The oscillatory mechanisms of working memory maintenance

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von
Marcin Leszczyński
aus
Poznań, Polen

Bonn, 2016

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen
Fakultät der
Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. Nikolai Axmacher
2. Gutachter: Prof. Dr. Michael Hofmann

Tag der Promotion: 30.05.2016

Erscheinungsjahr: 2016

Declaration of Authorship

I, Marcin Leszczyński, declare that this thesis titled, 'Oscillatory mechanisms of working memory maintenance' and the work presented in it are my own. I confirm that:

- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.

Parts of the results and arguments presented in this thesis have been published in a form of scientific articles. The relevant articles include:

1. **Leszczyński, M**, Fell, J, Axmacher, N. (2015). Rhythmic working memory activation in the human hippocampus. *Cell Reports*, 13(6):1272-82.
2. Chaieb, L, **Leszczyński, M**, Axmacher, N, Höhne, M, Elger, CE, Fell, J. (2015). Theta-gamma phase-phase coupling during working memory maintenance in the human hippocampus. *Cognitive Neuroscience*, 6(4):149-57.
3. **Leszczyński*, M**, Wykowska*, A, Perez-Osorio, J, Müller, HJ. (2013). Deployment of spatial attention towards locations in memory representations. An EEG study. *PloS ONE*, 8(12).
4. **Leszczyński, M**, Myers, N, Akyürek, E, Schubö, A. (2012). Recoding between two types of short-term memory representation revealed by dynamics of memory search. *Journal of Cognitive Neuroscience*, 24(3): 653-63.
5. **Leszczyński, M**. (2011). How does hippocampus contribute to working memory processing? *Frontiers in Human Neuroscience*, 5: 168.

Signed: _____

Date: _____

Summary

Working memory (WM) is a cognitive process which allows for maintenance of information that is no longer perceived. Although theoretical models have recognized that working memory involves interactions across cell assemblies in multiple brain areas, the exact neural mechanisms which support this process remain unknown. In this thesis I investigate the neural dynamics in the human hippocampus, the ventral, dorsal and frontal cortex as well as the long-range network connectivity across these brain areas to understand how such a distributed network allows for maintenance of various information pieces in WM. The results described here support the model in which working memory relies on dynamic interactions across frequencies (the cross-frequency coupling, CFC) in a distributed network of cortical areas coordinated by the prefrontal cortex. In particular, maintenance of information during a delay period selectively involves the hippocampus, dorsal and ventral visual stream as well as the prefrontal cortex each of which represents different features.

The hippocampus contributes to this large network specifically by representing multiple items in working memory. In two independent experiments I observed that the low-frequency activity (a marker of neural inhibition) was linearly reduced across memory loads. Importantly, the hippocampus showed very prominent low-frequency power during maintenance of a single item suggesting that during this condition the neural processing was strongly inhibited. In turn, the broadband gamma activity was linearly increasing as a function of memory load. This pattern of results may be interpreted as reflecting an increased involvement of the hippocampus in representing longer sequences. Importantly, the low-frequency decrease was not static but fluctuated periodically between two different modes. One of the modes was characterized by the load-dependent power decreases and reduced cross-frequency coupling (*memory activation mode*) whereas the other mode was reflected by the load-independent high levels of power and increased coupling strength (*load-independent mode*). Crucially, these modes were temporally organized by the phase of an endogenous delta rhythm forming a "hierarchy of oscillations". This periodicity was essential for the successful performance. Finally, during the memory activation mode the WM capacity limit was inter-individually correlated with the peak frequency change as predicted by the multiplexing model of WM. All these effects were subsequently replicated in an independent dataset. These results suggest that the hippocampus is involved in WM maintenance showing periodic fluctuations between two different oscillatory modes. Parameters of the hippocampal iEEG signal correlate with individual WM capacity, specifically during the memory activation mode.

The ventral and dorsal visual stream each contributes to the distributed WM network by representing configuration and spatial information, respectively. Specifically, the alpha power in the ventral visual stream was decreased during maintenance of face identities. In turn, the alpha power was desynchronized in the dorsal visual stream while participants were maintaining face orientations. This shows that the alpha power double dissociates between the feature specific networks in

the ventral and dorsal visual stream. These effects are further interpreted as reflecting selective involvement of the dorsal and ventral visual pathway depending on the maintained features. Importantly, each of the visual streams was selectively synchronized with the prefrontal cortex depending on the memory condition and the alpha power. This corroborates a central prediciton from the gating by inhibition model which assumes that the increased alpha power serves as the mechanism for gating of information by inhibiting task redundunt pathways. Moreover, during maintenance of information the phase of alpha modulated the amplitude of high-frequency activity both in the dorsal and ventral visual stream. Additionally, the low-frequency phase in the prefrontal cortex modulated high-frequency activity both in the dorsal and ventral visual stream. These results suggest that both the dorsal and ventral visual streams are selectively involved during maintenance of distinct features (i.e. face orientation and identity, respectively). They also indicate that the prefrontal cortex selectively gets synchronized with the visual regions depending on the alpha power in that region and the maintained feature. Finally, the activity in the prefrontal cortex influences processing across long distance as evident from changes in the phase synchrony with the visual cortical areas and by modulating gamma power in the visual cortical regions.

It is also noted that the ventrolateral prefrontal cortex (vlPFC) contains information regarding abstract rules (i.e. response mapping). In particular, using a multivariate decoding approach I found that the local field potentials recorded from the vlPFC dissociate between different types of responses. At the same time I observed no evidence for the load-dependent or stimulus-specific changes in that brain region. The null effect should be treated with caution. Nevertheless, the current results suggest that the vlPFC may contribute to working memory by processing of abstract rules such as a mapping between the stimulus and the response. Furthermore, I found that the alpha power dependent duty cycle in the vlPFC constrains the duration of the gamma burst which has been suggested as a mechanism for neural inhibition. This finding is important because such a property of the alpha activity has never been observed in a brain region other than the primary sensory cortex.

Together, the results presented in this thesis support a model according to which the working memory is a complex and highly dynamic process engaging hierarchies of oscillations across multiple cortical regions. In particular, the hippocampus is important for the multi-item WM. The dorsal and ventral visual streams are relevant for distinct visual features. Finally, the prefrontal cortex represents abstract rules and influences processing in other cortical regions likely providing a top-down control over these regions.

Acknowledgements

I would like to thank Leila Chaieb, Amirhossein Jahanbekam, Daniel Labbé, Paweł Łupkowski, Hui Zhang and Monika Ziółkowska for helpful discussions and comments on an earlier version of the manuscript. I would also like to thank Jürgen Fell and Nikolai Axmacher for many helpful discussions and their support in the data analysis.

Contents

List of Figures

Abbreviations

ANOVA	Analysis of Variance
BOLD	Blood Oxygen Level Dependent
BA	Brodmann Areas
CDA	Contralateral Delay Activity
CFC	Cross Frequency Coupling
CO	Control condition
dlPFC	dorsal lateral Prefrontal Cortex
EcoG	Electrocorticography
EEG	Electroencephalogram
EPSP	Excitatory Postsynaptic Potential
ERP	Event Related Potential
fMRI	functional Magnetic Resonance Imaging
FoA	Focus of Attention
ID	Identity condition
iEEG	intracranial Electroencephalogram
IPSP	Inhibitory Postsynaptic Potential
lPFC	lateral Prefrontal Cortex
LTM	Long-Term Memory
MEG	Magnetoencephalography
mHFA	modulated High Frequency Activity
mLFA	modulating Low Frequency Activity
MNI	Montreal Neurological Institute
OR	Orientation condition
PAC	Phase Amplitude Coupling

PET	Positron Emission Tomography
PLV	Phase Locking Value
PPC	Posterior Parietal Cortex
RE	Thalamic Reticular
ROI	Region Of Interest
RT	Reaction Time
SAS	Supervisory Activating System
STM	Short-Term Memory
SVM	Support Vector Machine
TC	Thalamocortical
TMS	Transcranial Magnetic Stimulation
vlPFC	vetral lateral Prefrontal Cortex
VWM	Visual Working Memory
WM	Working memory
2AFC	Two Alternative Forced Choice

Chapter 1

Cognitive models of working memory

1.1 Introduction

Working memory (WM) refers to the temporary retention of information that is no longer perceived and extends to the processing of new and previously-stored information. The WM concept assumes a limited capacity system that maintains information over short temporal intervals and as such it is an interface between perception and long-term memory [LTM; Baddeley, 2000]. Working memory supports several other cognitive functions like attention [Akyürek et al., 2010, Awh et al., 2006], aforementioned long-term memory [Axmacher et al., 2010c, Cowan, 1995, Oberauer, 2002, Ruchkin et al., 2003] or movement preparation [Pearson et al., 2014]. According to Baddeley [Baddeley, 2003] the term *working memory* was first used by Pribram and colleagues in the context of cognitive theories referring to the computer metaphor of the mind [Miller et al., 1960]. Analogous to random-access memory (RAM), the term was meant to reflect a flexible workspace buffer where information from the long-term memory (a metaphorical hard disk) is loaded and used for further operations. The term *working memory* was popularized in the 1970s by Baddeley and Hitch [Baddeley and Hitch, 1974] who used it

to emphasize the difference between their multi-component model and the earlier model of short-term memory storage [Atkinson and Shiffrin, 1968].

In the early 1970s the model of passive short-term storage (sometimes referred to as the modal model) became insufficient to explain the observed data. There were two main sources of criticism. First, this theory assumed that the human memory consist of two storage systems. The short-term store was able to hold the limited number of information over short time intervals. The long-term store in turn was capable of storing large amounts of information over long time intervals. Importantly, this model assumed that relation between the two stores is sequential. Information first passes the short-term store before it gets to long-term storage. However, neuropsychological research have shown that this assumption is violated. In particular, it has been observed that patients with a severe deficit to short-term storage were nevertheless able to store information in long-term memory as good as healthy controls would do [Shallice and Warrington, 1970]. This clearly spoke against the idea of a single short-term storage which passes information to the long-term storage.

The second source of evidence comes from the study of Baddeley and Hitch [Baddeley and Hitch, 1974]. The authors were interested in understanding whether the short-term storage is indeed only a passive store or if it also actively contributes to other cognitive functions. They used a dual task design to investigate if loading short-term storage would influence performance in another complex task. They asked participants to maintain a set of digits (up to eight) which should fully load the short-term storage. Subsequently, the subjects were required to perform complex reasoning or learning tasks while they were holding these digits in their memory store. For example, the authors presented participants with simple yes/no questions or asked to study and recall unrelated words. Baddeley and Hitch hypothesized that if there is only one kind of short-term storage and it is critical for performing complex cognitive tasks (instead of only passively storing information), then the additional digit memory span should be detrimental for performance in these cognitive tasks. Although, performance dropped with memory load, participants were still able to do both tasks fairly well. The authors concluded that there

must be multiple systems for storing information and that loading one of these stores does not necessary interfere with other cognitive functions. They further proposed a new concept of multi-component working memory with multiple storage systems and the central executive which flexibly controls allocation of memory between the storage systems [Baddeley and Hitch, 1974, Smith and Kosslyn, 2007]. There are three major differences between the Baddeley and Hitch model and previous unitary theory of short-term memory. First, their model consists of multiple components. Second, it combines storage with active processing of information. Finally, unlike the unitary approach, their model describes WM as a more general process contributing to several other cognitive functions beyond simple storage of information [Baddeley, 2003]. The new concept of WM was introduced to emphasize that this is an active process which contributes to many other cognitive functions.

1.2 The multi-component model of working memory

In 1974 Baddeley and Hitch proposed a fundamentally different model of WM which consisted of two storage systems (the phonological loop and the visuospatial sketchpad) and control system (the central executive). The model was further extended in 2000 to include the third storage component named the episodic buffer [see Figure 1.1; Baddeley, 2000].

1.2.1 Phonological loop.

The *phonological loop* refers to an articulatory process which is composed of a short term storage that captures memory traces for few seconds before they fade away and a rehearsal process which allows for refreshing of a memory trace. The concept of the phonological loop is important because it is one of the first attempts to conceptualize the active working memory rehearsal process. Imagine you were

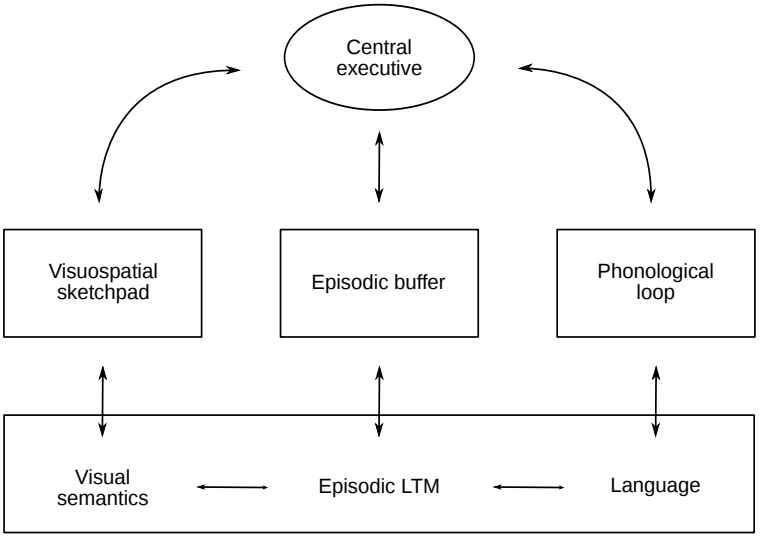

FIGURE 1.1: The multi-component model of working memory. The extended version of the model proposed by Baddeley [Baddeley, 2000]. The central executive and three "slave systems".

given a phone number: 1-2-5-6-9-4-8. Close your eyes and try to remember the digits. Some people report that while reading the sequence they hear their "inner voice". While their eyes are closed and no visual input is provided, they report that the "inner voice" repeats the digits. This metaphor is central to the understanding of the phonological loop. Since the rehearsal takes place in real time and it is a serial process - extending the number of items in the sequence will lead to a point in which the first item will stay too long in the phonological storage before it can be rehearsed and will fade away. Thus, the consequence of the serial articulatory rehearsal is a limit in the number of items that might be simultaneously maintained in memory (this is also refereed to as the limited memory span).

The concept of the *phonological loop* is supported by studies showing that the working memory span decreases while participants are required to rehearse words consisting of more syllables [Baddeley, 2000, Baddeley et al., 1975]. Longer words require more time to articulate. They will stay longer in the phonological store before they could be rehearsed and in turn will fade away leading to shorter memory span. Furthermore, the effect of word length on memory span disappears while preventing participants from any rehearsal by asking them to repeat irrelevant verbal

material [Baddeley, 2003]. This suggests that the word length becomes irrelevant for the memory span once the opportunity to verbally rehearse is blocked. Neuroimaging [Smith and Jonides, 1997] and neuropsychological [Trost and Gruber, 2012] studies suggest that there exist separate neural mechanisms for storage and rehearsal. The former was hypothesized to engage primarily the temporoparietal area (Brodmann Area BA 40), whereas the later has been related to the Broca's area [BA 6/44; Smith and Jonides, 1997, Trost and Gruber, 2012].

1.2.2 Visuospatial sketchpad.

The visuospatial sketchpad is a visual equivalent of the phonological loop. It plays an important role in the process of maintenance and manipulation of visuospatial information. If you were asked to recall pictures hanging on the walls of a room you are familiar with (other than the one you are currently in), you would be using the visuospatial sketchpad to scan information in your spatial working memory. Representations in visuospatial memory have been shown to share many properties with actually perceived visual representations and involve overlapping brain networks. For example, cueing attention towards representation in the visuospatial sketchpad results in a behavioral benefit just like it does while cuing attention to an external stimulus [Awh and Jonides, 2001, Griffin and Nobre, 2003]. Furthermore, studies, including my own, have shown that attending to a certain feature in an environment and attending to a feature in memory representation involves the same neural mechanisms [Griffin and Nobre, 2003, Lepsien and Nobre, 2007, Leszczyński et al., 2012, 2013].

Imagine that you are presented with four images in four quadrants of a computer screen. Next, images disappear from the screen and are masked with white noise patches to cancel afterimages. Finally, a cue indicating a certain location is presented. One very surprising result from such a task is that attention cued to the location where the stimulus **had been previously presented** results in a behavioral benefit as measured with decreased reaction time (RT) and increased accuracy for the cued, as compared to a random location. You will perform better

when discriminating the cued image than stimuli presented on other locations. This is a very surprising result because the cue only informs where to focus attention once the stimuli are offset. These results suggest that the representation maintained in the visuospatial sketchpad share spatial properties with the perceptual representation and that attention is similarly important for both. Specifically, allocating attention to either of the representations is associated with strengthening of a neural response and the behavioral benefit [Awh and Jonides, 2001, Griffin and Nobre, 2003, Lepsien and Nobre, 2007, Leszczyński et al., 2012, 2013].

1.2.3 Dissociated phonological and spatial storage.

The technique which allowed the discovery of the dissociated visuospatial sketchpad and the phonological loop is called the *selective interference* paradigm. The idea is very simple – introducing a secondary task during a maintenance interval of a working memory will lead to interference between the two tasks (i.e. lower the accuracy in the primary memory task or both). The assumption of selectivity further posits that the interference depends on the material of the secondary task. Imagine again the room you are familiar with and imagine you were asked to recall images hanging on the walls of this room. If I would additionally ask you to do a second task which involves manipulation of a visuospatial information, then the interference should be larger compared to a condition when the secondary task involves manipulation of a phonological information [J. et al., 1996]. This indeed has been shown. For example, Logie et al. [1990] compared working memory span for visual matrix patterns and visually presented letter sequences. They assumed that the former relies rather on the visuospatial sketchpad and the later on the phonological loop. These two tasks were combined with a secondary task which was either an arithmetic task (mental addition of aurally presented digits) or a task which involved manipulating of visuospatial material (filling squares in an imagined matrix according to aurally presented instructions). Importantly, the authors observed a condition specific disruption of performance. In other words accuracy was disrupted in the letter span task by the secondary arithmetic tasks but not

by a secondary visuospatial task. In contrast, the working memory span for visual matrix was specifically disrupted by the secondary visuospatial task. These results support the notion of separate and specialized visuospatial and phonological mechanisms in working memory [Logie et al., 1990].

Other evidence supporting for the concept of distinct visuospatial and phonological stores comes from a similar approach where the response is designed to relay either on the visuospatial or on the phonological information. In particular, these experiments required participants to answer questions regarding the outline of an imagined capital letters. Think for example of a capital letter *E*. Now imagine you are asked to navigate around it clockwise starting at the lower left. At each corner you need to answer yes/no question whether you are in the extreme top or the extreme bottom corner. This is a fairly easy task in which you are using a visuospatial information. The most important experimental manipulation was that some participants were asked **point** to the words *yes* or *no* printed in front of them. The other group had to **speak** the words *yes* or *no*. These experiments assume that a response which requires pointing towards a certain direction (i.e. one of the words) relies more on the visuospatial information than the type of response which requires speaking the relevant word. Indeed, longer response times were observed when participants were required to point compared to speaking condition. This in turn suggests that the pointing condition resulted in a stronger interference [Brooks, 1968, Smyth and Scholey, 1994].

Although the dissociation between these two storage systems have been well established, recent studies provided evidence that the separation is not rigid and in fact the transfer between the phonological and the visuospatial representation is likely to happen. In two experiments my colleagues and I used the technique of the event-related potentials (ERPs) to investigate a visuospatial working memory retrieval. Participants were presented with a visuospatial array of colored squares followed by a blank screen presented for various durations (maintenance interval). Next, a centrally presented abstract probe indicated one of the squares and participants were required to recall the color of the probed square. We measured an ERP component called the N2pc (negative in polarity ERP observed 200ms after

the stimulus onset over the posterior contralateral side) which has been related to allocation of the visuospatial attention [Luck and Hillyard, 1994]. First, we observed that the retrieval from the visuospatial storage is associated with allocation of attention just like it does during visual search [see also Dell'Acqua et al., 2010, Gazzaley and Nobre, 2012, Kuo et al., 2009]. Importantly, the N2pc (a marker of visuospatial attention) was linearly decreased with increasing maintenance interval. This suggests that the allocation of attention to the visuospatial array was prominent at short but attenuated at longer maintenance intervals. We further hypothesized that the allocation of attention would be constant if we blocked the possibility of verbal rehearsal. Such a pattern of results would indicate that the attenuation of visuospatial attention allocation likely reflects a process of recoding between the visual and verbal representations. To this end, we used the selective interference method described above and asked participants to rehears verbal information during the same visuospatial working memory task. Strikingly, when we blocked the phonological loop we observed that attention was allocated equally well at short and long maintenance intervals. These results show that inherently visuospatial information is recoded into the phonological loop which in turn suggests that the independence of two stores is limited [Leszczyński et al., 2012].

1.2.4 Central executive.

The central executive is a core part of the multi-component model. Yet our understanding of this component is relatively poor. It was initially designed as a general processing unit involved in processing of complex rules which have no sensory properties [Baddeley, 2003]. One popular approach is to understand the central executive in the context of the Norman and Shallice [1986] model of attention control. This model assumes that behavior is influenced by two mechanisms. First, it might be controlled by habitual patterns and schemes implicitly cued by the environment. Second, it might be also controlled by a limited capacity control system which interrupts routines and the schematic behavior. Norman and

Shallice named this attention system the supervisory activating system (SAS). A support for such an interpretation of the central executive was provided by studies with patients who suffered from the frontal lobe lesions. For example, Godefroy and Rousseaux [1996] compared performance of patients with the frontal lobe deficit to that of healthy controls. The authors presented participants with auditory and visual stimuli. The participants were required to respond to the stimuli of one modality and ignore the other one. The patients with frontal cortex lesion made more responses to the stimulus presented in the irrelevant modality (the so called commission errors) compared to the control group. These results suggest that the frontal cortex is important for controlling the allocation of attention to the auditory and the visual system. The association with the central executive of the memory system is therefore very tempting. However, it is important to note that Baddeley [1998] himself opposed against associating the frontal cortex with the central executive. He claimed that the central executive should not be related to a single brain region but rather to a class of processes and functions which itself is not unitary and cannot be explain by lesion to single cortical area [for a comprehensive review see also Andres, 2003].

When I asked you to recall paintings on the walls of a room you are familiar with, the visual information you experience was recalled form the episodic store (unless you were sitting in that room at this very moment) into the visuospatial sketchpad. This information consist of what you have previously experienced. It is bound together and stored in the form of a memory episode. How this information is retrieved and brought back to the visuospatial sketchpad or to the phonological loop was unexplained with the initial multi-component model of working memory. Similarly, the model was not able to explain how it is possible to experience and maintain objects as multi-modal phenomena. For example a visual experience of a sports car and an auditory experience of its engine require an interface between these two types of information which the initial model was missing. The necessity to account for such a phenomena, led Baddeley to extend the multi-component model and to introduce the episodic buffer [Baddeley, 2000]. This is a third storage system which provides interactions between both WM and LTM, as well as

between the visuospatial sketchpad and the phonological loop. This interface allows explaining how it is possible to store and bind together information in a form of multi-sensory episodes. The introduction of this episodic buffer system allowed a resolution to the state of aporia in which the model assumed no interaction between visuospatial and articulatory systems. Such an assumption strongly violated our everyday observations that we indeed do experience surroundings in a multi-sensory form.

1.3 State-based models of working memory

The multi-component model has been very prominent throughout working memory research. It brought a new perspective and shaped the understanding of working memory over the last forty years. Recently, however, a new type of models (the so called state-based models; sometimes also refereed to as the embedded-processes models or concentric models) have become increasingly popular [for reviews see D'Esposito and Postle, 2015, Eriksson et al., 2015]. Unlike the multi-component model they conceptualize information maintained in working memory as a dynamic state (i.e. the state of being in the spotlight of attention) rather than as a passive structure. These models assume that allocation of attention to an internal representation is a mechanism responsible for maintenance of information in the working memory. The internal representation might be of a different kind – i.e. semantic, episodic or sensory. These models brought an important conceptual shift. Importantly, they are largely consistent with the Baddeley-Hitch model. They offer however a more general information processing framework and therefore a more parsimonious explanation of the data. In particular, terminology used to formulate these new set of models is broader and more inclusive.

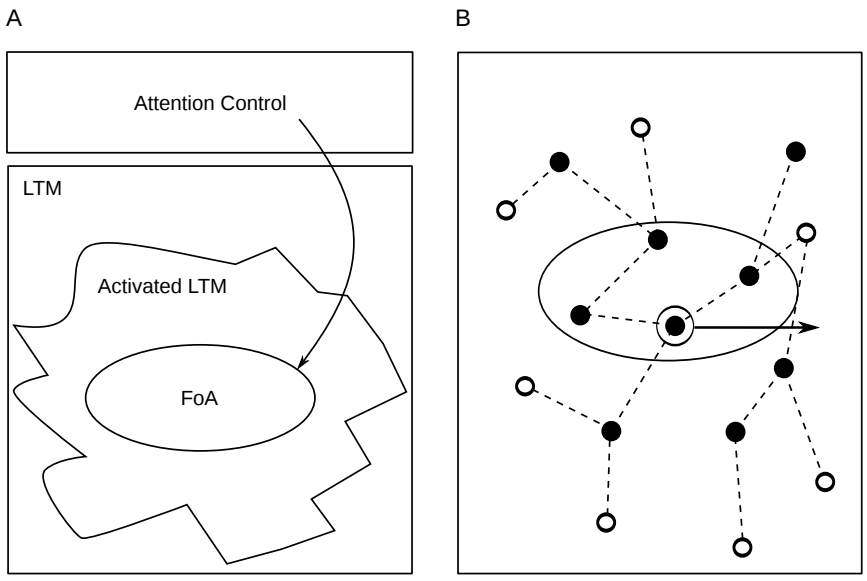

FIGURE 1.2: Two prominent examples of the state-based models. (A) Cowan's model which posits that a subset of long-term memory (LTM) might be directly accessed by the Focus of Attention (FoA). Adapted from [Cowan, 2008]. (B)The model proposed by Klaus Oberauer decomposes the FoA into a larger region of accessed items (large oval) and the region of a direct access where only a single item is activated (indicated by small circle). Dots linked by the dashed lines represent a network of the LTM representations some of which are activated [indicated by black lines; Adapted from Oberauer, 2002].

