Mapping*, vol. 32, pp. 1029–1035, 2010.

[144] S. Seoane, M. Heuvel, Acebes, and N. Janssen, "the subcortical default mode network and alzheimer's disease: a systematic review and activation likelihood estimation meta-analysis," 2023.

[145] J. Peirce, J. R. Gray, S. Simpson, M. MacAskill, R. Höchenberger, H. Sogo, E. Kastman, and J. K. Lindeløv, "Psychopy2: Experiments in behavior made easy," *Behavior research methods*, vol. 51, pp. 195–203, 2019.

[146] R. Cusack, A. Vicente-Grabovetsky, D. J. Mitchell, C. J. Wild, T. Auer, A. C. Linke, and J. E. Peelle, "Automatic analysis (aa): efficient neuroimaging workflows and parallel processing using matlab and xml," *Frontiers in neuroinformatics*, vol. 8, p. 90, 2015.

[147] T. M. Inc., "Statistics and machine learning toolbox," 2022.

[148] S. B. Eickhoff, T. Paus, S. Caspers, M.-H. Grosbras, A. C. Evans, K. Zilles, and K. Amunts, "Assignment of functional activations to probabilistic cytoarchitectonic areas revisited," *Neuroimage*, vol. 36, no. 3, pp. 511–521, 2007.

[149] JASP Team, "JASP (Version 0.19.3)[Computer software]," 2025.

[150] E. Koechlin and C. Summerfield, "An information theoretical approach to prefrontal executive function," *Trends in cognitive sciences*, vol. 11, no. 6, pp. 229–235, 2007.

[151] T. S. Braver, "The variable nature of cognitive control: a dual mechanisms framework," *Trends in cognitive sciences*, vol. 16, no. 2, pp. 106–113, 2012.

[152] N. Cowan, "The many faces of working memory and short-term storage," *Psychonomic bulletin & review*, vol. 24, pp. 1158–1170, 2017.

[153] C. Bundesen, "A theory of visual attention.," *Psychological review*, vol. 97, no. 4, p. 523, 1990.

[154] P. Jolicœur and R. Dell'Acqua, "The demonstration of short-term consolidation," *Cognitive psychology*, vol. 36, no. 2, pp. 138–202, 1998.

[155] J. Duncan, "The multiple-demand (md) system of the primate brain: mental programs for intelligent behaviour," *Trends in cognitive sciences*, vol. 14, no. 4, pp. 172–179, 2010.

[156] R. C. O'Reilly and M. J. Frank, "Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia," *Neural computation*, vol. 18, no. 2, pp. 283–328, 2006.

[157] L. Pessoa, "How do emotion and motivation direct executive control?," *Trends in cognitive sciences*, vol. 13, no. 4, pp. 160–166, 2009.

[158] M. E. Raichle, "The brain's default mode network," *Annual review of neuroscience*, vol. 38, no. 1, pp. 433–447, 2015.