1.3.1 The temporary activation of the LTM representation.

One prominent example of the state-based class of models originates from research on the temporary activation of the long-term memory representation [Cowan, 1995, Oberauer, 2002, Ruchkin et al., 2003]. It is assumed that the STM and the LTM are different states of the same representation. In particular, the model suggests that the activated long-term memory representation is sufficient to explain short-term memory properties with no need for separate subsystems to conceptualize storage across different time scales [Crowder, 1993]. Cowan [1995] suggested that working memory maintenance might be conceptualized by two states: the focus of attention (FoA) and activated representation of the LTM. Few items might be held in the focus of attention with a limit of about four [Cowan, 2001]. If

attention is shifted to another information, the representation of information previously available in the FoA transforms into LTM representation which is virtually unlimited but susceptible to interference and temporal decay. Consequently, the capacity limit in working memory results from the limited number of items that are simultaneously attended. In a more recent version of this model Oberauer [2002] distinguishes two forms of FoA which correspond with two distinct functional states of the representation in WM. First, the activated part of LTM is accessible indirectly through associations that are within the current focus of attention. Second, the region of direct access keeps the number of items available to potential access in the ongoing cognitive process. Finally, the focus of attention keeps just one item available for the next cognitive operation. Despite the differences regarding interpretations of the FoA these models share the very same idea that the WM is a state consisting of the activated LTM part and the focus of attention, which selects chunks of information from the LTM and makes it available for other cognitive processes (see Figure 1.2 A and B).

The main advantage of this new model is its simplicity and the reduced number of assumptions needed to explain the same results. For example, the concept of *activated memory* is very inclusive and refers not only to the phonological and the visuospatial information but also to other cognitive and sensory modalities like memory, tactile or gustatory. Even the extended Baddeley-Hitch model with the episodic buffer would need to introduce additional slave systems to account for these additional modalities. The state-based models do not require such additional assumptions because the idea of the activated memory is inherently mulitmodal. This makes the old model less parsimonious compared to the state-based models. Similarly, the interaction between working memory, long-term memory and attention might be relatively easy to reconcile within the Baddeley-Hitch model but it would require additional assumptions. The studies on attention and working memory have been particularly important to illustrate the parsimony of the new model. Many of these experiments use a cueing approach well-known from visual research. Participants are presented with two sets of stimuli (for example a set faces and a set of houses or a list of words printed in red and a list of words printed

in green). During maintenance interval a cue is presented which indicates the set of stimuli relevant for the retrieval phase [for two prominent examples see Lepsien and Nobre, 2007, Oberauer, 2001]. There are many interesting questions regarding the influence of attention on memory representation. For example, what happens to the uncued stimuli? Are they forgotten and removed from memory or are they still available? The state-based model suggests that attention may flexibility shift between various elements of the memory content (see Figure 1.2). Indeed, psychophysics [Oberauer, 2001] and neuroimaging [Lepsien and Nobre, 2007] studies consistently showed that information stored in working memory may be selectively modulated by allocating attention to and away from the relevant representations. Oberauer studied the influence of the uncued (i.e. irrelevant) set size on reaction times and observed that it takes approximately 600ms to remove these items from the current focus of attention. The study by Lepsien and Nobre [2007] further showed that information removed from the immediate focus of attention is not completely lost and may be still retrieved if attention at some later point is cued towards this representation.

1.3.2 The hypothesis of sensory recruitment.

The sensory recruitment version of the state-based model assumes that mechanisms involved in the sensory representation of a perceived information are also engaged during maintenance of this information in the working memory [for review see Awh and Jonides, 2001, Pasternak and Greenlee, 2005]. This is a particularly interesting version of the state-based model because it assumes that WM information is maintained with a shift of attention to representations in the sensory cortices. One important point of support for the sensory recruitment model comes from research on sustained attention during retention interval. Awh et al. [2000] found that a location in WM is maintained by the covert allocation of attention to this location during the retention interval. In another study [Harrison and Tong, 2009] used an approach to decode content specific representations from the fMRI (functional Magnetic Resonance Imaging) signal and found that the high fidelity

information about the maintained color and orientation both are stored in the primary visual cortex. Along the lines of this model an EEG research including my own have shown that the retrieval of information from WM results in an allocation of the visual attention towards a previously memorized location even if the task does not require it [Astle et al., 2010, 2009, Kuo et al., 2009, Leszczyński et al., 2012, 2013]. The common finding across all these studies is that maintenance and retrieval of information from working memory is associated with attention orienting towards a representation stored in memory. This corroborates the state-based model by showing the importance of visual attention during memory.

1.4 Cognitive models of capacity limits

1.4.1 The magical number 7 ± 2.

In the mid half of the last century, Georg Miller published his groundbreaking article which started a new line of psychological research on the limits of WM capacity [Miller, 1956]. First, Miller suggested that there is a "limit to accuracy with which we can identify absolutely the magnitude of a uni-dimensional stimulus variable" [Miller, 1956, p.8]. He implied that there is a limit to the number of items of a given kind that might be accurately maintained over short intervals. Second, he introduced a notion of a *chunk* which expresses an interesting observation that simple items can be grouped and merged into a form of a more complex units. The concept of the *chunk* was introduced to understand the results observed by authors like Hayes [1952] who found that a memory span changes as a function of information input. In details, he presented participants with different kinds of materials (e.g. binary digits, letters and words) at the constant rate of one item per second. Hayes observed that the amount of information in the memory span is not constant at about 2.6 bits as was previously believed but depends on the input information. For example, a four digit sequence 1-9-3-5 might be grouped into one number "1935" a sequence of letters M-A-R-C-I-N might be grouped into a word "MARCIN". These groupings are probably best observed in a natural

language where letters are effortlessly chunked into words and words are grouped into phrases and sentences. Consequently, Miller points out that the short-term storage although limited to about seven chunks, is not rigid but rather prone to strategies like chunking and grouping which increases the number of maintained bits of information [Hayes, 1952, Miller, 1956].

1.4.2 The magical number 4 ± 2.

Nearly fifty years later Cowan [2001] reviewed studies on the memory capacity limits and came to a conclusion that the number of 7 ± 2 items suggested by Miller is overestimated. The author proposed a more standardized way of estimating capacity limits. He reasoned that the possible factors biasing the estimate are chunking and recoding of information to other formats. Therefore, carefully designed experiments should limit these possibilities. The change detection is an example of such a controlled experimental condition which was used by Cowan himself to estimate the capacity.

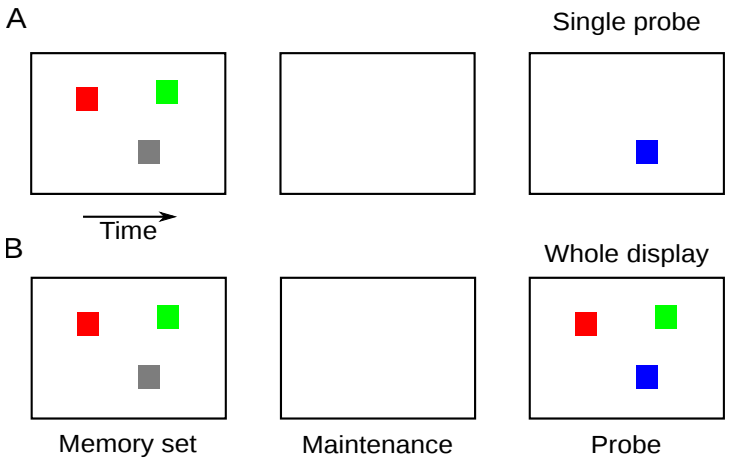

FIGURE 1.3: Two popular examples of the change detection task. Initially a memory set is presented with the number of colors patches as a memory set (here the memory load of 3 is presented as an example). Subsequently, a blank screen follows. Finally a probe is presented. Two most popular ways of probing memory is a single probe (A) and a whole display probe (B). In the former only one location is probed whereas in the later the full set of memory items is presented. Participants in both need to indicate if the color of the probe matches or not the color of items in the memory set.

According to Cowan [2001], there is an abundant evidence showing a lower capacity limit. If chunking possibility is limited (as for example in the color change detection task; see Figure 1.3) experimental evidence shows capacity limit around three to five items. Consistently with his working memory model he implied that the capacity limit originates from a restricted spotlight of attention which can only hold about four items at a time.

1.4.3 The change detection task.

Past research on working memory was dominated by the verbal content (like digit or letter spans as stimuli material). Recently, research on working memory has shifted from the verbal to the visual domain [Luck and Vogel, 2013] with simple (colors, orientations, abstract shapes) or complex (natural scene images of faces, houses, and landscapes) visual features. One of the main reasons for this shift is an increased popularity of the change detection paradigm which allows for a reliable estimation of the VWM (visual working memory) limits (see Figure 1.3). The change detection task typically involves initial presentation of a memory set which might consist of one or more colored squares. Subsequently, visual mask (e.g. white noise) might be shown followed by a retention interval. Finally, a probe is presented which consists of either the same stimulus array as was initially presented or a modified one. In a series of experiments Luck and Vogel [1997] showed that the VWM is limited to about four items. Importantly, the limit was similar irrespective of whether participants were maintaining set of colors, orientations or both. This last result was particularly interesting because it suggested that participants store integrated objects rather than the individual features.
The change detection experimental procedure allowed not only to measure capacity, but recently also the precision with which an item is memorized. This is achieved by continuous manipulation of colors on a color circle during memory encoding and retrieval [for review see Ma et al., 2014]. Participants are presented with a number of color patches to remember. After retention, the probe is displayed which is an individually adjusted by the participant color on the color

circle. Note that unlike the two alternative forced choice (2AFC) protocol this task allows measuring accuracy as a continuous, circular variable.

1.4.4 Are the working memory resources discrete or continuous?

Two models have been suggested within the change detection framework to explain the limited resources of the visual working memory – the discrete slots [Zhang and Luck, 2008] and the continuous resource model [Bays and Husain, 2008, Ma et al., 2014]. Both theories assume that only a limited number of items ($K_{\max}$) might be stored in the visual working memory. However, they have different explanations concerning the mechanism of this limit. According to the slot-based hypothesis, if the number of items in a visual input exceeds the $K_{\max}$, than only $K_{\max}$ number of items is stored in working memory and no other representation is maintained. The model assumes that items are selected according to their priority maps. The resource-based model in contrast assumes that the capacity limit is flexible, so does the $K_{\max}$. When a visual input exceeds the $K_{\max}$, resources are flexibly allocated across all the presented items with less resources per single item. While the slot model anticipates that items exceeding $K_{\max}$ are not stored at all, the continuous resource model predicts that these items are indeed stored but with reduced overall precision.

The results up to date are mixed and support both models [for reviews see Luck and Vogel, 2013, Ma et al., 2014]. In further experiments, Zhang and Luck [2008] presented their participants with an array consisting of several color patches randomly distributed across the display. After a retention interval of 900ms, participants were asked to indicate a color of a probed patch. Importantly, the probe was continuous and the participants' task was to respond to whichever color was presented by clicking on a color wheel. This continuous report task allowed the authors to estimate the precision with which items were stored in the visual working memory by estimating the distribution of response errors (the circular difference between the actual and the reported color). The authors hypothesized that if a

probed item is stored in WM then the error should be small and normally distributed around the correct color. If however, a probed item is not stored, they expected errors to be uniformly distributed across the wheel of colors. Interestingly, the authors observed that an increasing number of memorized items resulted in a vertical shift of the whole distribution but no significant change in its width. This suggests that increased memory loads decrease the probability that a probed item is stored, but does not reduce the precision of stored representation which supports the slot-based model.

Bays and Husain [2008] also presented their participants with a display consisting of color patches. After a short retention interval of 500ms, a probe was shown. Here, the probe was one of the memorized stimuli and observers responded to whether it was displaced to the right or left with respect to memorized location. Here, the slot and resource models differ regarding predictions for the large memory set sizes. In particular, the slot model predicts that participants will make errors when the set size is large (i.e. nothing is stored apart from information in small number of fixed slots), irrespective of the size of displacement. The continuous resource model predicts that participants would still perform well for the large set sizes given that the displacement is large enough. The authors indeed observed very good performance for large set sizes, particularly a during larger displacement which supports the resource model. Altogether the results up to now are mixed and in favor of both models.

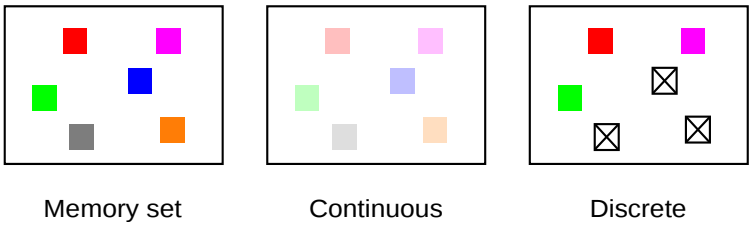

FIGURE 1.4: Two models of the visual working memory resources. The difference between the continuous resources model and the discrete slots model. Adapted from [Luck and Vogel, 2013].

1.5 Maintenance of multiple-items

Another set of classical studies were conducted by Saul Sternberg in the late 1960s [Sternberg, 1966, 1969]. He briefly presented his participants with a number of items (e.g. digits) followed by a maintenance interval during which participants were required to maintain the presented stimuli. Subsequently, the author showed a probe and participants were requested to respond as fast as possible whether the probe matched any of the stimuli in the memory set. Sternberg measured the participants' reaction times (RT), defined as a time needed from presentation of the probe to pressing a button indicating the response. Importantly, he manipulated the number of items in the memory set. He observed that the RT increased linearly as a function of memory set with a rate of increase at about 38ms per item [Sternberg, 1966]. Sternberg suggested that the time required to respond should reflect the total time needed to (1) form a perceptual representation of the probe; (2) compare the probed representation with memorized set; (3) make a binary selection between "match" and "non-match" and (4) execute a response. Since all factors but (2) were constant across conditions he reasoned that the increase in RT reflects rising demands in comparing the probed representation with the memory set. In particular, he suggested that information held in short-term memory is accessed serially (one-by-one) at a high speed. Given that the rate of increase in RT was about 38ms, Sternberg suggested that this is the amount of time needed to access each item. Thirty years after Sternberg's discovery a multiplexing model of WM (see chapter 2.3) has been proposed. This model explains the linear increase in the reaction time as well as the neurophysiological and the computational mechanisms behind the linear increase in RT.

1.6 Summary

Cognitive models of the working memory are required to account for several of its principal features like limited capacity, flexibility and importance for other cognitive functions. Early models assumed that WM was a hierarchical structure.

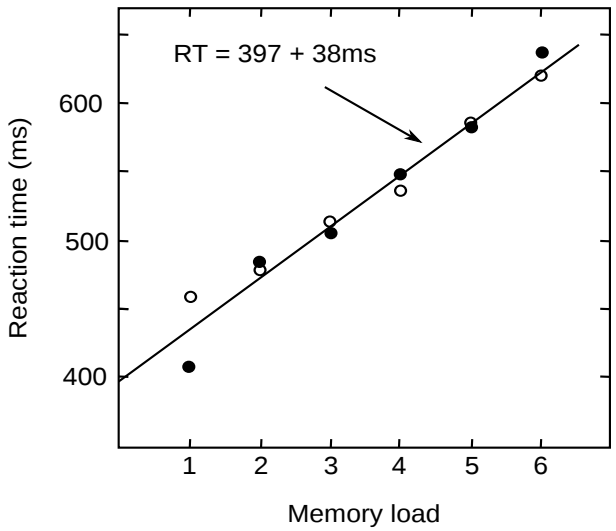

FIGURE 1.5: The linear increase in reaction time across memory load. The results from Sternberg experiment show a linear increase in reaction time across memory loads. Importantly, the rate of increase is constant at about 39ms per item. Adapted from [Sternberg, 1966].

Limited capacity was a feature inherent to its storage elements. Flexibility was introduced by a meta-level structure (the central executive) that controlled the so called "slave systems". The contribution to other functions was achieved by assuming interactive connections across components. Newer models conceptualize WM as a state rather than as a structure. They fuse boundaries between WM and other cognitive functions by assuming that the maintenance is achieved by allocating attention to long-term or sensory representations. They also offer a more parsimonious explanation of the limits to WM assuming it results from a limited focus of attention.

Chapter 2

Neuronal models of working memory

2.1 The distributed cortical network supports working memory

In the previous chapter I described the working memory as an interactive cognitive function involving multiple processes like storage, attention allocation, control of attention, long-term memory. As it is unlikely that all these processes are performed by a single cortical area, one may expect that working memory relies on multiple brain regions. Indeed, studies have shown that distinct brain parts are involved in working memory [for review see Eriksson et al., 2015, Fuster and Bressler, 2012]. In particular, the maintenance has been suggested to involve multiple brain regions (see Figure 2.1). Owen et al. [1996] compared performance of patients with frontal lobe lesions to the performance of three other groups: patients lesioned to the temporal lobe, and patients who had undergone an amygdalo-hippocampectomy, as well as healthy controls. The authors measured performance in two types of task: spatial and verbal working memory. They observed that patients lesioned to the frontal lobe had difficulties performing the spatial working memory task but not the visual one (a WM task that relied on maintenance of

shapes and colors). Crucially, lesion to the temporal cortex affected visual working but not spatial the WM. This suggests that the frontal patients have difficulties specifically during the maintenance of spatial information and temporal patient have feature specific deficit in maintenance of non-spatial visual information. Similarly, Pisella observed that parietal patients showed a very specific impairment in performance during maintenance of location compared to maintenance of color and shape task [Pisella et al., 2004].

This suggests that WM involves a distributed network and that distinct parts of the brain contribute to the maintenance of distinct WM features. In other words these effects show content specificity – patients with lesions located in a given part of the WM network are impaired during maintenance of a very particular type of material. This distributed pattern and cortical specificity is important because early animal models of working memory although recognized that distinct materials may be maintained by separate cortical networks but only limited this specificity to distinct networks in the prefrontal cortex (PFC).

In the next parts of this chapter I will focus on studies investigating a role of the lateral PFC, the hippocampus as well as the sensory cortices each of which is particularly important for the results described in the current thesis.

2.2 The search for the WM code in the lPFC

The lateral prefrontal cortex (lPFC) has been traditionally viewed as a brain structure central for the maintenance of information in WM [Curtis and D'Esposito, 2003, 2004, D'Esposito, 2007, Sreenivasan et al., 2014, Stokes, 2015]. One prominent theory [Goldman-Rakic, 1995] posits that persistent activity of neurons in the lPFC is the underlying neural mechanism for WM storage and reflects maintained information itself [Goldman-Rakic, 1995]. This early model suggests further that distinct parts of the lPFC are involved in maintenance of different types of materials. More recent views suggest that high-fidelity WM representations are stored in modality specific sensory areas rather than in the lPFC. Activity in the lPFC in

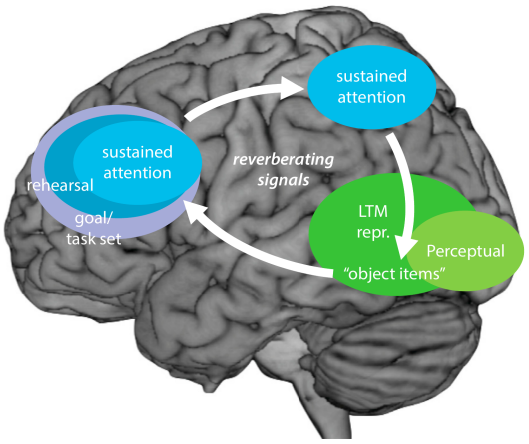

FIGURE 2.1: The distributed network involved in the visual working memory maintenance. Processes involved in working memory maintenance are distributed across the number of cortical regions including frontal (important for maintenance of task set, attention guidance), as well as parietal cortices (relevant for attention allocation to internal representation and maintenance of spatial information) and other sensory areas (i.e. the dorsal visual stream, primary visual cortex or the MT+ involved in maintenance of relevant features). Adapted from [Eriksson et al., 2015].

turn has been suggested to reflect attention allocation to an internal representation [Curtis and D'Esposito, 2003, Eriksson et al., 2015, Sreenivasan et al., 2014]. Another hypothesis suggests that dynamic coding in the lPFC reflects more abstract representations like a response set or task contingencies [Meyers et al., 2012, Stokes, 2015, Stokes et al., 2013]. These abstract representations might be used to bias stimulus-specific information in the sensory brain regions. Both of these more recent views suggest that local and large-scale network interactions across brain areas are essential for understanding the WM.

2.2.1 The signal intensity approach.

Early studies on the monkey model of WM investigated overall signal (i.e. spiking rate) change across the delay period (see Figure 2.2A). The elevated activity was assumed to reflect the storage of information. Many neurons in the lPFC have

indeed been observed to fire in sustained way during the working memory retention interval [Goldman-Rakic, 1995, Kubota and Niki, 1971]. Research conducted on lesioned monkey lPFC indicate impaired WM performance while leaving other functions intact [Funahashi et al., 1993]. These results led Goldman-Rakic to propose a model in which the lPFC and its dorsal part in particular (dlPFC; BA46) is critically involved in representing sensory information during WM delay (time interval when no stimulus is presented). The sustained firing has been suggested to reflect storage of relevant features [Funahashi et al., 1989, Goldman-Rakic, 1995]. Studies measuring single unit activity in the dlPFC reported that the WM delay signal changes parametrically with the duration of maintenance interval as well as the number of items maintained [Funahashi et al., 1989, Fuster and Alexander, 1971]. In particular, Funahashi and colleagues have shown that duration of persistent firing increases with increasing delay interval. The approach which looks at signal intensity was very productive with many studies investigating the lPFC during memory delay. Studies within the signal intensity framework led researchers to develop two models describing the functional specialization of the lateral prefrontal cortex. The *content specific model* posits that the dorsolateral PFC (dlPFC) is involved in maintenance of spatial information (i.e. locations) whereas the ventrolateral PFC (vlPFC) is involved in maintenance of non-spatial (i.e. objects) information [Curtis and D'Esposito, 2004, Levy and Goldman-Rakic, 2000]. This model in essence assumes that the dlPFC and the vlPFC are functional extensions of the dorsal "where" and the ventral "what" visual streams [Goodale and Milner, 1992, Milner and Goodale, 1995].

Funahashi with collaborators trained monkeys to remember the spatial location in periphery of the visual field and subsequently to make a saccadic eye movement to the memorized location. They observed that lesions to the dlPFC disrupted memory-guided eye movements to the spatial location where the target was presented [Funahashi et al., 1993]. Passingham also trained animals in a working memory task but instead of probing a location he probed the color feature. He observed that monkeys lesioned in the dlPFC performed equally well before and

after the surgery. In contrast, animals lesioned to the vlPFC were severely impaired [Passingham, 1975]. These results support the content specific view of the lateral PFC with the dorsal part being relevant for maintenance of the spatial information and the ventral part for the non-spatial materials.

In turn the *process specific model* postulates that the dlPFC and the vlPFC contribute to distinct processes during a working memory task. In particular, the dlPFC was hypothesized to contribute to monitoring and manipulation of the stored information. The vlPFC in turn was suggested to play a role in encoding and retrieval [Curtis and D'Esposito, 2004]. Supporting evidence for this model comes from a study by Owen and colleagues who used a positron emission tomography (PET) to study the cerebral blood flow in human subjects performing two different tasks. One of the tasks involved maintenance and subsequent reproduction of several items. The other task required continuous updating of maintained information. An increased blood flow was observed in the dlPFC but not in the vlPFC during the maintenance of information. The reversed was observed during updating of information in WM [Owen et al., 1999]. Since both experiments used spatial material these results were interpreted as supporting the process specific account of the lateral PFC.

In contrast to these early studies conducted on the monkey model of WM, research looking at human subjects lesioned in the lPFC have shown no detrimental effect to working memory span [D'Esposito et al., 2006]. Furthermore, research using fMRI in human subjects have demonstrated that maintenance of increasing number of items does not necessary lead to increased response in the dlPFC [Jha and McCarthy, 2000, Rowe et al., 2000, Rypma and D'Esposito, 1999]. For example, Rowe et al. [2000] used fMRI to distinguish two sources of activity during a working memory task. One associated with maintenance itself and the other related to selection of information from the memory. Importantly, the maintenance related activity was observed bilaterally in the BA8 and intraparietal cortex but not in the dlPFC (BA46). The selection of a target from the memory in turn has been observed to engage the dlPFC. The authors concluded that the dlPFC is indeed

involved in WM but it rather contributes to some other function instead of representing sensory information.

These findings lead to the development of a model which suggests that sustained activity in the dlPFC reflects more abstract features (i.e. task-set, executive control or attention to an internal representation). This model presumes that the dlPFC plays a role in the maintenance of task-relevant contingencies. It controls and top-down modulates a broad range of networks including sensory areas where the high-fidelity representations are stored and motor areas where responses are executed [Curtis and D'Esposito, 2003, Miller and Cohen, 2001, Smith and Jonides, 1999].

2.2.2 Multivariate coding approach.

Increased availability of the computing power facilitated usage of the multivariate models in the analysis of neuroscientific data [Haynes and Rees, 2006, Norman et al., 2006]. In the context of WM this allowed for conducting experiments which study the coding properties of large networks and investigate the high-dimensional dynamics of active neural representations which are hold in memory during the delay period. Importantly, this approach provides an evidence that content related activity is not necessarily reflected by the sustained and elevated global response as previously assumed [for an excellent review see Postle, 2015]. For example, studies which used multivariate decoding approach reported the stimulus-specific activity during the delay period in the occipital cortex even if no sustained signal intensity change was observed in that brain area [Harrison and Tong, 2009, Serences et al., 2009]. Harrison and Tong [2009] presented their participants with orientation gratings to memorize. A cue that followed indicated which of the two gratings to be retained. The authors used fMRI with the pattern classification approach and observed orientation specific activity even in the absence of above baseline global BOLD (blood-oxygenation-level dependent) response. Furthermore, these activity patterns in the areas V1-V4 predicted the memorized stimuli. Importantly, the orientation selective patterns were prolonged during the delay interval.

The study shows that the fine-tuned information is sustained in the sensory areas during memory maintenance even if the overall activity change is low. Lee et al. [2013] used a similar approach to show that the extrastriate areas contain information about the visual whereas the lPFC about non-visual and abstract features of the very same stimuli maintained in memory. These results further corroborate the concept of the lPFC representing abstract features and lower-level and the modality specific cortices maintaining high-resolution sensory information. It also supports the notion that the brain areas specialized in object recognition and perception contribute to maintenance of these objects. Furthermore, these representations are stored in highly distributed manner across many interconnected networks [Sreenivasan et al., 2014]. The multivariate approach has also shown that the lPFC activates independently from a WM task. Its activity was observed during a broad range of cognitive functions, indicating that a role of the lPFC is not specific for WM [Meyers et al., 2008, Stokes et al., 2013, Wallis et al., 2001]. Altogether, findings suggest that the lPFC is sensitive to representations of multiple simultaneously maintained goals and it may play a role in top-down control over representations stored in a distributed network including object recognition and sensory cortices. Signals in the lPFC might provide a source for biasing attention between representations at these lower-sensory levels. The precise mechanism of this influence remains unknown.

These studies motivated the development of the *dynamic coding* framework which suggests that WM activity does not necessary depend on a global, stationary and persistent neural response but rather on high dimensional interactions across large neural populations [Stokes, 2015]. It is sometimes refereed to as the "activity-silent" WM to emphasize that there is no need for the signal intensity change in order to code information in WM (see Figure 2.2B). Such an "activity-silent" WM has been hypothesized to represent information in a form that does not demonstrate itself on a global level (i.e. no sustained response). The "activity-silent" WM network rather stores information as a change to a dynamic state of a population response or as dynamics in the functional connectivity (see Figure 2.2B). One candidate for a mechanism of the dynamic coding in WM is the frequency

specific coherence. In particular, dynamic shifts in the functional connectivity might be achieved by phase aligning networks of increased excitability which in turn increases the likelihood of information transfer [a phenomenon known as communication through coherence: Fries, 2005, 2015]. Indeed, Buschman et al. [2012] observed that the lPFC frequency-specific coherence interactions represent various rules relevant for guidance of behavior. The important conclusion from studies using this approach is that the WM specific information may be coded even when no signal intensity change is observed.

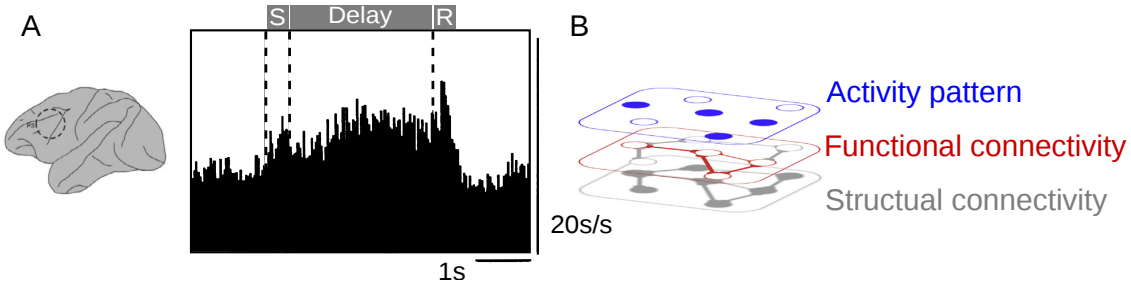

FIGURE 2.2: Signal intensity and activity-silent approach to working memory. (A) An example of the signal intensity approach. The neurophysiological recordings from the monkey prefrontal cortex showing neural activity during a WM task. The histogram presents a sum of neural activity for preferred stimulus averaged across 46 neurons in the principal sulcus. (B) A hypothetical layered neural state which consists of an activity-silent state (i.e. the activity pattern which does not fully manifest itself in the structural connectivity). The dynamic coding model suggests that changes to these silent-states could play an important role in WM. Adapted from [Funahashi et al., 1989, Stokes, 2015].

2.3 The neural mechanisms of capacity limits

Studies with non-human primates revealed sustained neural response during the maintenance interval across many neurons in the inferotemporal cortex [Fuster and Jervey, 1982, Miller et al., 1993]. An analogous sustained brain response has been observed in the human EEG recorded over the posterior parietal and occipital sites. It has been called the contralateral delay activity (CDA) because it is observed during the delay period on the side contralateral to the location of memorized material [Vogel and Machizawa, 2004, Vogel et al., 2005]. In particular

it is measured at posterior parietal, lateral occipital and posterior temporal electrode sites contralateral to the location of memorized stimuli. This suggests that an underlying neural mechanism is anatomically unspecific and rather a global phenomenon observed across many brain networks. Vogel and Machizawa [2004] observed that the amplitude of the CDA increases with increasing working memory set size. Importantly, the amplitude rises only up to about four items and saturates around an individual WM limit. The authors observed that individual WM capacity limit was strongly positively correlated with the increase of CDA's amplitude from memory load 2 to memory load 4. Vogel et al. [2005] further observed that the individual limit in the number of items memorized in WM (indexed by the CDA), strongly depends on the efficiency in selecting the relevant information. In fact, participants with high and low WM capacity were able to maintain about the same number of items. However, the former group of subjects was more efficient in selecting the relevant items (i.e. ignoring irrelevant distractors). In contrast, the low-capacity participants although capable to memorize a similar number of items were more likely to maintain both relevant and irrelevant items.

Another model describing the limits in WM capacity has been proposed by Lisman and Idiart [1995]. Unlike Vogel's approach it relies on the dynamic interactions between theta and gamma oscillations (see Figure 2.3A). The authors suggested that information might be sustained by increased membrane excitability repeated on each cycle of a network oscillating with the theta (5-12Hz) frequency. According to the model each memory item is stored as a neural assembly firing within the gamma cycle whereas the whole memory set is refreshed with the theta-alpha oscillation (5-12Hz). The model is able to explain several observations reported by the psychophysics experiments. For example, it explains the limit in WM capacity to about 7 ± 2 items as observed by Miller (see chapter 1.4). According to the model this limit reflects the ratio between gamma activity ($\sim$40Hz; representing individual items) and theta-alpha oscillations (5-12Hz; refreshing memory set). This ratio between frequencies imposes a limit in the number of items that might be refreshed in a single low-frequency cycle. The items represented by cell

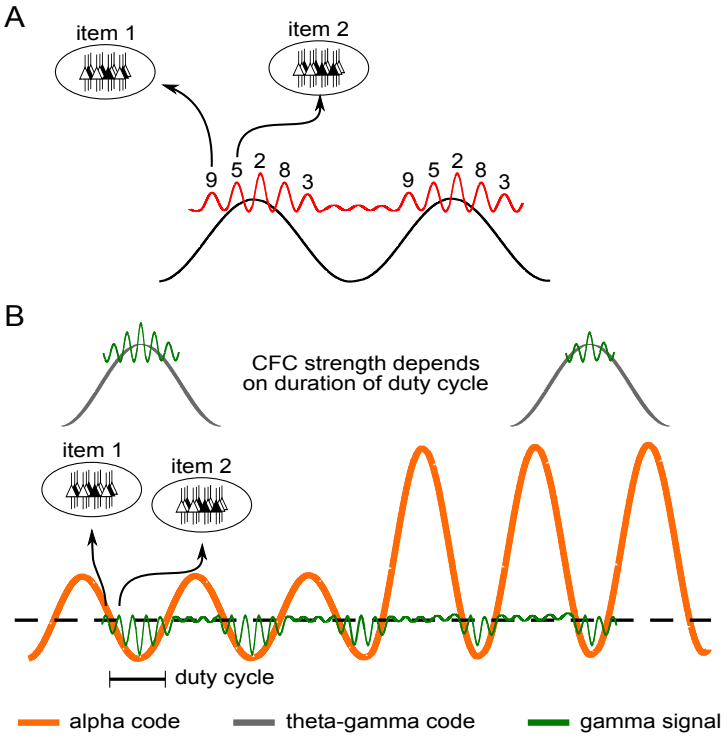

FIGURE 2.3: The multiplexing model of working memory. (A) The initial multiplexing model developed by Lisman and Idiart [Lisman and Idiart, 1995]. The model suggests a coupling between the phase of low-frequency oscillations in theta and alpha (5-12Hz) range and the gamma oscillations (∼40Hz). Each item is represented as a neural assembly synchronized in the gamma cycle. Consecutive gamma cycles represent corresponding items in a memory sequence. The model further suggests that the whole memory set is reactivated during successive theta cycles. (B) Suggested extension of the multiplexing model. The strength of CFC is not uniform during maintenance and depends on the length of the alpha duty cycle. The length of the duty cycle relies on power in the alpha frequency range. Low alpha power (orange on the left) results in longer duty cycles, whereas high alpha power (orange on the right) is associated with shorter duty cycle. The length of the duty cycle constrains gamma power (indicated by green oscillations), which in turn may be reflected as modulations to the theta-gamma code. In particular, long duty cycles may be observed as decreases in strength of the theta-gamma modulation index (modulating activity is broadly distributed across the low-frequency phase; upper left). In contrast, short duty cycles may be reflected by increases in CFC strength (upper right).

assemblies firing after the preferred window of membrane excitability will not be reactivated on consecutive theta-alpha cycles and therefore will fade away. Thus, the model predicts the inter-individual correlation between WM capacity and the oscillatory signal properties. In particular, the longer the preferred window of membrane excitability (indexed by the frequency of individual theta dominant oscillation), the higher individual WM capacity. Alternatively, one might also expect that the faster synchronization of neural assemblies representing individual items (indexed by the dominant frequency of gamma oscillations), the higher individual WM capacity. Furthermore, the model explains the results of Sternberg experiments on the maintenance of multiple items in WM (see chapter 1.5). If representation of each item is stored by cells firing in the gamma cycle and individual items are reflected by firing at different phases of the theta-alpha oscillations then larger memory set size increases the number of cell assemblies locked to the consecutive theta-alpha phases. In other words, larger memory load is associated with increased portion of the slow oscillation which already is associated with a gamma activity representing individual items. This chain of representations increases exactly by the duration of gamma cycle with each item in the memory set. This explains Sternberg's observation that the reaction time increases linearly at about 38ms per memory item. Note that this explanation requires an additional assumption that items are recalled serially one-by-one which indeed was proposed by Sternberg. The multiplexing model of WM offers several clear-cut predictions. The number of gamma sub-cycles within the theta cycle during WM maintenance should scale with number of items maintained. The distribution of gamma activity across theta cycle should be non-uniform, resulting in the cross-frequency coupling (CFC). Indeed, several studies observed increased CFC during WM maintenance [Axmacher et al., 2010b, Chaieb et al., 2015, Leszczyński et al., 2015, Sauseng et al., 2009, Siegel et al., 2009] corroborating the concept of the CFC as a potential mechanism for storing multiple items in WM.

2.4 PFC vs. MTL in memory: a failed dichotomy?

2.4.1 The traditional model of memory processes.

The famous case of the patient HM (Henry Molaison) who suffered from the pharmaco-resistant temporal lobe epilepsy and was surgically treated with a bilateral hippocampal resection has driven research on memory for many years [Corkin, 2002, Scoville and Milner, 1957]. After removal of both hippocampi he developed an anterograde amnesia and lost the ability to register new episodic memories. In contrast, his working and procedural memory were intact. This clear dissociation between the long-term declarative memory and short-term working memory was so striking that it made a foundation for the classical taxonomy of memory processes [Gazzaniga et al., 2002, Squire and Zola, 1996]. The long-term declarative memory has been suggested to depend critically on the medial temporal lobe (MTL) and the hippocampus in particular but not on the PFC. Working memory in contrast has been exclusively associated with the prefrontal cortex but not with the MTL or hippocampus. Assuming a double dissociation between these two cognitive functions and the brain structures this model has been very productive in guiding research for decades. However, recent experiments both in human subjects and rodents have questioned the main assumptions of this model in two ways.

2.4.2 Aporetic findings.

First, there are other non-declarative forms of episodic memory [i.e. statistical learning: Schapiro et al., 2014, 2012] and more general cognitive processes like object representation [Schapiro et al., 2014] or decision making [Wimmer and Shohamy, 2012] which have been identified to relay on the hippocampus [for review see Shohamy and Turk-Browne, 2013]. These results are difficult to explain within the classical model which expects the hippocampus to exclusively contribute to the episodic memory [Squire and Zola, 1996]. Second, and for the purpose of this thesis more important, working memory has been also found to depend on the

hippocampus. In particular studies using neuroimaging [Nichols et al., 2006, Ranganath and D'Esposito, 2001], intracranial EEG [Axmacher et al., 2010b,c, 2007, Leszczyński et al., 2015, van Vugt et al., 2010] have shown that the hippocampus activates during WM maintenance of multiple items. Furthermore, together with my collaborators I have found that the relevant parameters of EEG signal recorded from the hippocampus predict individual WM capacity limits [Chaieb et al., 2015, Leszczyński et al., 2015]. Studies have also shown that rats [Aggleton et al., 1992] and human participants [Nichols et al., 2006, Olson et al., 2006] with lesioned hippocampus have deficits during maintenance of multiple items or associations of multiple features. For example, Olson and colleagues compared the performance of patients with the medial temporal lobe amnesia and a control group of healthy participants in a task that required maintaining either three objects, three locations or object-location conjunctions in WM. The authors observed that patients and controls performed equally well in memorizing objects and locations separately but were impaired while maintaining conjunctions of both features. The authors suggest that the hippocampus is critical for relating or binding separate features into conjunctions at both short and long time intervals. Similarly, Nichols et al. [2006] observed that performance of amnesiac patients with variable lesions in the medial temporal lobe was impaired in WM task that included maintenance of a face. Although face was a singular stimulus but remembering it requires an integration of multiple features (i.e. its shape, relations between eyes and eyebrows, etc.). These studies suggest that the hippocampus is not only activated during WM tasks, but it is also necessary under certain conditions which include the maintenance of multiple items or a number of features. Along the lines of these findings Ranganath and Blumenfeld have doubted the double dissociation between short-term and long-term memory and questioned if theories of memory need to propose a distinct storage mechanisms for both [Ranganath and Blumenfeld, 2005]. In particular, the authors reviewed studies showing that working memory indeed do rely on the medial temporal lobe when complex or novel visual objects need to be maintained.

2.4.3 A hypothetical role of the hippocampus in WM.

The emerging doubts on the double dissociation and increasing amount of research showing hippocampal contribution to WM begs the question as to what is being represented in the hippocampus during WM maintenance? What is the role which hippocampus plays during maintenance of multiple items? One may imagine a twofold role for the hippocampus during working memory maintenance. First, a *domain general* involvement of the hippocampus would assume that it does not represent any particular feature maintained in working memory but is rather driven to encode information to episodic memory or plays some form of an executive function [for a similar idea see Shohamy and Turk-Browne, 2013]. Alternatively, a *feature specific* account would suggest that the hippocampus codes a particular feature during a working memory task. This feature would likely be an abstract and complex association across items (for example temporal or spatial relations). Importantly, the feature specific account would assume that the hippocampus is involved only in a subset of working memory tasks. These would be the type of task which requires such a relational representation. This in turn explains why the performance in some WM tasks is spared after the hippocampal lesion.

In sum, the double dissociation between the brain structures (hippocampus, PFC) and cognitive functions (episodic memory, working memory) with the the hippocampus being essential for the episodic memory but not for the working memory becomes questionable. There is accumulating evidence pointing towards hippocampal involvement beyond episodic memory in a broad range of cognitive functions spanning working memory [Axmacher et al., 2010b, Chaieb et al., 2015, Leszczyński et al., 2015] but also perception, decision making [for review see Shohamy and Turk-Browne, 2013] or conflict processing [Oehrn et al., 2015].

2.5 The sensory cortices.

As outlined in chapter 1.3, the state-based cognitive model of working memory predicts the involvement of sensory cortical areas in representing individual items in WM. The sensory cortices are hypothesized to contribute high-fidelity information about each individual item stored in the WM. Indeed, several studies have reported that the visual [Ester et al., 2009, 2015, Harrison and Tong, 2009, Riggall and Postle, 2012, Sarma et al., 2016], auditory [Bigelow et al., 2014, Scott and Mishkin, 2015, Scott et al., 2014] or somatosensory areas [Harris et al., 2002, Katus and Eimer, 2015, Katus et al., 2014, 2015] participate in task- or content-specific maintenance in WM. Furthermore, various parts of the visual cortex have been shown to contribute to maintenance of the content-specific information. For example, studies have shown the involvement of the posterior parietal cortex (PPC) to the maintenance of spatial information [Christophel et al., 2015, 2012]. The dorsal visual stream have been found to selectively engage in the maintenance of orientation [Jokisch and Jensen, 2007]. The content-specific memory signals have been also observed in the motion sensitive visual (MT+) area [Christophel and Haynes, 2014]. One very interesting and important example of a study showing contribution of the visual cortices to WM is the one by Christophel and colleagues. They presented their participants with two complex color patterns during each of the 192 trials. Subsequently, the cue indicated which of the two images a participant should remember over a first delay period of 8s. After this time interval a cue informed a participant to rotate the memorized image either clockwise or counter-clockwise. A second delay followed during which participants were maintaining the rotated mental image. Finally, participants performed a similarity discrimination task during which they indicated which of the two newly presented patterns was similar to the image after rotation. This procedure allowed the authors to probe if an image was accurately memorized and rotated. The study allowed investigating the representation of remembered stimuli before and after mental rotation. Performing multivariate decoding analysis of the BOLD fMRI signal the authors observed that the occipital and the parietal cortex contained the stimulus specific

information regarding both memorized stimuli before but also after the mental rotation. This is very interesting because it shows that not only memorized representation is coded in the visual cortex but also representation which is changed due to mental operations in working memory (i.e. mental rotation).

2.6 The gating by inhibition model in the context of working memory.

The gating by inhibition model is a generic framework describing a role of the alpha oscillation (8-14Hz) in shaping functional neural architecture [Jensen and Mazaheri, 2010]. Activity in the alpha range is a very prominent rhythm observed across many regions in the human brain. It has been suggested as a marker of functional and neural inhibition [Jensen et al., 2014, Klimesch, 2012, Klimesch et al., 2007]. Prominent alpha is usually interpreted as reflecting attenuated neural activity in a network (some refer to this as disengagement of a cortical region) which is being inhibited by some other brain area (e.g. the top-down influence of the lPFC). Indeed, spikes [Haegens et al., 2011] and high frequency gamma power [Spaak et al., 2012] are modulated as a function of both phase and amplitude of alpha activity (see Figure 2.4A). Therefore, this model attributes inhibition to both the amplitude of the alpha oscillation and to its phase [Jensen et al., 2014]. The gating by inhibition account further suggests that instantaneous information propagation to networks that show strong alpha power is reduced [Jensen and Mazaheri, 2010]. Consequently, the power in the alpha band might play a role in network connectivity blocking task irrelevant pathways (see Figure 2.4B,C).

In the context of working memory this account assumes that networks actively involved in the maintenance of information will exhibit lower levels of alpha oscillations [Jokisch and Jensen, 2007]. It also assumes that a reduced amount of information will be exchanged with networks showing high alpha power (see Figure 2.4C, right panel). To test some of these predictions Jokisch and Jensen recorded MEG (Magnetoencephalography) signal during a single item working

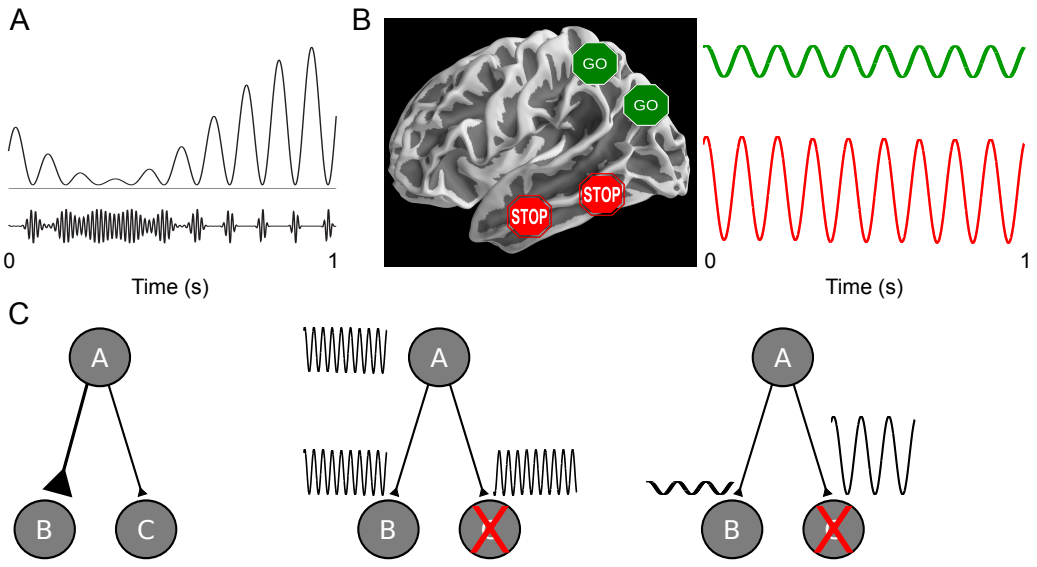

FIGURE 2.4: The alpha inhibition and the gating by inhibition hypotheses. (A) The alpha inhibition model posits that neural activity (i.e. spikes and gamma oscillations) are modulated by the alpha power and phase. (B) The gating by inhibition suggests that information is routed to these networks which show decreased alpha power (i.e. disinhibition). (C) Three possible principles of information gating through a network. Synaptic connection may be strengthened between A and B but not between A and C. Information may be gated through increased neuronal phase synchronization between A and B but not between A and C. Finally, gating by inhibition suggests that information may be propagated when some routes are suppressed by functional inhibition (alpha power increase). Increased alpha power inhibition in the node C gates information between A and B. Adapted from [Jensen and Mazaheri, 2010].

memory task. Participants viewed rotated images of faces. Depending on the condition they were instructed to maintain the identity of a face or its orientation. Passive viewing was used as a control condition. The task was designed to selectively engage either the dorsal (orientation) or ventral (identity) visual stream. The authors observed pronounced alpha oscillations over parieto-occipital sensors during maintenance of identity as well as during the control condition. This activity was strongly reduced when participants maintained orientation of faces. The gamma activity mirrored the results of alpha oscillation and was strongest during the maintenance of orientation as compared to the identity and control. This was one of the first evidence supporting the alpha inhibition hypothesis by showing that a decreased alpha activity is indeed associated with disinhibited brain networks (i.e. irrelevant for the task). Another concept important for the alpha

inhibition model is the notion of the *duty cycle* which refers to the phase of alpha oscillations where the inhibition is weakest (see Figure 2.4A). The model suggests that this time interval of minimal inhibition increases as the alpha power decreases [Jensen et al., 2014, Spaak et al., 2012]. Therefore, it posits that the alpha power increases should be accompanied by shorter duty cycles, and alpha power decreases by longer duty cycles, respectively [Jensen et al., 2014, Spaak et al., 2012]. Since the duty cycle has been related to a very specific phase (the trough) of alpha oscillations, the model further predicts a separation of duty cycles by periods of inhibition (the peak of alpha rhythm). Consequently, the amplitude of high-frequency activity is modulated by the phase of low-frequency activity, a phenomenon described as phase-amplitude cross-frequency coupling [CFC; Axmacher et al., 2010b, Leszczyński et al., 2015, Siegel et al., 2009].

2.7 Summary

The neural models of working memory have recently gone through a change with respect to assumptions they make on the type of process that working memory is. From assuming WM is a stable and stationary process they moved to interpreting WM as a dynamic, highly dimensional and distributed process. This results in a major shift from investigating sustained neural response to investigating dynamic updating of representations. Efficient WM performance has now been assumed to relay on dynamic interactions across many brain areas rather than on a processing in a single brain area. Consequently, the lPFC lost its exclusive role in working memory. Finally, this also leveraged the working memory research targeting other than the PFC brain areas (i.e. the hippocampus or sensory cortices).

Chapter 3

What are the neural oscillations?

Rhythm is a ubiquitous phenomenon in our lives. For example, ticking clocks, vibrating buildings, sounds and light may be considered oscillatory phenomena. Our daily routines are organized into a rhythmic interplay between sleep and awake. Markets fluctuate between the states of boom and depression. Therefore, it should not be a surprise that the oscillations are also observed in the brain across different species. The mammalian brain shows oscillations ranging from 0.05Hz to 500Hz. They consist of periodic fluctuations of the extracellular field potentials [Buzsáki and Draguhn, 2004].

3.1 The origins of neural field potentials.

There are various neuronal activities which elicit currents that might be observed extracellularly. These electric currents from many cellular processes superimpose at a specific location generating a potential (measured in Volts), relative to some reference. Although several contributing sources have been identified (i.e. fast action potentials, calcium spikes, intrinsic currents and resonance, spike afterhyperpolarizations, gap junctions, neuron-glia interactions, ephaptic effects), synaptic activity is known to be the main source of the extracellularly recorded potentials

[Buzsáki et al., 2012]. These potentials are refereed to as the electroencephalogram (EEG) while the recording electrode is placed on the scalp and the electrocorticogram (EcoG) when recorded by electrodes placed directly on the cortical surface or by the depth electrodes (this is also refereed to as the intracranial EEG, iEEG). The LFP reflects the same potential recorded with electrodes of the small size inside the brain and the MEG is the magnetic field elicited by the same activity.

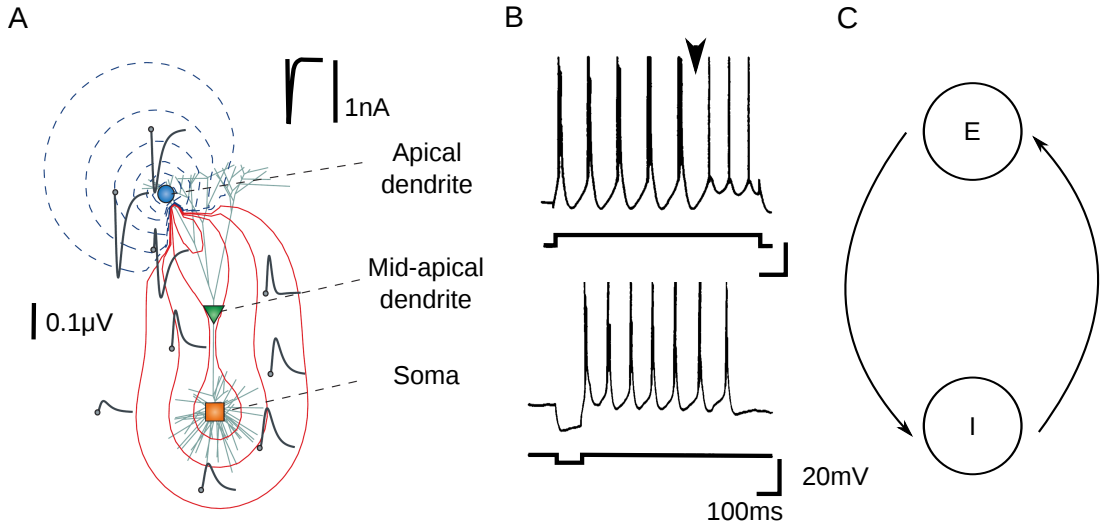

FIGURE 3.1: Generation of the local field potential and the neural oscillations. (A) Spatial distribution of the local field potential in response to an excitatory synaptic current input (a sink, presented here in blue). Red and blue contour lines reflect positive and negative values of the LFP, respectively. Apical dendrite, mid-apical dendrite and some, blue, green, orange, respectively. Adapted from [Buzsáki et al., 2012] (B) Rhythmic firing pattern of neurons from layer 5. Bursting response to step current injection (upper panel) followed by oscillatory single-spiking pattern (indicated by an arrow). Lower panel presents another neuron from layer 5 with rhytmic bursting response to transient current pulse (lower trace). Adapted from [Silva et al., 1991] (C) Recurrent inhibition model with networks of excitatory E and inhibitory I neurons which are mutually connected (arrows).

An extracellular electrode records potential difference which is generated mostly by the cortical pyramidal neurons (see Figure 3.1A). At rest these cells show unequal distribution of ions across the membrane which results in a steady and homogeneous negative potential difference at about -70mV measured intracellularly. During excitatory postsynaptic potential (EPSP) positive ions (i.e. Na^+)

flow from the extracellular to the intracellular space resulting in a local extracellular sink (positive charge enters the cell; Figure 3.1A blue part of the dendrite). The reverse happens during the inhibitory postsynaptic potential (IPSP). In order to balance this extracellular sink an opposing flow of ions from the intracellular to extracellular space takes place along the neuron (this is known as the passive current also referred to as the return current flow). The dipole may result from these two opposing ionic movements. Although all neurons contribute to the potential, their relative contribution depends on the shape of the cell. Most pronounced contribution comes from the pyramidal cells which have long apical axons and generate strong dipoles due to spatial separation of the sink/source and the return currents. Single dipoles however are too small to be measured by most of the extracellular techniques. Even the dipoles generated by the pyramidal cells are too small. The recorded potential is in fact a summation across several of such parallel dipoles. Therefore, there are two important factors determining the magnitude of the extracellular potential – spatial alignment of neurons with long apical dendrites and temporal synchrony. Spatial alignment is important because the dipoles with radial arrangement will cancel each other. The temporal synchrony is important because the dipoles in order to scale from summation must temporally overlap [Buzsáki et al., 2012].

3.2 The origins of neural oscillations.

The exact mechanism generating oscillatory potentials remains unresolved. The simplest possible explanation would be to assume that there are intrinsically oscillating cells which entrain the whole network. Indeed, many neurons have been found to exhibit such a property. In particular, the cortical layer 5 pyramidal cells show intrinsic oscillations of low frequency (5-12Hz). Silva et al. [1991] recorded in vitro activity of neurons in slices of rats' sensorimotor cortex and observed that many neurons in the layer 5 respond rhythmically to various intracellular current pulses. Interestingly, the rhythmic response could be induced both by prolonged and by a brief intracellular current pulse [see Figure 3.1B; Ritz and Sejnowski,

1997, Silva et al., 1991]. Gray and McCormick [1996] described the so called chattering cells which were found in the cortical layer 2/3 of the anesthetized cats' striate cortex in vivo. These cells responded to the intracellular current stimulation with repetitive burst discharges. The frequency of this intrinsic oscillation in the chattering cells was found to cover a broad range of the gamma spectrum from 20Hz to 80Hz. Additionally, these cells had a tendency to fire repetitive bursts of action potentials in the gamma frequency range in response to the preferred stimulation.

Another possibility is that oscillations are the property of a network resulting from interactions among cells. Ritz and Sejnowski [1997] suggested several possible mechanisms based on network interactions. For example, a recurrent inhibition model posits that a neural oscillation may arise from mutually connected networks of excitatory and inhibitory cells (see Figure 3.1C). In such a network the oscillations are organized by reciprocal recruitment of excitatory and inhibitory cells. Destexhe and colleagues developed a computer model of a network consisting of the thalamocortical (TC) and the thalamic reticular (RE) neurons which was based on the results from ferret thalamic electrophysiology. They observed that the thalamocortical cells elicited EPSPs in the reticular neurons which in turn elicited IPSPs in the TC cells. Such an interaction between mutually connected thalamocortical and thalamic reticular neurons resulted in sustained oscillating potentials [Destexhe et al., 1996]. Another popular model suggests that intrinsically oscillating cells may entrain a larger network. These intrinsically oscillating cells may serve a pacemaker role for the network.

Koepsell and colleagues suggested further possibilities. First, oscillations may be entrained be the temporal structure of an external stimulus (e.g. beats in the auditory cortex). Indeed, oscillations in the auditory cortex have been shown to follow the temporal structure of the auditory stimulus [Lakatos et al., 2008]. Second, they may originate from periodic movements of mechanical sensors [i.e. rodent's whisker system or human eye or hand movements Koepsell et al., 2010]. This hypothesis is supported by findings which show frequency specific entrainment resulting from a rhythmic finger tapping [Morillon et al., 2015, 2014].

Although the exact mechanisms critical for generation of oscillating potentials are still debated, there is fair amount of evidence that oscillations are relevant physiological phenomena which play several important functions [Buzsáki and Draguhn, 2004, Gray and McCormick, 1996, Ritz and Sejnowski, 1997]. There is also gathering consensus that oscillating potentials reflect cyclic variations in the neural excitability [Schroeder and Lakatos, 2009].

3.3 A system of oscillations.

Neural networks exhibit oscillations in several frequencies. Penttonen and Buzsáki [2003] and later also Buzsáki and Draguhn [2004] suggested that the mean frequencies of the observed oscillations form a linear increase on a natural logarithmic scale (see Figure 3.2). Several rhythms can coexist in the same network and at the same time. Rhythms can also interact with each other. Conventionally, there are five main frequency bands - delta (1-3Hz), theta (4-7Hz), alpha (8-12Hz), beta (13-30Hz), gamma (31-150Hz). Despite these well studied conventional frequency bands there is a continuously rising interest in the slow (0.1Hz-1Hz) and infraslow (0.01Hz-0.1Hz) oscillations [Monto et al., 2008, Vanhatalo et al., 2004].
It is important to note that these frequency bands and their limits is an arbitrary convention. For example, research using the rodent model often refers to the theta oscillation as spreading over a broader spectral window from 4Hz to 10Hz [Buzsáki and Draguhn, 2004, Penttonen and Buzsáki, 2003]. In turn, research investigating oscillations in the human subjects often refer to the oscillations at 10Hz as the alpha rhythm which is very prominent in the human neocortex [Berger, 1929, Jensen et al., 2012, 2014, Klimesch, 2012]. Given that the oscillatory potentials strongly depend on the cellular architecture it is important to be cautious while interpreting rhythms of the same frequency observed across distinct cortical regions.

Furthermore, activity across distinct frequency bands is not independent. Lakatos et al. [2005] observed that spontaneous oscillations in the auditory cortex of macaque monkeys are organized hierarchically. They observed that the amplitude

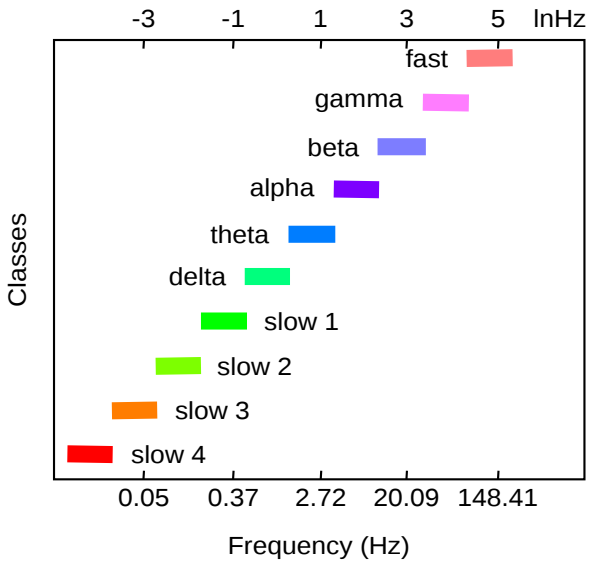

FIGURE 3.2: The linear progression of the frequency classes on the logarithmic scale. Adapted from [Buzsáki and Draguhn, 2004, Penttonen and Buzsáki, 2003].

of the theta and gamma activities were periodically fluctuating. Importantly, these fluctuations had a particular spectro-temporal structure. The activity in gamma fluctuated at theta frequency and the activity in theta fluctuated at delta frequency. The effect was observed across all cortical layers. These results show both that the magnitude of oscillations depends on power in other frequencies (power-power correlation) but also that the power of higher frequencies relates to the phase of lower frequencies (phase-to-power correlation). These phenomena are of great importance to research on working memory (see Figure 2.3). It is important because the multiplexing model of WM predicts that sequences of items are represented in WM as gamma bursts nested in the phase of low frequency (see chapters 2.3).

There is large amount of research aiming to describe functions which are supportd by these different oscillations. Some of the most important are listed by Buzsáki and Draguhn [2004] and include input selection and plasticity, binding cell assemblies, consolidation and combination of learned information, representation by phase coding. For the purpose of the current thesis I will focus on the theta, alpha and gamma oscillations which have been found relevant during various forms of

working memory task.

3.4 A role of theta oscillation in working memory.

Theta activity has been studied for long in the context of the hippocampal place cells firring patterns and its relevance to spatial navigation and memory processes [for review see Buzsáki and Moser, 2013, Kahana, 2006]. Studies in human subjects have also reported a role of theta activity during a working memory task [Axmacher et al., 2010b, Chaieb et al., 2015, Jensen and Tesche, 2002, Klimesch, 1999, Liebe et al., 2012, Raghavachari et al., 2001, 2006, Sauseng et al., 2010, 2004, 2009].

3.4.1 The power of theta.

Raghavachari et al. [2001] and Raghavachari et al. [2006] recorded intracranial EEG while their participants were performing a modified Sternberg task (similar to this depicted in Figure 5.5A) with a list of 1-4 consonants presented in a sequence. After the maintenance interval participants were presented with a probe and were required to indicate whether the probed letter was presented in the memory set. The authors observed elevated theta power lasting during the whole task. This effect was observed across several distributed anatomical locations in the occipital, temporal and frontal lobes. Interestingly, increased power of the oscillation started at the beginning of a trial, during encoding and terminated at the end of the retention interval. Thus, the authors concluded that activity in the theta range reflects cognitive gating of brain oscillation during WM task. The power of theta has been also observed to increase as a function of memory load. Jensen and Tesche [2002] used MEG to observe a parametric increase in the theta power across the number of items maintained in WM. This effect was measured during the whole retention interval over the electrode sensors localized in the frontal

cortex. Importantly, the power was again reduced following the task suggesting that the gating effect is load specific. These results show that the theta oscillations are important during a task itself and that the frontal theta plays an active role in WM maintenance.

3.4.2 The phase of theta.

The phase of theta oscillation has also been shown to play a relevant role in the context of working memory maintenance. Sauseng et al. [2009] conducted an EEG study in which they used a cued spatial change detection task. First a cue was presented which informed participants regarding the side of the display which they should attend. They were required to encode stimuli only from this side of the screen and ignore the stimuli presented on the contralateral side. The cue was followed by a change detection task (see Figure 1.3) with bilateral presentation of stimuli and 900ms of maintenance interval. The authors observed an increased theta phase-locked gamma burst on the side contralateral to the cued side. This suggests that theta-phase gamma power may be relevant for maintenance of information in WM. Although the phase-locked gamma power was not modulated by memory load itself, the authors observed that the synchronization between theta and gamma phases was increasing with rising memory load. A similar role of the theta phase has also been demonstrated in the human hippocampus by my colleagues and me [Chaieb et al., 2015]. To this end, we calculated m:n phase-phase coupling between the phases of theta and gamma. This measure quantifies how precisely the high frequency oscillation is locked to the low frequency rhythm. This is different from the phase-amplitude coupling because it looks at relation between phases of both oscillations rather than between phases and power. Similarly, to Sauseng et al. [2009] we observed a prominent increase in the theta-gamma phase-phase coupling during maintenance of multiple-items as compared with an inter-trial interval. Importantly, the coupling was absent for load 1 supporting the idea that the hippocampus contributes specifically to maintenance of multiple items [Chaieb et al., 2015]. Axmacher and colleagues observed increased hippocampal

theta-gamma phase to power coupling during maintenance of images presenting unfamiliar faces. Interestingly, the frequency of theta activity which was modulating gamma burst decreased across memory loads. This was interpreted in line with the theta-gamma multiplexing model which posits that longer theta cycles are needed to represent increasing number of items in a memory sequence [Lisman and Idiart, 1995, Lisman and Jensen, 2013]. According to these findings a local theta may be a mechanism for binding items into a coherent memory set. Sauseng and coworkers observed that a long-range theta synchrony may also contribute to binding of items into a coherent memory set. In particular, they found that long-range theta phase coherence between the frontal and temporo-parietal sites was increased during encoding and retrieval of information in WM as compared to the inter-trial interval. The authors concluded that the long-range theta phase coherence is relevant for coactivation of distributed networks involved in encoding and retrieval of a working memory task [Sauseng et al., 2010, 2004]. Similarly, Liebe et al. [2012] found increased inter-regional phase coherence between the theta oscillations in the frontal cortex and the V4 of a monkey performing a working memory task. Furthermore, the phase of the frontal oscillations modulated single unit activity in the visual area V4.

In sum, the oscillations in theta range are important for coding information in working memory. Theta phase may be the mechanism for organizing multiple items into distinctive memory sets by synchronizing activity in higher frequencies [Lisman and Idiart, 1995, Lisman and Jensen, 2013]. A similar role may be attributed to the long-range theta phase coherence between distributed cortical regions. Increased coherence across such a distributed network may support coactivation of neural assemblies involved in working memory [Liebe et al., 2012, Sauseng et al., 2010, 2004].

3.5 A role of alpha oscillation in working memory.

The alpha (8 to 12Hz) rhythm is the most prominent neural rhythm in the human brain during an awake state. It was first observed and described by Hans Berger who found that the power of this rhythm increases while participants have their eyes closed [Berger, 1929]. This early finding developed to the classical thinking about the alpha oscillation as reflecting the state of relaxation and task unoccupied brain [Klimesch, 1999, 2012].

Although the precise physiological model of the alpha rhythm remains to be found, a large number of studies interpret an increase in power of alpha as a marker of functional inhibition [Bonnefond and Jensen, 2012, Jensen et al., 2012, Jokisch and Jensen, 2007, Klimesch, 2012, Klimesch et al., 1999, Sauseng et al., 2009]. Recent findings indeed relate the power but also the phase of alpha oscillation to the attenuation of neural activity [Haegens et al., 2011, Spaak et al., 2012]. Due to its inhibitory property the alpha has been suggested to support the spatial working memory indirectly by suppressing distracting environmental input [Freunberger et al., 2011]. This hypothesis was motivated by findings in the cued change detection task. Such a task design allowed dissociating the activity related to the maintenance of relevant information from the activity associated with ignoring the redundant distractors. Importantly, the performance has been observed to benefit from efficient filtering of the irrelevant information [Bonnefond and Jensen, 2012, Sauseng et al., 2009, Vogel et al., 2005]. The power of alpha during maintenance interval increases on the side ipsilateral to the cued hemifield. This suggests that the alpha power may be involved in suppressing task irrelevant information. This was corroborated by findings showing that, the magnitude of amplitude followed the number of distractors presened on the contralateral side. The lateralized alpha power was also predictive of individual spatial working memory capacity [Sauseng et al., 2009]. The importance of lateralized alpha oscillation for performance was further investigated by repetitive transcranial magnetic stimulation over the side

contralateral to the hemifield where distractors were presented. The 10Hz alpha power entrainment over the occipital sites where the irrelevant information is processed improved WM performance, most likely due to increased efficiency in the distractor suppression. In contrast, the TMS contra-lateral to the target side harmed performance by suppressing relevant target maintenance. This effect was specific to the alpha frequency range since the 15Hz stimulation had no behavioral effect [Freunberger et al., 2011, Sauseng et al., 2009]. Another study showed that this effect is not limited to spatially presented information but generalizes to the temporal expectation of distractors. Bonnefond and Jensen [2012] used a modified Sternberg's task during which they presented sequentially a series of four to be remembered stimuli followed by a distractor at the end of the sequence. The authors observed that both the power and the phase of alpha oscillation adjusted in anticipation of the upcoming distractor. Importantly, the increase of the predictive alpha magnitude was beneficial for behavioral performance as measured with the reaction times. Participants with higher anticipatory alpha showed faster reaction times [Bonnefond and Jensen, 2012]. This is in line with the hypothesis that the alpha oscillation serves a mechanism for filtering out irrelevant and distracting information.

A similar role for the alpha power has been also atributed to the dorsal visual stream where the alpha was found to reflect filtering of irrelevant stimuli. In a study by Jokisch and Jensen [2007] each trial started by presenting a participant with a single, rotated image of a face. Depending on the condition a participant was instructed to remember either the orientation or identity of a face. Importantly, each of the images contained both information (i.e. identity and orientation). Therefore, in each trial some information needed to be ignored whereas another information needed to be maintained as relevant for the task. Similarly, to the results mentioned above regarding the spatial and sequential working memory the relevant information was also observed to elicit decrease in the alpha oscillations. In particular, a lower alpha power was observed in the dorsal visual stream when a participant needed maintain information about the orientation of an image [Jokisch and Jensen, 2007].

Together, prominent alpha power is usually interpreted as functional and neural inhibition. The increased alpha power have been viewed as the process of blocking irrelevant information by attenuating the respective network. Conversely , a decrease of alpha power marks a release from neural inhibition indicating increased involvement of the cortical region.

3.6 A role of gamma activity in working memory.

There is a long tradition of research associating gamma activity with the visual perception [Bertrand and Tallon-Baudry, 2000, Gray and McCormick, 1996, Jensen et al., 2007, Tallon-Baudry and Bertrand, 1999]. Although the cellular mechanisms and the function of gamma activity remain unclear, it is however, believed that the broadband gamma activity reflects synchronous, rhythmic firing of large neural assembles [Jensen et al., 2007, Mukamel et al., 2005]. Furthermore, the synchrony in the gamma range is well suited to provide an optimal time frame for neuronal communication. Imagine there is a neuron which receives an input from several other cells. The response of this neuron is enhanced if inputs from the other cells are synchronized and the resulting excitatory postsynaptic potentials are integrated. The duration of the EPSP is approximately 10ms, which makes gamma (with the duration of a single cycle at about 10-30ms) the ideal candidate for a mechanism of such a termporal integration [Jensen et al., 2007].
Tallon-Baudry et al. [1998] used a delayed-matching-to-sample task where they presented subjects with abstract shapes as stimuli. Participants were required to press a button if the probed stimuli matched the memory sample. Importantly, the authors compared maintenance during which the memory representation had to be actively sustained with a dimming condition. The dimming condition consisted of the same sample stimuli as presented at the beginning. However, this task did not have any memory load. Participants, were required to only actively view the stimulus which was not probed at the later stage. In order to match

the response and alertness conditions participants were required to press a button whenever the fixation cross changed its luminance. This comparison is very important because it keeps the visual features of used images as well as attention and response requirements constant and only varies the need to actively sustain the representation of the perceived items. The authors recorded scalp EEG while participants were performing both tasks and observed that the gamma activity over the occipital sites was increased compared to the baseline during active maintenance condition but it did not differ from the baseline during the passive viewing with no memory load. Because the delay period used in that study was transient and constant it was difficult to interpret this result as reflecting a typical working memory sustained response pattern. Thus, Tallon-Baudry et al. [1999] performed another study during which they increased the length of maintenance interval and observed that the occipital gamma effect was in indeed sustained.

Axmacher et al. [2007] recorded intracranial EEG signal from the human hippocampus while their participants were performing a modified version of the Sternberg working memory task. To this end, images of unfamiliar faces were presented to participants followed by maintenance interval. Subsequently, a probe image was presented and participants performed a matching-to-sample task. The authors observed that the gamma-band activity in the human hippocampus was reduced compared to prestimulus baseline but only when participants maintained a single item. This pattern of effects is in line with the hypothesis suggesting that the hippocampus is particularly involved in processing of multiple items in working memory (see chapter 2.4). van Vugt et al. [2010] further investigated the role of gamma power for WM maintenance in the human hippocampus. They also recorded intracranial EEG signals from the human hippocampus while their participants performed a modified version of the Sternberg working memory task. Importantly, they used images of filtered synthetic faces and digits as memorandum. The authors observed that the gamma power in the hippocampus was increasing across memory loads during the delay period. The effect was stronger during the maintenance of faces as compared to the maintenance of digits. This effect of a relative difference between the gamma power during maintenance of faces

and digits is supports the idea that the hippocampus is sensitive to the number of information maintained in WM. In particular, images of faces are much more complex and contain much more features which have to be remebered as compared to simple digits. This effect of gamma power increase across memory loads for both types of stimuli was also replicated by me and my co-workers [Leszczyński et al., 2015].

Chapter 4

Motivation

4.1 How is the hippocampus involved in working memory maintenance (Experiment 1 and 2)?

The alpha-inhibition model suggests that activity in the alpha range reflects the degree to which a cortical region is involved in processing of a given feature (see chapter 2.6 for more details). In particular, an increased alpha power reflects stronger inhibition as measured by the number of spikes [Haegens et al., 2011] and the amount of high frequency activity [Spaak et al., 2012]. Abundant evidence have shown that networks actively involved in the task exhibit lower levels of alpha oscillations [Bonnefond and Jensen, 2012, Jensen et al., 2012, Jokisch and Jensen, 2007, Klimesch, 2012, Sauseng et al., 2009, Thut et al., 2006]. This effect is often reffered to as the functional inhibition property of the alpha. However, it remains unknown whether this feature generalizes to the hippocampus. The functional inhibition model would predict that the increasing number of items maintained in WM would lead to a linear power decrease in the alpha frequency range. I used data recordings from the human hippocampus to test this prediction. The results of these tests are described in chapters 6.1 and 6.2.

The multiplexing buffer model (see chapter 2.3 for more details) assumes that representations of individual items (via high-frequency oscillations) do not occur during the entire cycle of low-frequency oscillations. Instead, these high-frequency oscillations are thought to be restricted to the phase ranges that correspond to higher levels of neural excitability [the so called duty cycle Jensen et al., 2012, 2014, Mehta et al., 2002]. Between successive duty cycles, excitability is actively reduced in order to avoid overlap and interference between the representations of item sequences. The model predicts that gamma oscillations are not uniformly distributed across the theta cycle but are restricted to phase of the highest excitation. This in turn results in the cross-frequency coupling. The hippocampus has been observed to exhibit such patterns during WM of novel materials [Axmacher et al., 2010b]. Here, I tested whether the hippocampus also shows an increased CFC during maintenance of highly familiar information (i.e. digits). This question is important because representations of familiar information have been hypothesized to rely on networks other than the hippocampus [i.e. the neocortex Hasselmo and Stern, 2006].

I further explored the dynamics of the duty cycle and its relations to the strength of the cross-frequency coupling. The multiplexing model predicts that maintenance of longer sequences requires a reduction of alpha power, so that more items can be represented during longer duty cycles. Therefore, a high WM load should correspond to a decrease alpha power (increased involvement) and decrease of CFC strength, because the modulated frequency would be more broadly distributed across the entire cycle of modulating low frequency activity (mLFA).

4.2 The neural mechanisms of individual limits in working memory (Experiment 1 and 2).

One important prediction from the multiplexing buffer model of WM [Lisman and Idiart, 1995, Lisman and Jensen, 2013] is that the parameters of cross-frequency coupling (frequency for phase, frequency for amplitude or the ratio between these

two frequencies) would correlate inter-individually with WM capacity limits. This prediction was tested during iEEG signal obtained during WM maintenance of both highly familiar (chapter 6.1) and novel (chapter 6.2) stimuli. The model predicts that WM capacity should be maximal if the duration of individual low-frequency cycles were as long as possible (lower frequencies of mLFA), if high-frequency cycles were as short as possible (higher frequencies of modulated high frequency activity; mHFA) or if CFC ratios were high. Furthermore, the m:n phase-phase coupling was also explored as a potential mechanism limiting individual WM capacity.

4.3 A role of alpha oscillations in the ventral and the dorsal visual stream (Experiment 3).

Successful working memory (WM) requires both maintenance of relevant and inhibition of irrelevant information. The alpha rhythm (9-13Hz) has been found to inhibit redundant visuo-spatial information during WM maintenance. Importantly, this effect has been limited to the parieto-occipital sites [Bonnefond and Jensen, 2012, Jensen et al., 2012, Jokisch and Jensen, 2007, Klimesch, 2012, Sauseng et al., 2009, Thut et al., 2006]. In these cortical networks it has been suggested that the alpha band activity reflects the allocation of resources by inhibiting and disinhibiting specific cortical regions. While this has been explored in the parieto-occipital lobes, it remains elusive if the alpha oscillations plays a similar role in the higher sensory cortices (i.e. the ventral visual stream). Here, I tested whether such an inhibitory role of alpha activity generalizes to the ventral visual stream. I also used the iEEG to replicate previous results obtained with another method (the MEG) which shows the functional inhibition role of alpha in the dorsal visual stream.

4.4 The gating by inhibition hypothesis (Experiment 3).

The prefrontal cortex has been suggested as a region which supports WM by performing a top-down control over the sensory cortices. Previous research found that increased phase synchronization in low frequencies might be a marker of the top-down control of the PFC over the posterior brain areas [Liebe et al., 2012, Zanto et al., 2011]. Zanto et al. [2011] observed that the PFC modulates activity in the posterior cortex selectively biasing attention during memory encoding. Disrupting the PFC with transcranial magnetic stimulation (TMS) attenuated top-down driven attention signal over the posterior electrode sites and communication between the two regions. The precise neural mechanism of this top-down modulation remains unclear. The gating-by-inhibition models suggests a possible candidate mechanism. The model posits that local neural inhibition (marked by the alpha power) operates by blocking task irrelevant pathways and in turn limits information transfer to and from the inhibited cortical region. In particular, it predicts that increased alpha power in a given brain region should relate to a decrease of neural communication between this region and other areas involved in the task. In the current context this model suggests that the synchrony between the frontal cortex and the ventral visual stream would be increased during maintenance of identity as compared to orientation and control. In turn, the coherence between the frontal cortex and the dorsal visual stream should increase during maintenance of orientation relative to identity and control. This hypothesis although fundamental for our understanding of the oscillations however has never been tested.

4.5 The search for WM code in the prefrontal cortex (Experiment 4).

Initially, the prefrontal cortex has been suggested to play a role in representing individual stimuli during WM delay period [Goldman-Rakic, 1995]. This hypothesis often implicitly assumes that memory related activity is associated with parametric change of the signal intensity (e.g. change of the oscillatory power). This presupposition, however has been recently criticized for its lack of sensitivity and specificity [Postle, 2015]. An alternative approach benefits from the multidimensional dynamics of the signal (see the Methods section). Although the multivariate approach in studying WM delay have been very productive while applied to fMRI [for a review see Postle, 2015] its attempts in electrophysiology are rather sparse. Here, I used the data from the local field potential recordings acquired from the ventrolateral prefrontal cortex (vlPFC) of two macaque monkeys performing multi-item working memory task to look for load dependent changes during the maintenance. I initially investigated the univariate signal intensity change across memory loads. To further gain sensitivity I used a multivariate pattern classification approach to look at multidimensional, load specific changes during the maintenance. Both of these analyses aimed at testing a hypothesis that the vlPFC is involved during maintenance of multiple items in WM.

Furthermore, a recent model proposed a role of the lPFC in representing abstract rules rather than individual stimuli [see chapter 2.2; Stokes, 2015]. One example of an abstract rule in a typical WM task is the response mapping. For example, in the current experiment monkey had to press left button to indicate a match (when the probe matched the identity of memorized stimulus) and the right button to signal a non-match trial. It remains unknown, whether the local field potentials from the vlPFC indeed code such information. Therefore, I used the LFP data recordings acquired from the vlPFC of the two macaque monkeys performing a single-item WM to test if these two distinct abstract response mappings are indeed coded in the signal recorded from the vlPFC.

4.6 The alpha band duty cycle in the vlPFC.

The alpha rhythm has been suggested to inhibit the neural activity [single unit and broadband gamma Jensen et al., 2014]. The range of alpha phase when the inhibition is the lowest is sometimes reffered to as the duty cycle. Spaak et al. [2012] showed that the alpha duty cycle indeed modulates neural activity in the primary visual cortex. In particular, the authors observed that the size of the duty cycle (measured with the power of alpha - high power reflects short duty cycle, low power indicates long duty cycle) modulates the length of a gamma burst. Importantly, this effect has been only shown in the primary visual cortex. Here, I investigated whether the size of the alpha duty cycle also modulates the length of gamma burst in the vlPFC. This is an important question because the effect has never been observed outside of the primary visual cortex.

Chapter 5

Methods

Various neuronal activities elicit currents that might be measured extracellularly. Electric currents from many cellular processes superimpose at a specific location generating a potential, relative to some reference [for review see Buzsáki et al., 2012]. There are several recording methods of the extracellular events like electroencephalography (EEG), magnetoencephalography (MEG), electrocorticography (ECoG) also known as intracranial electroencephalography (iEEG), local field potentials (LFP). Here, I used recordings from two methods – the iEEG and the LFPs. I will first provide a more general overview of the signal, recording procedures and analytical tools used in the current dissertation followed by a more detailed description of the experimental and analytical procedures applied to derive the current results.

5.1 Intracranial EEG and the local field potentials.

The intracranial EEG offers a rare opportunity to investigate signals recorded directly from the brain of the awake, behaving human. The signal is obtained with electrodes placed directly on the cortex (strip or grid subdural electrodes) as well in deeper brain structures like the hippocampus, or amygdala (depth electrodes,

see Figure 5.1). This method allows bypassing the skull which heavily distorts the signal and is a source of major limitations in other methods like EEG. The iEEG signal is commonly recorded from relatively small size of the recording site (i.e. 1.3mm x 2.5mm for the depth electrodes) sampled at the rate of 1kH. The location of each electrode might be confirmed with post-implantation MRI and precisely localized in the standard MNI brain space (standard brain from the Montreal Neurological Institute which is an average across many MRI scans). To this end, the post-implantation MRI scans which contain information about each electrode position are normalized to the standard MNI space. The pre-implantation images are used to correct for possible distortions caused by artifacts in the post-implantation MRI scans. This process is important because it allows determining the location of each electrode in a standard space. There are several main advantages of iEEG. First, the signal recorded using an iEEG electrode consists of postsynaptic potentials recorded from a relatively small patch of cortex [$< 5mm^2$: Buzsáki et al., 2012]. This makes the method spatially specific. In other words the superimposed currents contributing to the recorded potentials are spatially limited and therefore the signal to noise ratio is higher. Second, unlike blood-based measures (e.g. fMRI) the iEEG signal has superb temporal and spatial resolution and reflects a more direct measure of neural processing. Together the iEEG combines excellent temporal and spatial resolution making it one of the best methods currently available to cognitive neuroscience research. It is particularly valuable to research that looks at deep brain structures like the hippocampus or the thalamus. Even with current analytical tools for source reconstruction measuring the electrical signals from the deep brain structures is very difficult with other popular methods like scalp EEG or MEG. However, there are also some limitations to the method.

One important limitation of the iEEG is its invasiveness. Since electrodes are inserted directly into the cortical structures stringent ethical concerns allow its implantation only for highly constrained medical purposes. One such example is the case of patients suffering from severe pharmaco-resistant epilepsy. Such patients undergo an invasive neurosurgical treatment to localize the focus of epilepsy. Intracranial electrodes allow determining (1) the location of epilepsy focus and (2)

whether the brain area around the focus might be dissected without any obvious damage to cognitive functions. Each patient is implanted with electrodes according to specific diagnostic purposes. Various types of electrodes are implanted in one or multiple brain areas (see Figure 5.1) and patients undergo the clinical monitoring during which they stay medication free. Once a clinical diagnosis has been made and prior to explanation of electrodes, patients may, if they are willing to, participate in further scientific experiments. Another limitation of the iEEG is that participants recruited for studies often suffer from sever neurological disease (i.e. epilepsy). The electrical activity of "pathological" brain might be altered due to the disease progression. Scientific studies involving experiments with patients and recording of the iEEG signal need to carefully consider these limitations. This is usually achieved twofold. First, only electrodes from the healthy parts of the brain are used for scientific research. All electrode sites localized within or close to the epilepsy focus are discarded from analysis (i.e. in the case of bilateral implantation in the hippocampus only electrodes localized on the side contra-lateral to the epilepsy focus are considered for analysis). Second, even after excluding electrodes overlapping with the epilepsy focus the remaining electrodes might still show epilepsy related signal properties (i.e. any epileptiform activity). The signal is therefore carefully inspected and epilepsy related artifacts are removed. One other limitation is that some patients, given their long history of neurological disease might be cognitively impaired [Axmacher et al., 2007]. Any experimental design has to account for such a possibility. Experiments should be relatively easy, short with possibilities to break any time. Patients' motivation as well as attention focused on the task should be encouraged.

The local field potential (LFP), similarly to the iEEG, is also an extracellular measure of superimposed electric currents and is usually considered as a signal below 500Hz [Einevoll et al., 2013]. The LFP signal compared to the iEEG is regarded as more localized with smaller networks contributing to the observed potential [i.e. $160\mu m^2$ as reported by Peyrache et al., 2015]. Although precise estimation of the network size which is sampled is uncertain [Einevoll et al., 2013], one advantage is that the signal to noise ratio is even higher than with the iEEG.

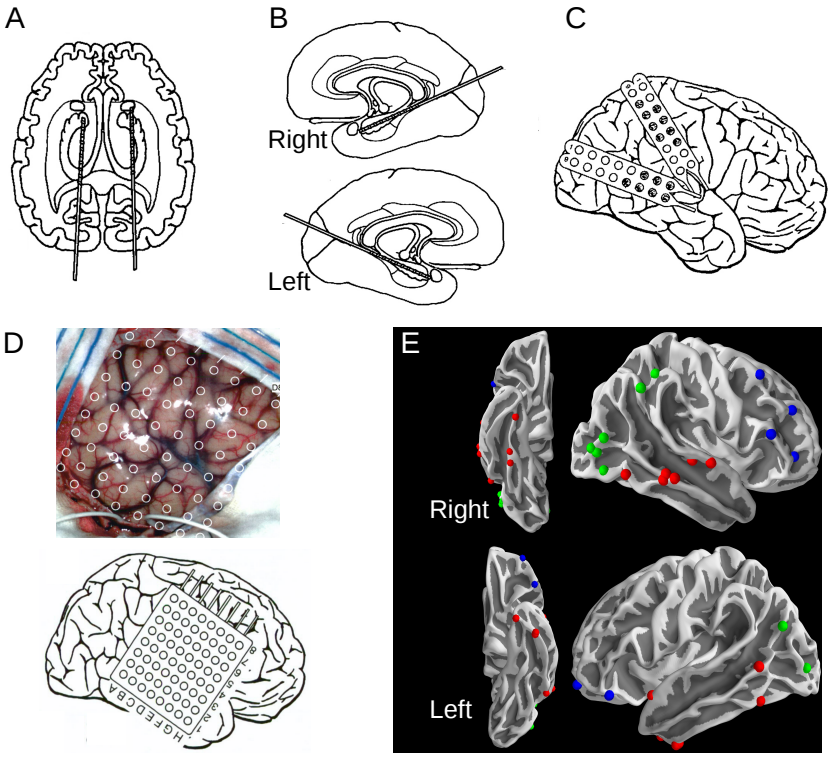

FIGURE 5.1: Examples of the intracranial electrode layouts. (A, B) An example of the depth electrodes implanted bilaterally to the hippocampus. The electrodes are presented on the axial plane (A) and the sagittal plane (B). (C) Lateral view showing an example of four strip electrodes. (D) An example of a grid electrode. An image taken during an implantation (upper panel) and schematic overview of the same grid electrode (lower panel). (E) An example of a group of patients with electrodes localized in three cortical areas of interest – the ventral visual stream (red spheres), the dorsal visual stream (green spheres), and the frontal cortex (blue spheres). Each sphere depicts a location of an electrode in a single patient.

This particularly applies to activity in higher frequencies (i.e. gamma frequency >30Hz).

5.2 Preprocessing of the data.

Preprocessing involves a set of steps which transform recorded raw signal into artifacts free, units of analysis. Preprocessing of the iEEG and LFP usually involves the following five steps: (1) re-referencing if necessary or important for the

hypothesis; (2) artifact detection and removal; (3) segmentation of the data; (4) baseline correction; (5) filtering.

5.2.1 Referencing the data.

Recorded voltage is a measure of the current flow between the electrode of interest and a specified reference site. Selection of the reference will affect observed signal and therefore it is of great importance. There are several standard procedures for referencing the (i)EEG signal. The most commare the linked mastoids, bipolar scheme or a common average. It is important that the selection is motivated by the research question because each approach comes with advantages and disadvantages. In the bipolar scheme each electrode is referenced to its neighbor increasing contribution of the local potential generators and canceling more global (i.e. visible in both channels) potentials. The common average in turn does not bear the risk of canceling global potentials. However, for the iEEG each patient has a distinct electrode layout with electrodes localized in different brain areas. This makes the reference signal highly diverse across patients and violates the assumption of common average reference (the assumption of equal electrode distribution). Additionally, a single noisy electrode (i.e. containing many epilepti form artifacts) in the case of the common average will propagate this noise to all other channels. The linked mastoids in turn is the average of electrodes placed on both mastoid bones. It does not bear the risk of canceling global potentials and it is common across all patients. Additionally, since electrodes placed on both bones are used the reference does not introduce any hemispheric bias. However, it does not come without limitations – its external location might introduce additional noise (i.e. electric noise).

5.2.2 Artifact removal.

Another important step is the removal of artifacts. The most pronounced types of artifacts in the iEEG recordings are those related to the epilepsy. The epileptiform

activity consists of a broadband spike followed by a slow frequency component which makes artifacts relatively easy to detect. The data is visually inspected and all time points contaminated by artifacts are marked and removed from further analysis.

5.2.3 Segmentation of the data.

The data epochs are usually created around the event of interest. Depending on the research question often the pre- and post-event samples are included. The duration of the segment also differs depending on the research question. For example, while analyzing data in the frequency domain one must include signals of a duration which is sufficient to cover the slowest frequency of interest. This usually requires at least three cycles of the considered frequency which is the minimum data length to reliably estimate that frequency.

5.2.4 Baseline correction.

Since most of the research questions in cognitive neuroscience look how the neural activity patterns change due to some external event (i.e. stimulus presentation), researchers are interested in a relative change of signal properties elicited by this event. The baseline correction gives such a relative measure by normalizing signal of interest to a part of the signal independent from the event (e.g. prior to the event or during rest). In the current thesis, unless it is specified differently, I used baseline period prior to the stimulus (the last stimulus in case of multiple items).

5.2.5 Filtering of the data.

Narrow band notch filter is used to discard frequencies considered as electric noise (i.e. 50Hz and its harmonics). Filtering is also used to extract the frequency specific activity in the case when activity in a specific frequency range is being investigated.

5.3 Analytical tools.

5.3.1 Time-frequency decomposition.

The traditional approach in analyzing EEG data has been to average preprocessed time domain data across large sample of trials without using any further analytical tools. This assumes that noise which is randomly distributed across trials will average out and the event related potential (ERP) will become visible. Since an ERP is composed mostly of a slow frequency power increase which is phase locked to the event of interest this approach ignores a great portion of higher frequency signal as well as these signals which are not phase locked [Fell et al., 2004]. This is particularly important for working memory research which investigates potentials during extended periods of maintenance rather than the stimulus evoked responses. The time-frequency analysis considers spectral and temporal complexity of the signal. It refers to the number of techniques [for overview see Cohen, 2014] most of which rely on convolution in the time domain. For example, the method of Morlet wavelets allows for point-by-point convolution of signal with a family of wavelets (sine waves convoluted with a Gaussian function) reflecting different frequencies. It allows extracting multiple frequencies at each time point. The slowest frequency is constrained by the length of data epoch and the highest by the Nyquist frequency (one-half of the sampling frequency). Importantly, the complex Morlet wavelets allow estimating instantaneous power and phase of the convoluted time-domain signal.

5.3.2 The cross-frequency coupling (CFC).

The cross-frequency coupling is a generic term reflecting several relations between two or more signals [Jensen and Colgin, 2007]. One version of the CFC particularly relevant for the current thesis is the phase-amplitude coupling (PAC). It reflects a relationship between the phase of slower and the amplitude of faster oscillations in which the high frequency amplitude is non-uniformly distributed across the

phases of the slower oscillation (see Figure 5.2 left). Importantly, pronounced PAC during maintenance of multiple items is predicted by the multiplexing model of working memory (see chapter 2.3 and Figure 2.3) as a mechanism for maintenance of information [Lisman and Idiart, 1995, Lisman and Jensen, 2013].

There are various ways of calculating PAC [Axmacher et al., 2010b, Canolty et al., 2006, Dvorak and Fenton, 2014, Penny et al., 2008, Tort et al., 2010]. The method used in the current thesis relies on Pearson's correlations between the filtered low- and the power of high-frequency signals. To account for possible inter-individual differences between the phases of the low-frequency signal at which the amplitude of the high-frequency signal was maximal, the low-frequency signal was shifted by the subject-specific "modulation phase" at which the strength of modulation was maximal for each participant [Axmacher et al., 2010b]. The modulation phase is calculated by extracting low- and high-frequency signals with one of the time-frequency convolution methods (i.e. Morlet wavelets or filter then Hilbert transform). This results in complex-valued time-series for low-frequency ($activity_{low}$) and high-frequency ($activity_{high}$). The phase ($phase_{low}$) of low- and modulus ($amplitude_{high}$) of high-frequency signals, respectively are used to construct a composite signal $z = amplitude_{high} * e^{i*phase_{low}}$ [Canolty et al., 2006]. Subsequently, the low-frequency signal is shifted by the individual modulation phase. Finally, Pearson's correlation coefficient between the real part of the phase-shifted low-frequency signal and the modulus of the high-frequency signal is calculated (see Figure 5.2 for illustration of these steps). The resulting correlation coefficient scores are Fisher-z-transformed for the statistical analysis.

5.3.3 The phase synchronization.

The oscillating potentials reflect cyclic variations in the neural excitability (see chapter 3.2). In other words oscillations create sequences of the excitation and inhibition restricted to short temporal windows. This periodic modulation of the postsynaptic excitability creates oscillatory modulation in the synaptic gain. The

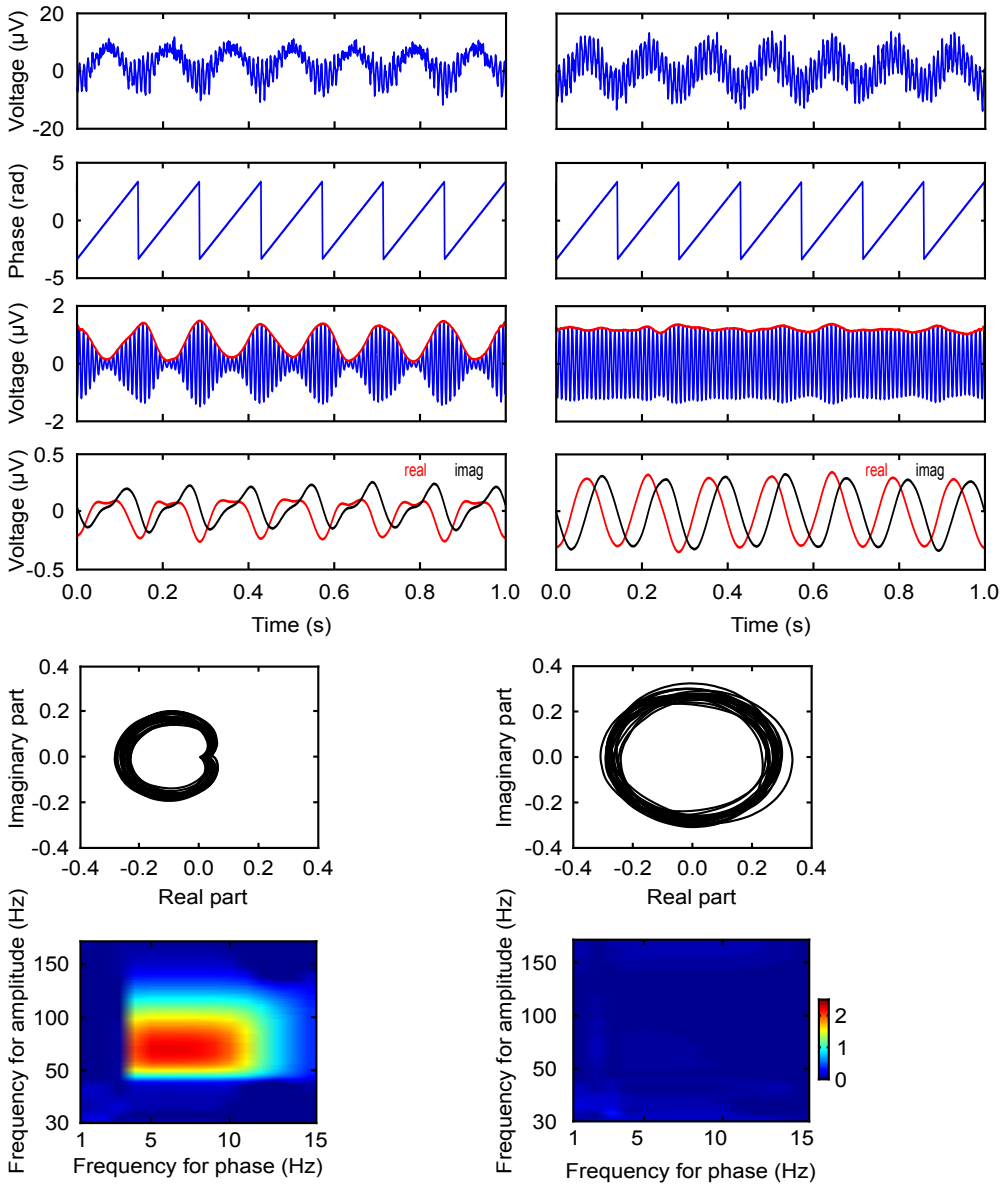

FIGURE 5.2: Steps in signal analysis for the cross-frequency coupling. Synthetic data with (left) and without (right) theta-phase gamma-power modulation. Both signals presented with the Gaussian noise superimposed. Synthetic data, time series of low frequency phases, envelope of high-frequency signal, real and imaginary parts (top to bottom, respectively). Lower two panels present modulation orbits depicting complex vectors obtained from consecutive time samples across the considered interval of 1s. The color map reflects the strength of modulation index across a broader frequency-frequency space. Adapted from Axmacher et al. [2010b].

neural inputs arriving during the high input gain benefit from the enhanced excitability or enhanced effective connectivity (see Figure 5.3). This is the essence of the Communication through Coherence model suggested by Fries [2005, 2015]. The model posits that the strong effective connectivity requires coherence between the involved groups of neurons. In the absence of coherence the neural inputs arrive at random point of the excitatory-inhibitory cycle and will have lower effective influence on the postsynaptic information transfer [Fries, 2005, 2015].
The phase synchronization usually refers to the measure of stability of phase difference between signals recorded with electrodes placed in two brain areas. The phase synchronization increases if phase difference between the two signals is consistent across trials [Fell and Axmacher, 2011].
Prominent synchronization have been observed in the theta [Liebe et al., 2012], beta [Buschman et al., 2012] and gamma [Fries, 2015] frequency range. In particular, the phase synchronization in the lower range has been proposed as a top-down mechanism of communication between the frontal and parietal brain areas [Liebe et al., 2012, Sauseng et al., 2004, 2005, Zanto et al., 2011].

One popular method to calculate the phase synchronization is the phase locking value [PLV; Lachaux et al., 1999]. The PLV may be calculated by extracting the data recorded from two brain areas which are hypothesized to engage in communication. The data is convoluted with wavelets or filtered and Hilbert transformed in the frequency of interest. This results in a complex-valued time series for each of the areas of interest. Next, phases are extracted from the complex-valued signals. The difference between these two time-series reflects locking of phases between the two electrode sites. If a cognitive task is associated with a coherent fluctuations of signals in both areas (either with no lag or with a certain delay), then the difference between the two time-series will be stable across trials. If in turn there is no coherency between phases of these two signals, then the phase difference will be randomly distributed.

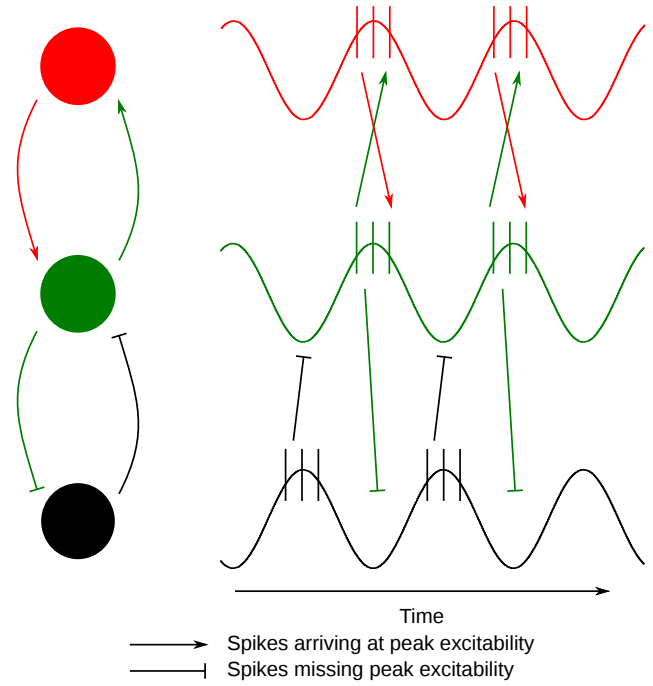

FIGURE 5.3: Communication through coherence. Spikes arriving at the peak excitability of the receiving network (arrowheads) form a condition for possible effective communication. The two coherently synchronized networks (represented by the red and green circles) form a minimal condition for communication. The networks represented by the green and black circles show fluctuations that are not coherent. Spikes arriving at the receiving network are inhibited by its own fluctuations in excitability and therfore a possible communication between these two networks is supressed. Adapted from Fries [2005].

5.3.4 The pattern classification.

The research on working memory have assumed for long that elevated signal intensity change during maintenance is the mechanism for WM storage (see chapter 2.2). Recently this assumption have been revisited [Postle, 2015]. One source of criticism cames from studies using the approach of multivariate pattern classification. The studies reported working memory dependent effects in the absence of the elevated delay activity which falsifies a long-standing belief that the elevated signal intensity change is necessary for coding of information in WM. This also made the method of pattern classification an interesting tool for working memory research. Let's consider an example to illustrate what might be a classification problem and how the method might help the brain research.

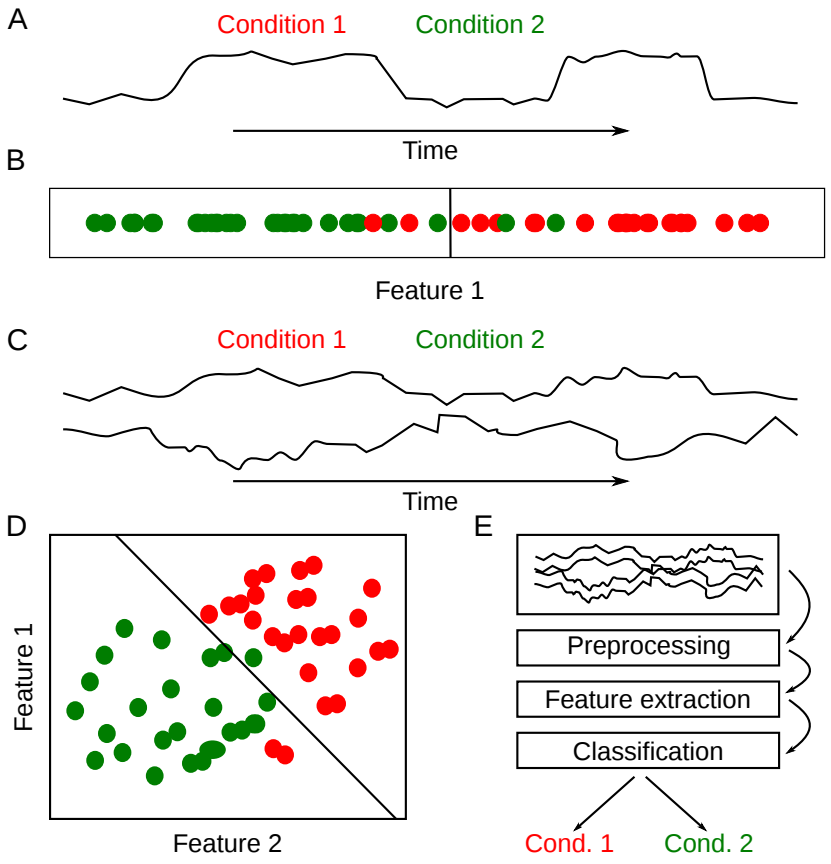

FIGURE 5.4: (A) Simulation of a simple data changing which varies as a function of experimental condition (Condition 1 and 2, red, green respectively). (B) Each dot represents one sample of the data observed for one of the two classes (Condition 1 and 2, red, green respectively). The position of each dot represents an intensity of some qunatitative measure (e.g. an intensity measure of an EEG potential). The two classes may be separated by setting a boundary between the distributions of points belonging to each of the classes. This is a very simple example with a one feature dissociating between the two classes. (C, D) In the case with two features a linear function in a two-dimensional plane may separate between the two classes. In the case of three features a two-dimensional plane may separate the clesses. (E) Illustration of the steps in pattern classification and a typical information flow indicated with arrows. The brain states to be classified are first recorded with an iEEG. The signals are preprocessed, features are extracted and finally the classification is performed (here either the Condition 1 or 2). Adapted from Duda et al. [2012].

Imagine you open a new email account. Initially you receive some spam. The email provider asks you to indicate unwanted information and after some time the number of spam emails gets lower. This is very likely due to pattern recognition algorithms which learn from your decisions to automatically classify an upcoming email as spam or not. How does it work? Each email might be characterized by several features (e.g. name of the sender, length of email, whether it contains your name or not, etc.). A pattern classification algorithm will creat a multi-dimensional vector across all these features to tell apart those emails that you indicated as wanted from these that you marked as spam. This is an example of supervised learning in which training dataset is provided with instances having accurate labels (spam vs. not a spam). Unsupervised learning would also search for patterns but only based on the inherent properties of the data without any labels provided. Imagine now that your task is to classify human behavior into remembering an object X and remembering an object Y. All you have is a number of electrodes inside of the human brain that record instantaneous electrical activity of the brain. The same set of algorithms might be used to try solving such a classification task.

There are four basic steps in the supervised learning approach: preprocessing (for the particualar case of the EEG signal see chapter 5.2, see also Figure 5.4), feature selection, training a classification model and testing generalization of the model [Duda et al., 2012]. Not every property of the incoming email is a good candidate with which an algorithm might efficiently predict if the email is spam or not. Thus, one has to select the most discriminative features and ignore those that contain little information for the classification. This is particularly important for the multi-dimensional datasets like brain recordings. For example, three seconds of iEEG recordings from a patient with 100 electrodes sampled at 1kHz would result in around 45000000 of potential features (i.e. 3000 milliseconds * 100 electrodes * 150 frequencies). Of course many of the time points and frequencies do not discriminate between the classes well. Therefore, one needs a method to limit the feature space. There are numerous approaches allowing reduction of the feature space and selection of most informative ones. For example, a conceptual

prior restriction on the feature space in frequency, time or anatomical (electrode) dimension. Alternatively the feature space can be reduced by data driven methods such as Principal Component Analysis or methods known from the inference statistics. A statistical method of the feature space reduction which is very popular in the cognitive neuroscience is an analysis of variance (ANOVA). In this case the number of features is reduced to those that pass a statistical threshold. The different classes serve as factors and the instances are independent observations. Features exceeding some threshold (usually $p < 0.05$) are treated as more discriminative and enter further analysis.

Training a classification model is the process during which an algorithm provided with data instances and associated labels develops a set of rules to predict unknown (untrained) instances (i.e. new, untrained trials of an experiment).

Finally, during testing a classifier is confronted with the "new" untrained instances of the data. Rules developed during learning part allows the classifier to make predictions regarding the class to which these new instances belong. A classifier is able to generalize to the new data if it predicts the classes of these new instances above chance level.

5.4 Materials and methods

5.4.1 Experiment 1. Maintenance of multiple familiar items in the hippocampus.

Participants. Data was recorded from a group of 9 pharmaco-resistant epilepsy patients (mean age of 35.8 years, ranging from 24 to 58, four female) implanted with intracranial EEG electrodes for diagnostic purposes. All patients were implanted bilaterally and had well-defined lateralized ictal onset zones (7 of them had a seizure onset zone within the left hippocampus, the other two in the right

hippocampus). Only electrodes localized in the seizure-free hemisphere were considered for analysis. Recordings were performed at the Department of Epileptology, University of Bonn, Germany. The study was approved by the local medical ethics committee, and all patients gave written informed consent.

Task. Participants performed a modified Sternberg WM task with digits serving as stimuli. Each trial contained a sequence of either one (load 1), three (load 3), five (load 5), or seven (load 7) digits. Each item was presented for 0.5s, followed by an inter-stimulus interval of a duration that was randomly jittered between 1.5s and 2s. Once the sequence had been presented, participants had to maintain the representation of the entire memory set for the duration of a retention interval lasting 3s. Next, a question mark informed participants to type all digits in correct order using a keyboard (see Figure 5.5A). Free recall of the entire sequences was tested because the hippocampus is known to be particularly relevant for the sequence memory [Eichenbaum, 2000, Kumaran and Maguire, 2006]. A total of 120 trials were presented (30 trials per load condition) in random order. During the experiment, iEEG signals from depth electrodes (see Figure 5.1A,B and Figure 6.1B) as well as from the linked mastoids were continuously recorded.

Recordings. Data from the hippocampus was recorded with multicontact depth electrodes. Each electrode had a diameter of 1.3mm, comprising 10 cylindrical platinum contact sites with a length of 2.5mm and an inter-contact spacing of 4mm. Data was sampled at 1kHz and referenced to linked mastoids. Only contacts contralateral to the epileptic focus were considered for analysis. Continuous data was visually inspected for artifacts (predominantly arising from epileptiform activity). All samples contaminated by artifacts were marked and removed from further analysis. For each patient, the recording sites were selected in a two-step procedure. First, the sites anatomically localized in the healthy hippocampus (anatomical criterion) were determined. To this end, I used individual post-implantation MRI scans and selected electrodes localized in the hippocampus based on anatomical landmarks [Henri M. Duvernoy, 2005, Naidich et al., 2007]. To further determine the anatomical location of each contact, I used post-implantation MRIs co-registered to the pre-implantation MRIs and normalized to

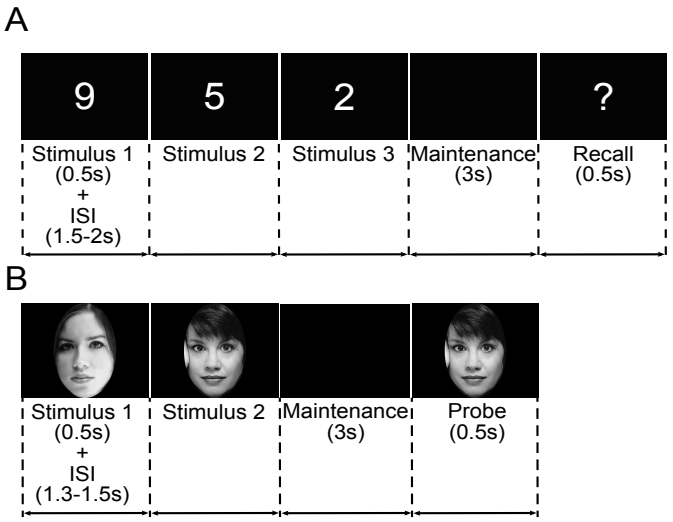

FIGURE 5.5: (A) Serial WM task with consecutive presentation of one, three, five or seven digits. Here an example of trial with three items is presented. Upon presentation of the question mark participants were instructed to type as many digits as they remembered in the correct order (B) Similar WM task used in the replication study. Participants were presented with sequence of one, two or three unfamiliar faces. Here an example of load two is presented. After the maintenance participants were presented with a probe face which was either one of the faces in the memory set or a new face (match, non-match, respectively). Here an example of a match trial is presented.

the standard MNI space. For the power analysis, I averaged across activity in all of these anatomically selected channels (5 ± 2 hippocampal channels [mean $\pm$ standard deviation]). Note that the results were qualitatively identical if a single channel was used (as for the analysis of phase data). For all analyses related to phase information, averaging across several channels might smear the results because of traveling wave properties [Bahramisharif et al., 2013, Lubenov and Siapas, 2009]. Thus, for phase analyses I selected one hippocampal channel in each patient with the highest amplitude of the hippocampal P300 component averaged across trials [Axmacher et al., 2010a].

5.4.2 Experiment 2. Maintenance of multiple novel items in the hippocampus.

Participants. A new group of patients (N = 8, mean age of 43 years, two female) implanted with electrodes in the hippocampus performed a WM task similar to Experiment 1. All patients had well-defined lateralized ictal onset zones (4 of them had seizure onset zone within the left hippocampus, the other four in the right hippocampus). There were three main differences regarding previous design: (1) faces of unfamiliar people served as stimuli; (2) participants were probed with a stimulus and were expected to indicate if the probed face was presented in the WM set (see Figure 5.5B) or if the face was new (two alternative forced choice task, 2AFC, instead of the whole report task); (3) memory load was 1, 2 or 4 faces. Timing of stimulus presentation and retention duration was the same as in Experiment 1. All preprocessing steps and the statistical analysis were same. Recordings were also performed at the Department of Epileptology, University of Bonn, Germany. The study was approved by the local medical ethics committee, and all patients gave written informed consent.

5.4.3 Experiment 3. Maintenance in the dorsal and the ventral visual stream.

Participants. Data was recorded from a group of 12 pharmaco-resistant epilepsy patients (mean age of 32.2 years, ranging from 18 to 55, six female) implanted with intracranial EEG electrodes for diagnostic purposes. All had well-defined ictal onset zones. Eight of them had a seizure onset zone within the left hemisphere, the other three in the right hippocampus and one had bilateral hippocampal onset zone. Only electrodes that were located outside of the focus zone were considered for analysis. Patients had electrodes implanted in the ventral (N = 12) and dorsal (N = 7) visual stream as well as in the frontal cortex (N = 5). Six patients had simultaneous recordings in the ventral visual stream and the prefrontal cortex. Three more had electrodes in the dorsal visual stream and the prefrontal cortex.

Recordings were performed in the Department of Epileptology, University of Bonn, Germany. The study was approved by the local medical ethics committee, and all patients gave written informed consent.

Task. Participants performed a delayed-matching-to-sample task assumed to engage either the ventral or the dorsal visual stream. Depending on the condition participants were expected to memorize the identity or spatial orientation of a face. Passive viewing with no memory load was used as a control condition. Each trial started with presentation of a face for 0.5s, followed by maintenance interval lasting 2.7s. Next, participants were presented with a second stimulus (0.5s) probing the accuracy of their memory (see Figure 5.6). This stimulus was either face (identity, control condition) or a box (orientation condition). The order of conditions was pseudo-randomized across participants. Participants were instructed to press one out of two buttons depending if they thought the probe matched or not the identity or orientation of the initial face. The response was not speeded, accuracy was emphasized. After the response, participants were provided with feedback. During the experiment, iEEG signals from the dorsal and the ventral visual stream as well as from the prefrontal cortex (see Figure 6.10) and the linked mastoids were continuously recorded.

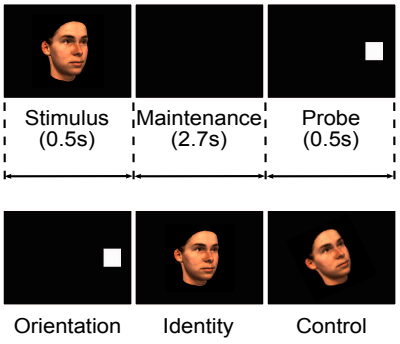

FIGURE 5.6: Trial sequence in experiment 3: maintenance in the dorsal and ventral visual stream. Single-item WM task (upper row). Depending on a condition participants were instructed to maintain either the orientation, identity or passively view the face. Subsequently, they were probed for the respective feature. The probe could either match or not the memorized feature. Here a match probe for orientation condition is presented. Lower panel shows example probes – a match probe during orientation and identity and a probe during control condition.

Recordings. Data was recorded using multicontact strip and grid electrodes. These electrodes were made of stainless steel and had contacts with a diameter of 4mm and an inter-contact center-to-center spacing of 10mm. All data was sampled at 1kHz and referenced to the linked mastoids. Only contacts non-overlapping with the epileptic focus were considered for analysis. Continuous data was visually inspected for artifacts (predominantly arising from the epileptiform activity). All samples contaminated by artifacts were marked and removed from further analysis. For each patient, I selected recording sites in a two-step procedure. First, electrodes were anatomically localized in the healthy part of the ventral, dorsal stream and prefrontal cortex based on anatomical landmarks (anatomical criterion). Next, I selected one contact per region of interest (ROI) using functional criterion that was defined based on prior findings [for the dorsal stream see Jokisch and Jensen, 2007] or properties of the recorded signal [the biggest N200 event related potential within the ventral stream: Allison et al., 1999, McCarthy et al., 1999]. Importantly, all criteria were orthogonal to tested hypotheses and provided an optimal signal to noise ratio for the given region of interest (ROI). For examples of similar approach see [Axmacher et al., 2010a, Leszczyński et al., 2015].

5.4.4 Experiment 4. Maintenance in the vlPFC.

Participants. Two adult rhesus monkeys (*Macaca mulatta*, 1 male and 1 female, 11.2 and 8.0 kg, respectively) were trained on a version of the delay-match to sample task. The monkeys were seated in a primate chair. They used left and right hand to move electromechanical levers which indicated task relevant behavior. The head was restrained and water container placed near to the mouth for automatic fluid reward delivery. The monkeys were housed in large stables ($32.04m^3$) which double the size of housing space required by the European Convention for the Protection of Vertebrate Animals used for Experimental and other Scientific Purposes (ETS No.123 App. A). Drinking was controlled and water deprivation was performed strictly according to the rules set by the German Primate Center. A minimum fluid intake of more than 20ml/kg*day was followed. All procedures

were approved by the local authorities and are in full compliance with the guidelines of the European Community (EUVD 86/609/EEC) for the care and use of laboratory animals. The recordings were made available to me by courtesy of Dr. Matthias Munk.

Task and training. Each trial started with presentation of a memory set sample (see Figure 5.7). Depending on the memory load the sample consisted of 1, 2 or 3 items presented in a center of a screen. The whole memory set was always displayed within 0.9s (irrespectively of the load condition). Maintenance interval followed with a blank screen presented for 3s during which monkeys needed to keep the memorized set vividly represented. Finally, a probe was presented for 0.9s. The probe either repeated one of the items (match-probe) or consisted of a stimulus not presented in the memory set (non-match probe). Monkeys were trained to indicate if the probe matched or not the memory set by pressing left or right lever, respectively. Correct responses were rewarded with liquid. Stimuli were presented in the center of a 21-inch screen and consisted of familiar set of twenty colored fruit and vegetables drawings (each spreading around 6° - 9° of the visual angle, see Figure 5.7). Both monkeys were trained for many months on the simplest version of the task (load 1 with a single item as memorandum). Once monkeys' performance was above 80% of trials correctly performed, higher memory loads were introduced. The data was recorded with a grid of 4x4 tetrodes, sampled at 1kHz and referenced to a part of electrode tube touching the dura mater.

Pattern classification. For the multivariate pattern classification I used the support vector machine (SVM) classifier with 7-folds cross-validation procedure. I used independent samples t-test and ANOVA (threshold at $p < 0.01$) to reduce the feature space in two- (match vs. non-match) and three-class (load-dependent analysis) classification, respectively. I furthermore selected 500 most promising features showing highest t- or f-value. All classification analyses were tested against a distribution of very conservative surrogate data. To this end, the same pattern classifier was developed and the reference distribution of surrogate data was obtained by shuffling trial labels of the test set. This is a conservative approach

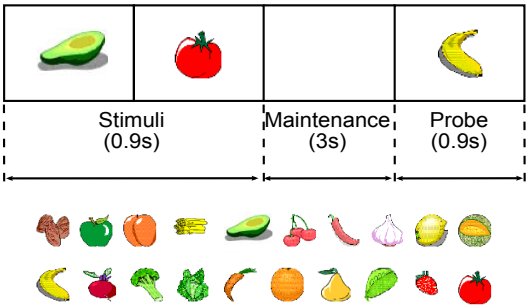

FIGURE 5.7: Trial sequence in experiment 4. Serial WM task with consecutive presentation of one, two or three objects. Here an example trial with two items is presented (upper panel) and a set of stimuli used in the experiment (lower panel). After the maintenance each monkey was presented with a memory probe which was either one of the objects in the memory set or a new object (match, non-match response, respectively). Here an example of a non-match trial is presented.

because training data and the whole training procedure is intact. This procedure was repeated 5000 times generating the reference distribution under the null hypothesis that the classifier did not learn to generalize to new data. If the classifier performance exceeded the 95th percentile of this distribution (corresponding to $p < 0.05$), I considered the classification model as being able to generalize to new dataset and therefore to decode a trial category above chance level.

Chapter 6

Results

6.1 Maintenance of familiar information in the hippocampus (Experiment 1).

Participants were presented with sequences of 1, 3, 5, or 7 digits maintained for 3 seconds. Subsequently, they were probed to recall the entire sequence in a correct order (see Figure 5.5A). During the experiment, intracranial EEG was continuously recorded from hippocampi of epilepsy patients, because previous results from neuroimaging [Nichols et al., 2006] and intracranial EEG [Axmacher et al., 2010b, van Vugt et al., 2010] have shown that the hippocampus activates during WM maintenance of multiple items or complex stimuli (e.g. faces). Furthermore, studies have shown that patients with hippocampal lesions do have deficits when they are asked to maintain multiple items or associations between multiple item features [Aggleton et al., 1992, Nichols et al., 2006, Olson et al., 2006]. Thus, these studies suggest that the hippocampus is not only activated during WM tasks, but is necessary under certain conditions.

6.1.1 Behavioral results

Accuracy (percentage of correct trials) and WM capacity was measured. First, accuracy as a function of WM load was calculated. One-way repeated measures ANOVA with "memory load" (1, 3, 5, 7 items) as factor revealed a significant main effect ($F(3,24) = 107.3$, $p < 0.0001$). Post-hoc paired-samples t-tests showed that accuracy in memory load 1 (93%) was higher than in load 3 (87%; $t(8) = 4.1$, $p < 0.005$); accuracy in load 3 was higher than in load 5 (78%; $t(8) = 4.2$, $p < 0.005$); and accuracy in load 5 was higher than in load 7 (23%; $t(8) = 9.6$, $p < 0.00001$; see Figure 6.1). Individual WM capacity was calculated as the maximum number of correctly maintained items, controlled for guessing (K_{max}). The K_{max} was at 5.54 ± 0.20 (mean $\pm$ S.E.M.) and varied between participants from 4.7 to 6.3. These results indicate that participants were able to maintain around 5 digits in a sequence.

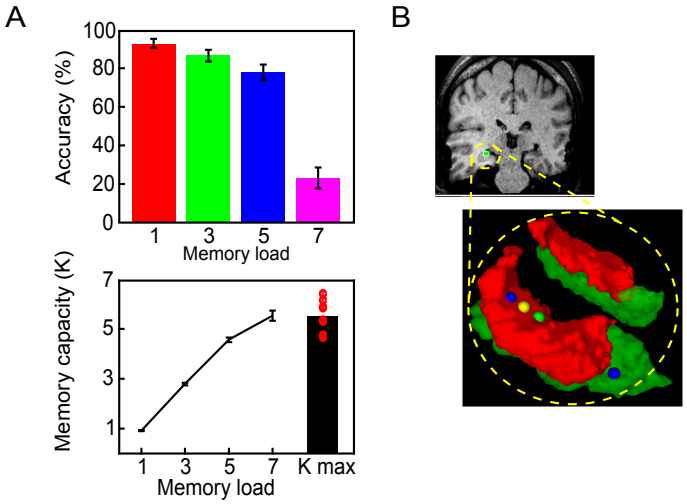

FIGURE 6.1: Behavioral results (Experiment 1) and the hippocampal electrode reconstruction. (A) Accuracy scores (upper panel) and WM capacity (lower panel) as a function of memory load and the maximal value of capacity (K_{max}). Dots indicate single-subject scores. Error bars show the standard error of the mean (SEM). (B) Coronal slice from one patient with the hippocampal electrode contacts which are marked by spheres. A mask from the AAL atlas was used for identification of electrodes in the hippocampus (red; the green mask indicates parahippocampal gyrus).

6.1.2 Time-frequency activity patterns

First, the activity patterns in the lower frequency range (4 - 30Hz) were inspected. I used surrogate-based cluster statistics to identify significant linear changes in power across memory loads (Maris and Oostenveld, 2007). Four significant clusters ($p_{corrected} < 0.05$ each) were observed in which power decreased linearly with increasing memory load (see Figure 6.2A). These four clusters had an average duration of 208 ± 35ms (mean ± standard deviation across the four clusters), covered frequencies in the alpha/beta range between 7Hz and 23Hz, and were distributed across the entire maintenance period with intervals of around 540ms between them (see Figure 6.2). Thus, the load dependent decrease of alpha/beta power was organized into discrete short and repetitive time intervals rather than being sustained over the entire maintenance period.

These periods of the linear power decrease were periodically interleaved with intervals without any indication of load-dependent power decrease (see Figure 6.2B). Abundant evidence has shown that decreases of alpha/beta power reflect release from functional inhibition [Bonnefond and Jensen, 2012, Haegens et al., 2011, Jensen et al., 2012, 2014, Jokisch and Jensen, 2007, Klimesch, 2012, Spaak et al., 2012, Thut et al., 2006, Waldhauser et al., 2012]. These results might be interpreted as reflecting a fluctuation between periods when the hippocampus shows a load-dependent low frequency power decrease and other periods during which a power is maintained constant across all load conditions. These periods of load-dependent reductions of the alpha-band activity may reflect an activation of working memory representations in the hippocampus (i.e. *memory activation*). This interpretation is further supported with results from incorrect trials and correlation analysis investigating individual capacity limits.

In addition to these effects in the alpha/beta range, and consistent with the interpretation of a load-dependent activation of the hippocampus, an increase in the gamma-band (30 – 150Hz) activity with load was observed. Because rapid power fluctuations in high frequency bands (>30Hz) can render cluster-based statistical methods ineffective, previous studies [Axmacher et al., 2007, Jokisch and Jensen, 2007] have quantified time-frequency effects in the high-frequency range by using

the analysis of variance on the data averaged over time and frequency. Therefore, I averaged power values in three consecutive (non-overlapping) time windows during the maintenance interval for the statistical analysis across the frequency range between 31Hz and 150Hz. Next, I calculated 3 x 4 repeated measures ANOVA with "time window" (1-1000ms, 1001-2000ms, 2001-3000ms) and "memory load" (1 item, 3 items, 5 items, 7 items) as factors and again analyzed the linear trends. This analysis revealed no main effect of "memory load", but a significant interaction between the "memory load" and "time interval" ($F(6,48) = 3.66$, $p < 0.005$). To further explore this effect, I calculated one-way ANOVAs separately for each time window, which revealed a main effect of memory load specifically during the last second of the maintenance phase ($F(3,24) = 3.80$, $p < 0.05$). In this time period, a significant linear increase in power across the memory loads was also observed ($F(1,8) = 4.90$, $p = 0.05$). In summary, the gamma-band activity linearly increased as a function of memory load specifically during the late maintenance phase, putatively reflecting increased hippocampal recruitment with rising memory load.

6.1.3 Phase-amplitude coupling

One of the predictions from the multiplexing model of WM posits interactions between the phase of low-frequency oscillations and the power of high-frequency activity as a mechanism for multi-item WM. A significant increase in the CFC as compared to the surrogate data was observed during the entire maintenance phase. The difference was found in a broadband frequency range of 2 – 16Hz for phase and 4 – 150Hz for amplitude ($t(8) = 3.96$, $p_{corrected} < 0.05$), and also within a smaller range of 5 – 9Hz for phase and 30 – 150Hz for amplitude that was predicted by the original model of the CFC ($t(8) = 3.99$, $p_{corrected} < 0.05$).
The initial observation that the load-dependent decrease of alpha/beta power (*memory activation*) occurs periodically in distinct temporal clusters led to a hypothesis that the CFC might show a similar pattern of periodic fluctuations. To this end, the CFC was calculated within 200ms intervals centered on significant

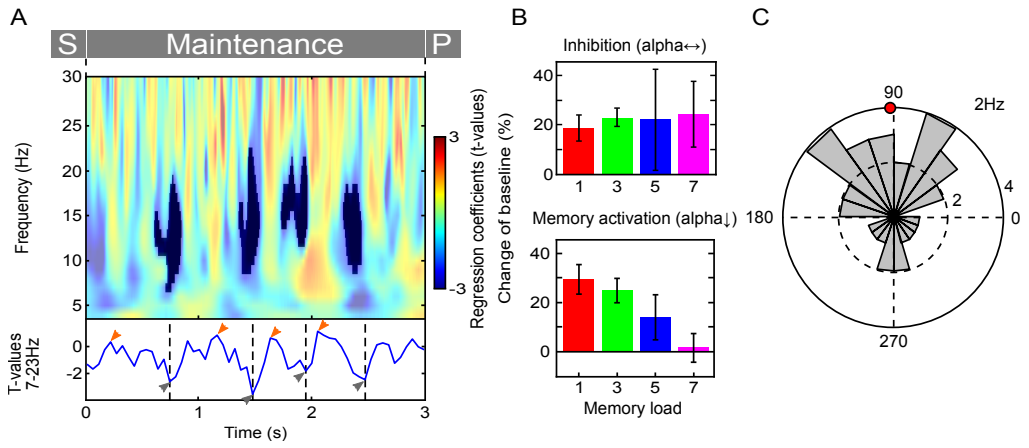

FIGURE 6.2: Load-dependent periodic power decreases. (A) Linear power decreases resulting from t statistic for regression coefficients calculated for each participant. The results were obtained by regressing the subject-specific spectro-temporal data on WM load. For each subject, data were averaged across trials of the same condition (loads 1, 3, 5, and 7). The color map reflects the results of this random-effects group-level t statistic (N = 9) showing that the power fluctuations calculated within each subject are consistent across the pool of patients. Bottom: time course of the linear regression averaged across frequencies where significant clusters were observed (i.e., between 7 and 23 Hz). Color arrows represent distinct periods, with gray depicting memory activation periods and orange reflecting interleaved intervals. The horizontal gray bar represents the temporal structure of the paradigm (S, presentation of the last stimulus; P, probe for recall). For the detailed analysis of single patients see supplemental Figure S1 in Leszczyński et al. [2015]. (B) Averaged alpha/beta power (7–23Hz) during periods with (bottom) and without (top) load-dependent power decrease (memory activation and interleaving periods, respectively) showing that the difference between the two types of intervals is due to different distributions of power across memory loads. Error bars indicate SEM. (C) Circular histogram shows clustering of the endogenous 2Hz oscillation phases during periods with load-dependent memory activation (polar plot with 20 phase bins, each covering 18°). Data pooled across nine patients and four clusters.

load-modulated alpha power reductions (*memory activation*) and interleaved intervals during which power was constant across memory loads. Indeed, the CFC strength was found to follow the alpha/beta load effects (Figure 6.3). To quantify this effect, CFC strength was directly compared during periods with and without significant load-dependent decreases of alpha/beta power ($t(8) = 2.3699$, $p < 0.05$; see Figure 6.3). These results were not biased by differences in signal-to-noise ratio, as average power did not differ between the load-dependent and load-independent intervals ($t(8) = 0.6126$, $p=0.5571$).

To summarize, the current data show that the CFC strength is lower during periods in which the alpha/beta-related inhibition is released with WM load. During these periods, the CFC may support a process of *memory activation* in which modulated high-frequency activity (mHFA) is distributed across broader phase ranges of modulating low-frequency activity (mLFA), supporting maintenance of multiple items.

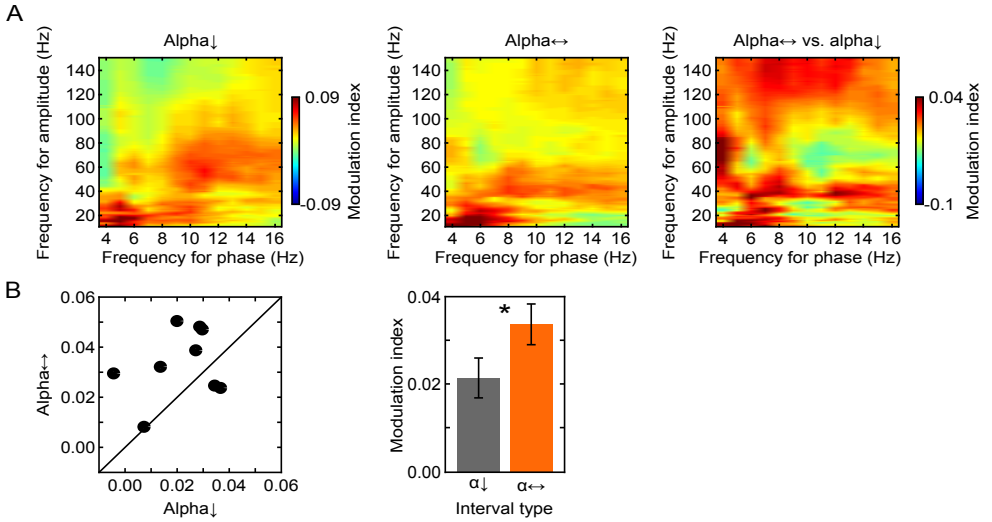

FIGURE 6.3: CFC difference across the two types of interval. (A) Grand average CFC during time intervals with (left) and without (middle) load-dependent release from alpha/beta-related inhibition and the difference between these periods (right). Color indicates CFC index. Higher CFC during periods without load-dependent disinhibition was observed. (B) The CFC strength is reliably lower during *memory activation* periods than the interleaved periods. This is evident by most of the dots being above the diagonal. Left: each dot reflects data from a single subject with CFC calculated separately during time intervals with load-dependent power decrease (x axis) and without (y axis). Right: bar plot reflects group average of the CFC during intervals with (gray) and without (orange) load-dependent power decrease. Error bars reflect SEM. The star indicates $p < 0.05$.

6.1.4 Relation to behavior

One of the predictions derived from the multiplexing model of WM is that individuals with higher peak frequency for amplitude (mHFA-max) or lower peak frequency for phase (mLFA-max) or higher ratio (CFC ratio) of both will show

increased working memory capacity limit. This is due to the assumption that nesting of high-frequency oscillations in the slow-frequency phase is the mechanism for maintenance of information in WM (see Figure 2.3). To test this prediction I calculated inter-individual correlation between individual WM capacity and CFC ratio, mLFA-max, and mHFA-max (Figure 6.4). These analyses were conducted separately for periods with and without load-dependent alpha/beta power effects (i.e., for *memory activation* periods and for the interleaved periods). The correlations were performed using Kendall's tau which is a particularly conservative rank correlation test to investigate correlations in small populations. It does not assume that the data are normally distributed [Abdi, 2007].

During the *memory activation* periods, a significant positive correlation between mHFA-max and WM capacity was observed (Figure 6.4; Kendall's tau = 0.59161, $p < 0.05$). No such relationship was found for the mHFA-max during periods without load-dependent release from inhibition (all tau < 0.2, all $p > 0.6$). Neither mLFA-max nor CFC ratio correlated with capacity during *memory activation* (all tau < 0.2, all $p > 0.5$) or during interleaved intervals (all tau < 0.6, all $p > 0.1$). In other words, patients with higher WM capacity showed modulations of the gamma-band activity at higher frequencies specifically during the *memory activation* periods (see Figure 6.4).

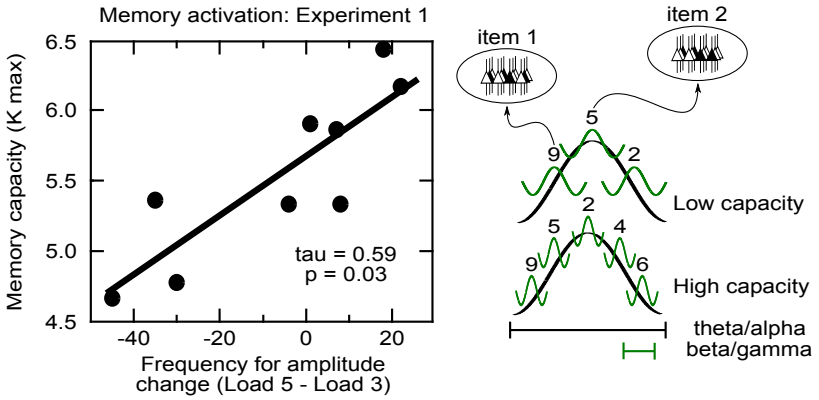

FIGURE 6.4: Behavioral relevance of the CFC. Left: during the *memory activation* periods, individual WM capacity correlates with load-dependent changes in the frequency for amplitude. Right: This correlation could be explained by faster cycles of high-frequency activity representing individual items in a sequence and therefore increasing individual WM capacity.

6.1.5 Delta rhythmicity.

The clusters showing load-dependent reductions of alpha/beta power occurred rhythmically about every ∼500ms. Fast Fourier transform of the time course of the t-scores for regression values indeed showed a peak in the power spectrum at 2Hz; indicating a dominant fluctuations at 2Hz. The details of this analysis are presented in the supplementary materials attached to my article [Leszczyński et al., 2015]. Four clear cycles of load-dependent power reduction were observed during the maintenance interval (Figure 6.2A). Next, I tested whether the fluctuations of intervals showing the release from alpha-dependent inhibition are consistently related to the phase of the endogenous slow oscillations [Lakatos et al., 2005]. Thus, I used the Rayleigh's test for circular data to investigate if the phase angles of 2Hz oscillations during the disinhibition intervals were non-uniformity distributed. This test showed that the 2Hz phases were indeed clustered ($z = 3.19$, $p < 0.05$) and showed a preferred direction of 90° (Figure 6.2C). Thus, intervals of *memory activation* occur during consistent phases of the endogenous 2Hz oscillations, indicating a hierarchy of hippocampal oscillations during working memory. More detailed analysis of the underlying delta rhythmicity may be found as the supplemental information to my article [Leszczyński et al., 2015].

6.1.6 Rhythmic duty cycles are critical for successful WM performance in the hippocampus.

To further assess the functional relevance of the observed periodicity, the analysis of power and the CFC were performed during incorrect trials (see Figure 6.5A). In these trials, no load-dependent reduction of alpha/beta power was observed. Interestingly, this was not due to an overall reduction of CFC, which was increased as compared to correct trials (0.033 vs. 0.022: $t(14) = 2.07$, $p = 0.05$ in Experiment 1). This suggests that the temporally distributed reduction of CFC rather than the absolute CFC strength is critical for successful WM performance.
Together, these results suggest that the hippocampus is indeed involved in WM

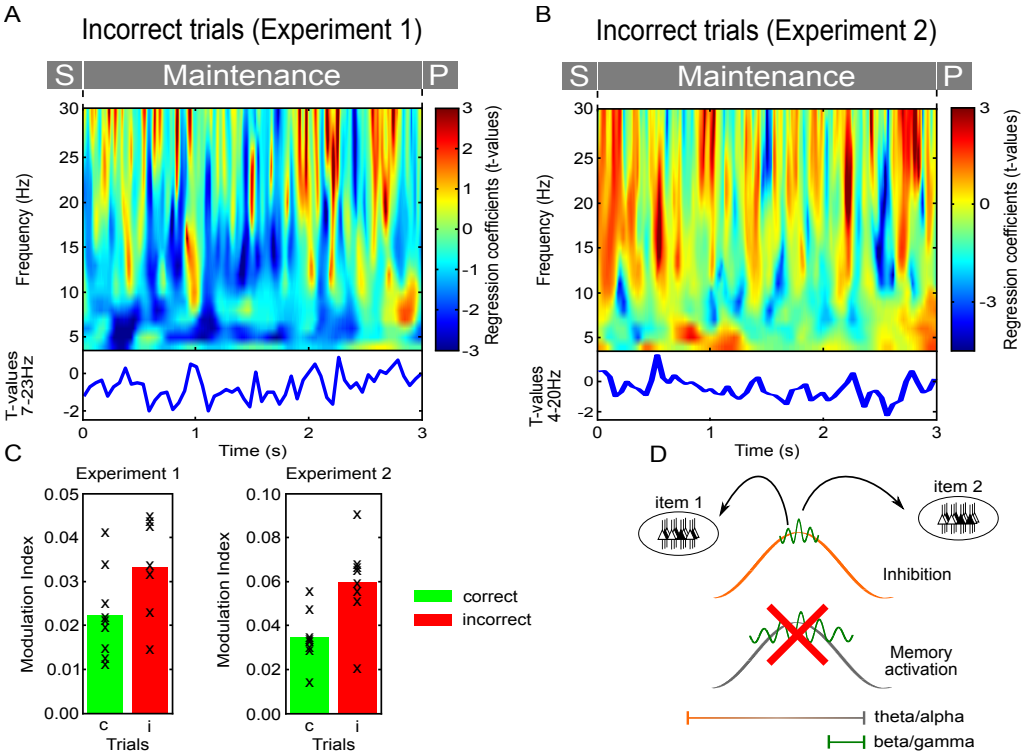

FIGURE 6.5: (A and B) Top: regression coefficients indicating no linear decrease in power with memory load in experiment 1 (maintenance of multiple numbers) and 2 (maintenance of multiple faces), respectively. Bottom: time course of averaged t values for regression coefficients between 7 and 23 Hz (A) and 4 and 20 Hz (B) in experiments 1 and 2, respectively. No significant differences were found. (C) Bar plots indicate increased CFC for incorrect (red) as compared to correct trials in the experiment 1 (left) and experiment 2 (right). Each "X" represents data from one single subject. (D) Schematic depiction of incorrect trials with lack of memory activation periods. Absence of the linear power decrease in the low-frequency range. There are no periods during which cycles of high-frequency activity are broadly distributed across larger phase ranges of low-frequency oscillations (corresponding to the lower levels of CFC).

maintenance. It shows load-dependent low frequency power reduction and high gamma activity linear increase. These results likely reflect increased neural assemblies being involved to represent sequence of increasing length. Furthermore, these changes are not static but undergo periodic fluctuations between two distinct modes which are organized by the endogenous delta rhythm. The reduction of the CFC strength and low frequency power is critical for successful WM performance. Finally, an individual load-dependent change in the frequency for amplitude predicts the number of items an individual is capable to remember. Subsequently, these results were successfully replicated with another WM task.

6.2 Maintenance of unfamiliar information in the hippocampus (replication study, Experiment 2).

The periodicity described above was a very surprising result. In fact most models of working memory rather assume that maintenance is a steady, uniform process. Because this finding was so unexpected, the second experiment aimed to replicate the observed effects. The second experiment differed according to three elements compared to the previous design (see Methods 5.4). One important difference was the stimuli used. Instead of highly familiar and simple features like digits the current experiment used complex images of unfamiliar faces. This difference is important because the hippocampus involvement in working memory has been suggested to depend on the novelty of the maintained materials [Hasselmo and Stern, 2006].

6.2.1 Behavioral results.

One-way repeated measures ANOVA with memory load (1, 2, 4 items) as factor revealed a significant drop in accuracy across memory loads ($F(2,14) = 4.18$, $p < 0.05$; see Figure 6.6). A post-hoc paired-sample t-test showed that accuracy in the load 1 condition was higher than accuracy in load 4 ($t(7) = 2.42$, $p < 0.05$). There was no difference between accuracy in load 1 and 2, and between load 2 and 4 (all $t < 1.85$, all $p > 0.10$). Individual WM capacity K_{max} was at 1.79 ± 0.80 (mean $\pm$ S.E.M.) and ranged between 0.53 and 3.33 (see Figure 6.6). The lower values of accuracy and capacity (compared to the Experiment 1) likely reflect the difference in the overall difficulty between the maintenance of simple digits and complex faces.

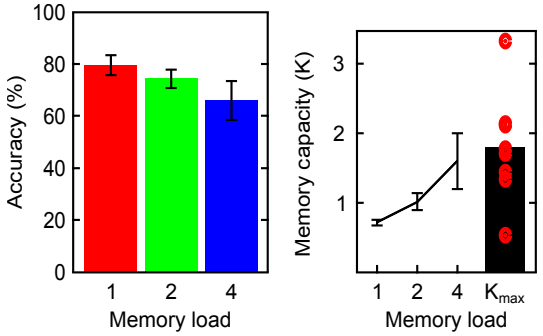

FIGURE 6.6: Behavioral results (replication study, Experiment 2). Accuracy scores (left) and WM capacity (right) as a function of memory load and the maximal value of capacity across load conditions (K_{max}). Dots indicate single-subject scores. Error bars show SEM.

6.2.2 Time-frequency activity patterns.

The same surrogate-based cluster statistics was performed as in the first experiment to identify significant linear changes in low frequency power across memory loads. Three significant clusters ($p_{corrected} < 0.05$ each) were found in which power decreased linearly with increasing memory load (see Figure 6.7). These three clusters had an average duration of 627 ± 91ms (mean ± standard deviation across the three clusters), covered frequencies in the theta/alpha/beta range between 4Hz and 20Hz, and were distributed across the maintenance period with a center to center distance of 660ms. Thus, similar to Experiment 1, the load dependent decrease of low frequency power was organized into discrete short and repetitive time intervals rather than being sustained over the entire maintenance interval.

Next, I tested if gamma power also increased as a function of memory load. To this end, a similar analysis of variance on broadband gamma (31-150Hz) was performed. I calculated a 3 x 3 repeated measures ANOVA with "time window" (1-1000ms, 1001-2000ms, 2001-3000ms) and *memory load* (1 face, 2 faces, 4 faces) as factors. The analysis revealed significant linear effects of memory load $F(1,7) = 7.65$, $p < 0.05$) and interval ($F(1,7) = 8.28$, $p < 0.05$) but no interaction ($F < 1.9$, $p > 0.2$). Thus, the result of linear increase in the gamma-band power activity with memory load was replicated. This effect occurred throughout the entire maintenance phase and not only during the last second of the maintenance period

(even though it was numerically strongest in this interval). Note that gamma effects in both experiments were independent of the periodicity observed in low frequencies.

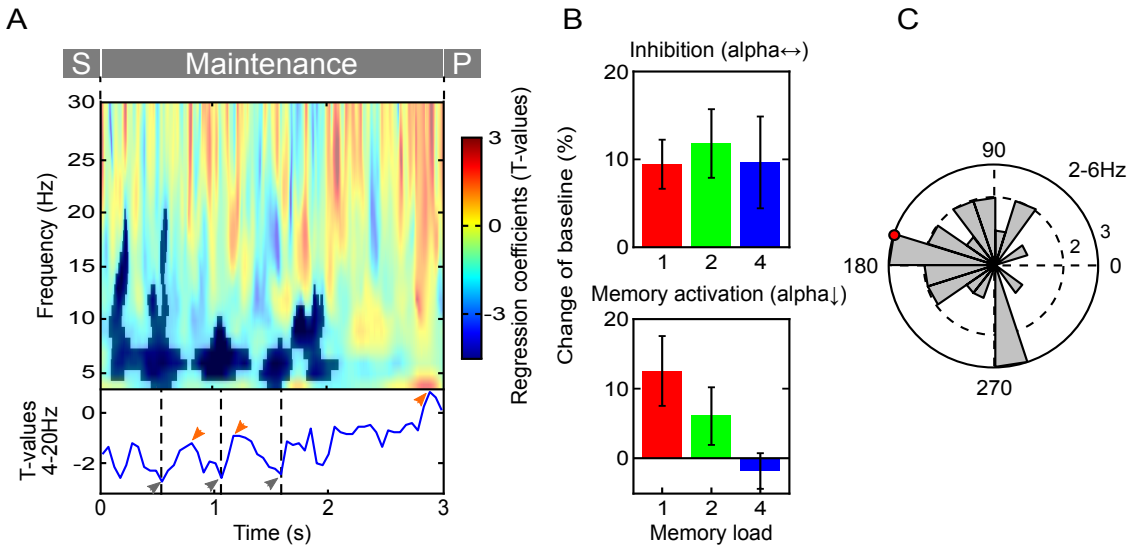

FIGURE 6.7: (A) Periodic fluctuations between periods with and without load-dependent release from low-frequency-related inhibition. Top: regression coefficients indicating load effects on theta/alpha/beta power effect. Bottom: time course of averaged t values for regression coefficients between frequencies where clusters were observed (4–20 Hz). Color arrows represent distinct periods, with gray depicting memory activation and orange reflecting interleaved intervals. (B) Averaged alpha/beta/theta power (4–20 Hz) during periods with (bottom) and without (top) load-dependent power decrease (memory activation and interleaving periods, respectively). Error bars indicate SEM. (C) Circular histogram showing clustering of endogenous 2 to 6 Hz oscillation phases during periods with load-dependent power decrease (memory activation). Data pooled across eight patients and three clusters.

6.2.3 Phase-amplitude coupling.

Next, the strength of CFC during memory activation and the interleaved intervals was compared. It was again reduced during *memory activation* periods as compared to the interleaving intervals ($t(7) = 2.57$, $p < 0.05$; see Figure 6.8). Again, I excluded that the CFC was artificially driven by differences in the signal-to-noise ratio. No difference in load independent power was observed across two types of interval: $t(7) = 0.79$, $p = 0.4551$). These results replicate what was observed in

the first experiment showing that the CFC strength differs across the two types of interval and is attenuated during *memory activation* intervals.

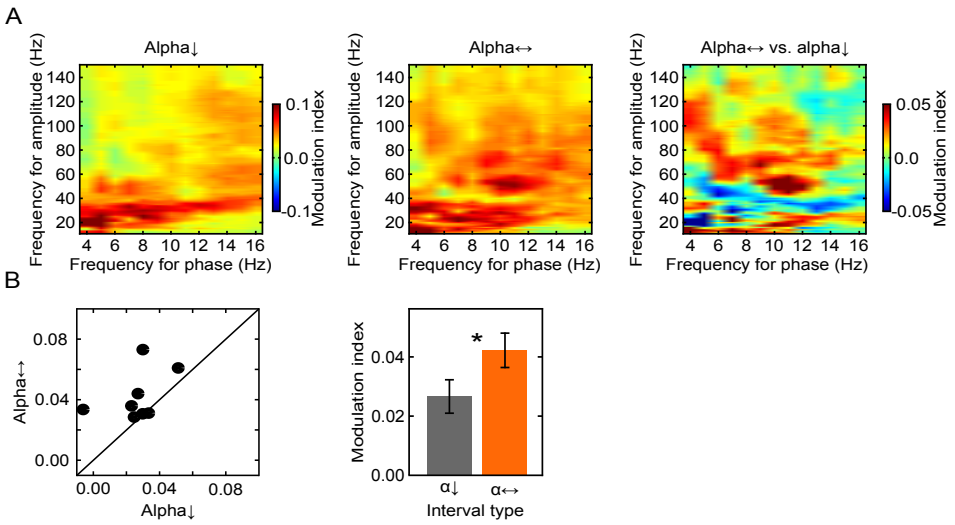

FIGURE 6.8: (A) Grand average CFC during time intervals with (left) and without (middle) load-dependent release from theta/alpha/beta-related inhibition (*memory activation* periods) and the difference between these periods (right). The color indicates CFC index. Higher CFC during periods without load-dependent disinhibition was observed. (B) The CFC strength is lower during memory activation periods than interleaved periods. This is evident by most of the dots being above the diagonal. Left: each dot reflects data from a single subject with the CFC calculated separately during the time intervals with (x axis) and without (y axis) the load-dependent power decrease. Right: bar plot reflects group average of the CFC during intervals with (gray) and without (orange) load-dependent power decrease. Each line depicts a single subject. Error bars reflect SEM. The star indicates $p < 0.05$.

6.2.4 Relation to behavior.

Finally, the behavioral relevance of the CFC for WM capacity was tested. The results exactly reproduced those observed in the previous experiment. Again, during the *memory activation* intervals, a positive correlation between the frequency of mHFA-max and WM capacity was observed (Kendall's tau = 0.62, $p < 0.05$; Figure 6.9). Thus, higher WM capacities are related to higher frequencies of mHFA-max, which corresponds to shorter cycles of the amplitude-modulated high-frequency oscillations. Putatively, longer memory sets can be maintained when individual

items are represented by faster mHFA cycles. No correlation between WM capacity and mLFA frequency (tau = 0.38, p = 0.26) or CFC ratio (tau = 0.214, p = 0.54) in both types of interval was observed.

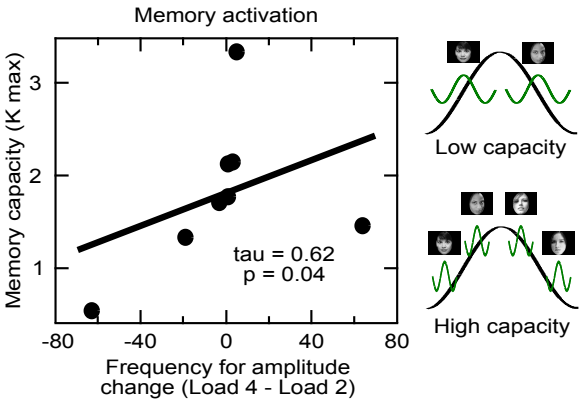

FIGURE 6.9: Left: during memory activation periods, individual WM capacity correlates with the load-dependent changes in the frequency for amplitude. Right: This correlation might be explained by faster cycles of the high-frequency activity representing individual items in a sequence and therefore increasing individual WM capacity.

To further explore the behavioral predictions of the multiplexing model the phase-phase coupling was considered. In particular, the multiplexing model of WM assumes that the neural assemblies are synchronized in the gamma cycle. During each gamma cycle the neural assemblies represent individual items maintained in memory (see Figure 2.3). Thus, the model predicts that the consecutive gamma cycles should be locked to very specific phases of the theta oscillation. This locking of gamma cycles might serve as a mechanism preventing interference across individual item representations within the memory set. Note that unlike phase-amplitude coupling reported above which looks at the distribution of gamma activity across phase of theta, the phase-phase coupling quantifies the precision with which the gamma cycles are locked to the consecutive theta phases. These are two related yet distinct phenomena. In particular, the gamma activity might be non-uniformly distributed across the whole theta cycle (high phase-amplitude coupling) but consecutive gamma cycles might not be precisely related to distinct phases of the theta (low phase-phase coupling). More detailed results on the phase-phase coupling observed in the current data set are reported by Chaieb

et al. [2015]. Importantly, these results show that the maintenance is associated with increased hippocampal phase-phase coupling for multiple items and decreased coupling for a single item which supports the concept of hippocampal involvement in the multi-item WM. Finally, the correlation between individual WM capacity limit and the coupling factor was calculated and quantified with the Kendall's tau. Increased capacity is associated with higher coupling factor across memory loads (tau=0.59, p=0.009). Therefore, this result converges on the findings reported above relating the frequency-for-amplitude to capacity limits. It further extends the current results by showing that the precise locking of high-frequency cycles to the consecutive phases of the low-frequency is also a factor limiting individual WM capacity.

6.2.5 Delta rhythmicity.

Next, the periodicity of *memory activation* intervals was again tested for any relations to the phase of endogenous slow frequency oscillations. A Fast Fourier transformation of the time series of regression t-scores revealed a peak in the spectrum between 2 and 6Hz. I extracted the phases at this frequency range and analyzed if phase values during memory activation periods were again clustered. Indeed, a Rayleigh test showed that the phase values were not uniformly distributed ($z = 5.21$, $p < 0.005$), but were clustered at around 164° (Figure 6.7C). This again replicates previous results showing that memory activation intervals occurred during consistent phases of underlying slow oscillations. This also reproduced the finding of an oscillatory hierarchy from the first experiment. For more information on the observed periodicity see supplemental information and Figure S4B, C in my article [Leszczyński et al., 2015].

6.2.6 Rhythmic duty cycles are critical for successful WM performance in the hippocampus.

Similar to previous experiment I also looked at incorrect trials. Again, no effect of load-dependent reduction of alpha/beta power was observed. Exactly as in experiment 1, the CFC was increased as compared to correct trials (0.059 vs. 0.034: $t(14) = 3.04$, $p < 0.01$). This suggests that the temporally distributed reduction of the CFC rather than the absolute CFC strength is critical for successful WM performance which replicates the results of Experiment 1 (see Figure 6.5B).

6.3 Maintenance of information the dorsal visual stream (Experiment 3).

Participants were presented with a single image of a face. Each of the faces were rotated -60°, -30°, 30°, 60° (relative to vertical axis) and 30°, 0°, 30° (relative to horizontal axis). Depending on the condition participants were instructed to maintain either the identity or orientation of the face. Passive viewing with no memory load served as a control condition. Subsequently, a second image was shown to probe the feature which was relevant for a given condition (see Figure 5.6). During the experiment, intracranial EEG was continuously recorded from the electrodes localized in the ventral (N=12), dorsal (N=7) visual stream and from the frontal cortex (N=6). Importantly, 6 patients had simultaneous recordings in the ventral visual stream and the prefrontal cortex. Three more had electrodes in the dorsal visual stream and the prefrontal cortex.

6.3.1 Behavioral analyses.

The accuracy (percentage correct) was measured for all three conditions ("orientation", "identity", "control"). One-way repeated measures ANOVA with memory condition as a factor revealed a significant main effect ($F(2,12) = 5.02$, $p < 0.05$).

Post-hoc paired-sample t-test showed that accuracy in "identity" was lower than control (87% vs. 95%; t(6)=3.167, $p < 0.05$). No other difference was observed (all $t < 1.56$, all $p > 0.15$). This shows that memory load in identity condition was higher as compared to control. Note that this analysis is only restricted to the group of patients (N = 7) with electrodes implanted in the dorsal visual stream. The results depicted in the Figure 6.10 reflects the whole group of patients (N = 12) recorded for the purpose of the current experiment.

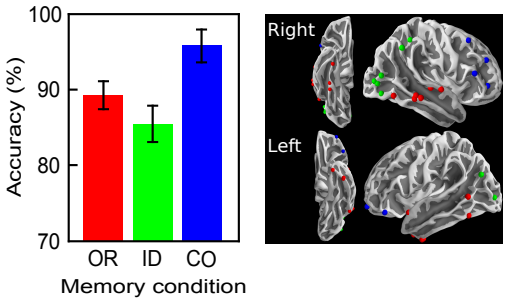

FIGURE 6.10: Behavioral results (maintenance in the dorsal and ventral visual stream, Experiment 3). Accuracy scores (left) as a function of memory condition (OR, ID, CO reflect maintenance of orientation, identity and control, respectively). The group overview of all dorsal, ventral and prefrontal electrodes (green, red, blue, respectively). Error bars show SEM.

6.3.2 Time-frequency activity patterns in the dorsal visual stream.

Next, I investigated whether the alpha frequency range (9-13Hz) activity patterns in the dorsal visual stream depend on working memory condition. This a priori hypothesis was derived from the previous MEG study by Jokisch and Jensen [2007]. The authors observed that the alpha activity over the parieto-occipital MEG sites was decreased during the maintenance of orientation as compared to identity and control. I used surrogate-based cluster corrected for multiple comparisons t-test (Maris and Oostenveld, 2007) to identify significant changes in the alpha power during memory maintenance across three conditions. Power was decreased during maintenance of orientation as compared to identity and control ($p_{corrected} < 0.05$;

see Figure 6.11). Importantly, no significant difference was observed between the alpha power during maintenance of identity and control. Although the effects were sustained through the entire maintenance interval, the clusters of significant change were most pronounced during an early part of the retention interval (around 500-900ms and 500-800ms after the stimulus offset for the orientation-control and orientation-identity, respectively). This replicates previous results of Jokisch and Jensen [2007] associating maintenance of orientation in WM with decreased alpha activity in the dorsal visual stream. As described above abundant evidence has shown that the alpha activity reflects functional [Bonnefond and Jensen, 2012, Jensen et al., 2012, 2014, Jokisch and Jensen, 2007, Klimesch, 2012, Thut et al., 2006] and neural [Haegens et al., 2011, Spaak et al., 2012] inhibition. Consequently, the decrease of power in the alpha frequency range relates to the release from neural inhibition (i.e. disinhibition). The current results add to evidence that activity in the alpha band reflects allocation of resources by inhibiting and disinhibiting activity in the dorsal visual stream (see Figure 6.11).
To quantify differences in the gamma frequency range I used 3x3 repeated measures ANOVA with memory condition (orientation, identity and control) and time interval (1-900ms, 901-1800ms, 1801-2700ms) as factors. Neither main effects (all $F < 0.54$, all $p > 0.54$) nor interaction was significant ($F(4,24) = 0.54$, $p=0.70$).

6.3.3 Intra-regional CFC in the dorsal visual stream.

Roux and Uhlhaas [2014] extended the multiplexing model of WM suggesting that theta-gamma and alpha-gamma cross-frequency coupling might provide distinct mechanisms for maintenance of different information in WM. Such an extended model posits that the alpha-gamma phase-amplitude coupling may play a role in suppressing of redundant visual features. To test this I calculated the CFC between the phase of alpha and the broadband gamma frequency (31-150Hz) both recorded from the same sites in the dorsal visual stream. To test if the observed CFC is stronger than chance I generated a reference distribution of the surrogate CFC data by independent shuffling labels of the low- and high-frequency trials.

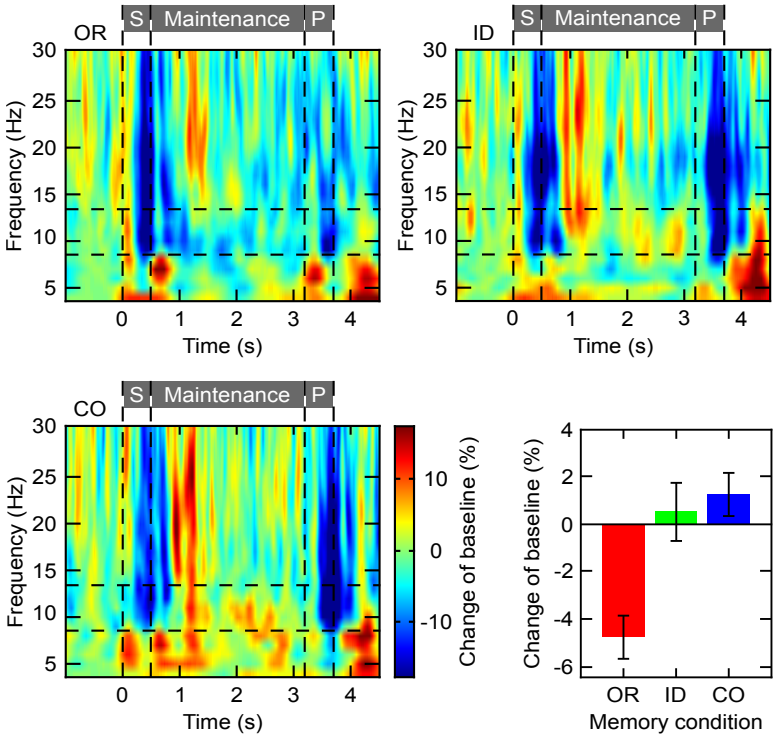

FIGURE 6.11: Time-frequency activation patterns in the dorsal visual stream. Group average of time-frequency activation patterns normalized to the time interval prior to stimulus presentation. The color maps reflect percentage change of baseline separately for each condition (OR, ID, CO reflect maintenance of orientation, identity and control, respectively). The grey bars above the spectrograms indicate trial sequence with the memory stimulus presentation (S), maintenance and the probe (P). The vertical dashed lines indicate time points of these events. The horizontal dashed lines mark the alpha power spectrum between 9 and 13Hz. The bar plot (lower right) presents baseline normalized alpha power averaged during the maintenance period (2.7s). Error bars show SEM.

This procedure keeps intact the spectral properties of the data as well as the actual analytic amplitude and phase time series at each time point while randomizing the trials structure between the two signals. This procedure was performed 100 times. If the observed modulation index is caused by spurious coupling, the empirical value should not be improbable given the reference distribution of surrogates obtained in permutation test. In other words, empirical value might be attributed a rank-based p-value given the reference distribution of surrogate data under the null hypothesis that there is no coupling. This procedure was performed separately for each subject leading to separate significance test for each participant.

The empirical modulation index exceeded the 95^{th} percentile of the surrogates (corresponding to the threshold of $p < 0.05$) for all seven subjects who had electrodes in the dorsal visual stream (see Figure 6.12). The mean modulated frequency was at 90 ± 28.6Hz (mean ± standard deviation). This result shows that the gamma power is indeed modulated by the phase of alpha in the dorsal visual stream. This corroborates the extended multiplexing model of WM as well as the alpha inhibition model of WM. How does the CFC relate to changes in the alpha power will be further investigated. In particular one might expect that the alpha duty cycle increases for task relevant condition (i.e. lower alpha power relates to increased duty cycle in the dorsal stream during the maintenance of orientation).

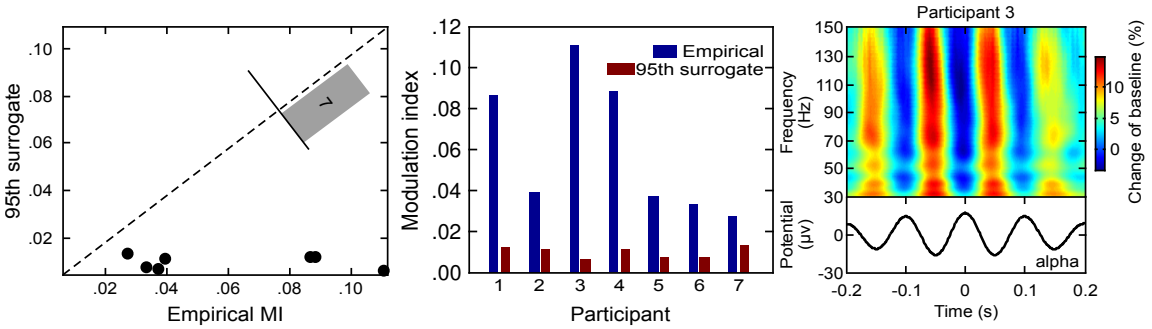

FIGURE 6.12: Intra-regional CFC in the dorsal visual stream. The alpha-gamma coupling in the dorsal visual stream for each of the 7 patients having electrodes in the dorsal visual stream. All seven patients (each represented by single dot) show increased empirical CFC (x-axis) compared to the 95^{th} surrogate (y-axis; left). The alpha-gamma coupling in the dorsal visual stream is reliably stronger than chance which is indicated by all dots being below the diagonal (left) and by all blue bars being above the red bars (middle). Single subject's example of the alpha-gamma phase-amplitude coupling (right panel). Normalized gamma (upper part of the right panel) is non-uniformly distributed across phase of the alpha oscillation (lower part). Time on x-axis is relative to the alpha peak. Color indicates percentage change of baseline. The gamma power is attenuated particularly strong during the peak of alpha oscillations.

6.3.4 Phase synchronization between the prefrontal cortex and the dorsal visual stream.

The gating by inhibition hypothesis [Jensen and Mazaheri, 2010] posits that the functional architecture might be shaped by oscillatory alpha activity. In particular, the model suggests that information is gated by inhibiting task redundant pathways and in turn information is routed to the regions relevant for the task. In other words, networks showing increased inhibition (i.e. stronger alpha power) are expected to show attenuated network connectivity patterns with other brain regions (e.g. prefrontal cortex). To test this prediction I calculated phase synchronization which has been suggested to support network connectivity [Fell and Axmacher, 2011, Fries, 2005, 2015, Liebe et al., 2012]. In particular, I focused on the theta phase synchronization which has been attributed a role in long-range connectivity [Fell and Axmacher, 2011, Liebe et al., 2012]. I calculated theta phase synchronization between electrodes localized in the prefrontal cortex and the dorsal visual stream for all experimental conditions. Indeed, all three subjects who had electrodes in both regions showed increased network connectivity during maintenance of orientation as compared both to the identity and control condition (see Figure 6.13). This result corroborates prediction of the gating by inhibition hypothesis in the dorsal visual stream by showing that theta phase synchronization between this region and the prefrontal cortex is increased while that cortical area undergoes alpha-band disinhibition.

6.3.5 Inter-regional CFC between the phase of PFC and the amplitude of the dorsal visual stream.

Models of working memory (see chapter 2) often assume that the prefrontal cortex plays an executive role over the sensory regions. The neural mechanism by which such a control is executed remains elusive. Here, I tested if the inter-regional phase-amplitude cross-frequency coupling might provide such a mechanism. The alpha-gamma CFC code is a good candidate because it has been suggested that the phase

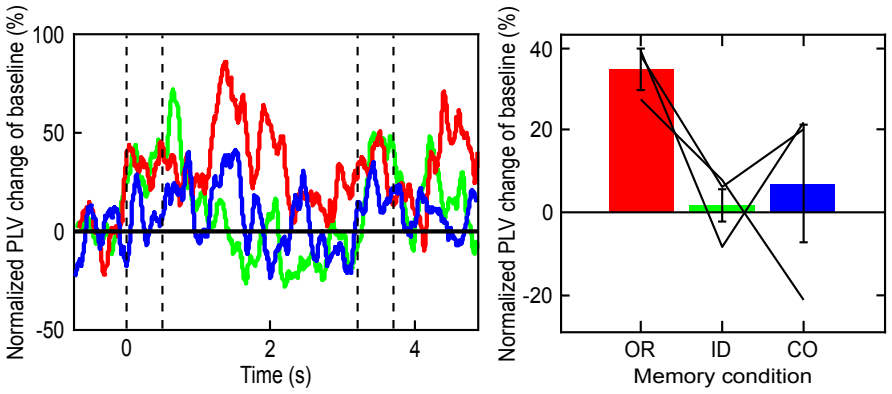

FIGURE 6.13: Phase synchronization between the prefrontal cortex and the dorsal visual stream. The average across all patients having electrodes in both regions (left; N=3). Red, green and blue reflects maintenance of orientation, identity and control, respectively. The vertical dashed lines indicate presentation of the memory stimulus and the probe. The right panel shows phase synchronization in the theta range averaged across the maintenance interval. Bars present average across patients and lines reflect data from single subjects. Error bars show SEM.

of low frequency oscillations (in particular the alpha range) reflect fluctuations in neural excitability [Fell and Axmacher, 2011, Lakatos et al., 2008, 2005]. These changes in neural excitability might then influence the amplitude of the high-frequency activity. Although this has been predominantly suggested as the intra-cortical mechanism, one might imagine that similar fluctuations in excitability are also observed to modulate neural activity across distant brain regions (e.g. between prefrontal cortex and the visual cortices).

To test this I calculated inter-regional phase-amplitude coupling with analytic phase time series in the alpha range obtained from the electrode sites localized in the prefrontal cortex and the amplitude time series derived from the dorsal sites. Here I also used the non-parametric, rank-based statistics with reference distribution obtained by shuffling trial labels and permuting 100 times.

For all subjects with electrodes both in the prefrontal cortex and the dorsal visual stream (N=3) the empirical value exceeded the 95^{th} percentile (corresponding to $p < 0.05$; see Figure 6.14). The mean modulated frequency was at 74 ± 44Hz (mean $\pm$ standard deviation). These results support the idea that the dorsal amplitude is modulated by the phase of prefrontal alpha which may be a mechanism for

top-down control of the dorsal visual stream.

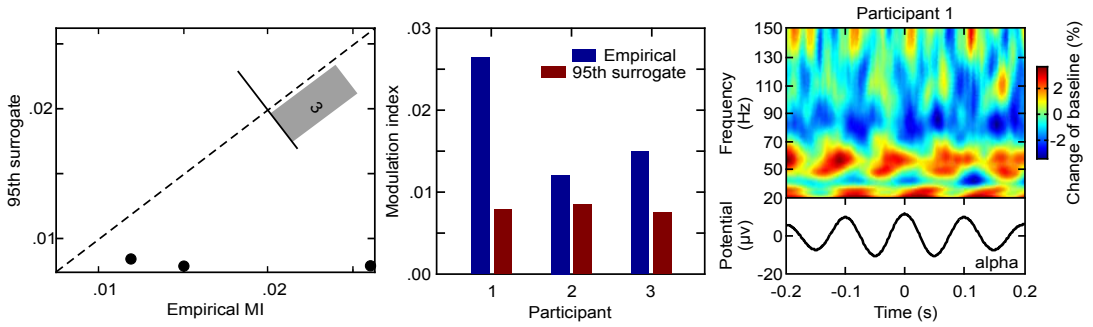

FIGURE 6.14: Inter-regional CFC between the PFC and the dorsal visual stream. Alpha-gamma coupling between the phase time series recorded in the prefrontal cortex and the amplitude time series obtained form the dorsal visual stream for each of the 3 patients having electrodes in both cortical areas. All three patients (each represented by single dot; left) show increased empirical CFC (x-axis) compared to the 95^{th} surrogate (y-axis). The alpha-gamma coupling between the prefrontal and the dorsal visual stream is reliably stronger than chance which is indicated by all dots being below the diagonal (left) and by all blue bars being above the red bars (middle). Single subject's example of the alpha-gamma phase-amplitude coupling (right panel). Normalized dorsal gamma (upper part of the right panel) is non-uniformly distributed across phase of the prefrontal alpha oscillation (lower part). Time on x-axis is relative to the alpha peak. Color indicates percentage change of baseline.

6.4 Maintenance in the ventral visual stream.

6.4.1 Behavioral results.

Accuracy (percentage of correct responses) was measured for three WM conditions. One-way repeated measures ANOVA with memory condition ("orientation", "identity", "control") as a factor revealed a significant main effect ($F(2,22) = 8.44$, $p < 0.005$). Post-hoc paired-sample t-test showed that accuracy in both memory conditions - orientation (90%; $t(11) = 3.52$, $p < 0.005$) and identity (86%; $t(11) = 3.81$, $p < 0.005$) was lower than accuracy in control (96%) condition. The accuracy did not differ between identity and orientation ($t(11) = 1.23$, $p > 0.2$). This result indicates that memory load was indeed higher in both orientation and

identity than in control condition. However, it did not differ between orientation and identity.

6.4.2 Time-frequency activity patterns in the ventral visual stream.

Subsequently, I looked whether activity patterns in the alpha (9-13Hz) frequency range in the ventral visual stream also depends on working memory condition. Since the ventral visual stream regions are relevant for processing of facial information [Allison et al., 1999, Druzgal and D'esposito, 2003, McCarthy et al., 1999, Postle et al., 2003, Ranganath et al., 2004] one might expect that the alpha-band activity would be attenuated for identity compared with remaining conditions (orientation and control). Using the same cluster-based surrogate statistics I indeed observed power decrease during maintenance of identity as compared to orientation and control ($p_{corrected} < 0.05$ each; see Figure 6.15). Importantly, the difference between alpha power during maintenance of orientation and control was not significant. Again, the effect was sustained through the maintenance interval with the peak difference observed during a late parts of the retention interval (1700-2000ms, 1600-2000ms after the stimulus offset for the identity-control and identity-orientation, respectively). These extend the previous results associating the maintenance of identity in WM with decreased alpha-band activity in the ventral visual stream. The results further support the functional-inhibition hypothesis of the alpha-band activity and provide the first evidence from the ventral visual stream supporting the model. The alpha-band activity both in the ventral and dorsal visual stream reflects double dissociation between feature specific networks. Again, to quantify any difference in the gamma frequency range I used repeated measures ANOVA with memory condition (orientation, identity and control) and time interval (1-900ms, 901-1800ms, 1801-2700ms) as factors. The main effect of interval was marginally significant ($F(2,22) = 3.44$, $p = 0.0863$). Neither the main effect of memory ($F(2,22) = 0.91$, $p = 0.48$) nor interaction ($F(4,44) = 1.20$, $p = 0.324$) was reliable. Post-hoc paired-samples t-test showed that broadband gamma

was marginally attenuated during the third interval as compared to first ($t(11) = 2.181$, $p = 0.0517$) and second ($t(11) = 1.982$, $p = 0.0730$). The difference between the first and second time interval was not reliable ($t(11) = 1.520$, $p = 0.1567$). Interestingly, this late gamma effect temporally overlaps with the time windows where the alpha differences were observed.

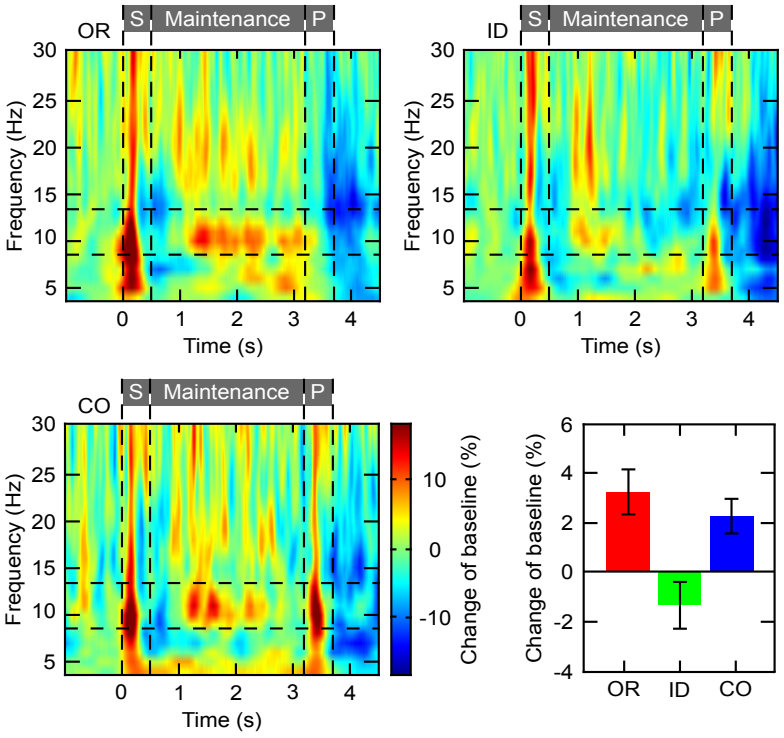

FIGURE 6.15: Time-frequency activation patterns in the ventral visual stream. Group average of time-frequency activation patterns normalized to time interval prior to stimulus presentation. The color maps reflect percentage change of baseline separately for each condition (OR, ID, CO reflects maintenance of orientation, identity and control, respectively). The grey bars above the spectrograms indicate trial sequence with memory stimulus presentation (S), maintenance and the probe (P). The vertical dashed lines indicate time points of these events. The horizontal dashed lines show the limits of alpha power spectrum between 9 and 13Hz. The bar plot (lower right) presents baseline normalized alpha power averaged during the maintenance period (2.7s). Error bars show the SEM.

6.4.3 Intra-regional CFC in the ventral visual stream.

Next, I quantified the intra-regional phase-amplitude CFC in the ventral visual stream and investigated if it is also increased above chance level. Again, I used

phase of the alpha-band and the broadband gamma frequency (31-150Hz) from the same sites in the ventral visual stream. I generated the reference distribution of surrogate data by repeatedly (100 permutations) shuffling labels of low and high frequency trials. Subsequently, I attributed a rank-based p-value to empirical data given this reference distribution. This procedure was performed separately for each subject leading to separate significance test for each participant. The empirical CFC modulation index exceeds the 95^{th} percentile of the surrogate distribution (corresponding to the alpha threshold of $p < 0.05$) for all twelve subjects who had electrodes in the ventral visual stream (see Figure 5.16). The mean modulated frequency was 88 ± 33.6Hz (mean ± standard deviation). This result shows that the gamma power in the ventral visual stream is also modulated by the phase of alpha which corroborates the extended multiplexing model of WM.

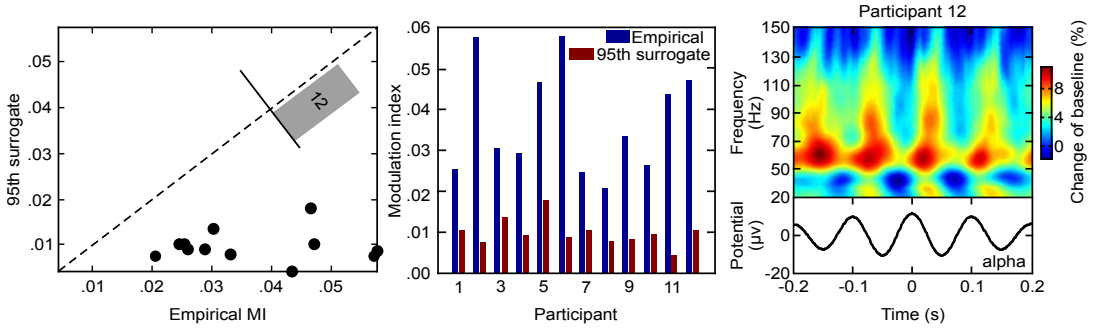

FIGURE 6.16: Intra-regional CFC in the ventral visual stream. Alpha-gamma coupling in the ventral visual stream for each of the 12 patients having electrodes in the ventral visual stream. All twelve patients (each represented by a single dot) show increased empirical CFC (x-axis) compared to the 95th surrogate (y-axis; left). The alpha-gamma coupling in the ventral visual stream is reliably stronger than chance which is indicated by all dots being below the diagonal (left) and by all blue bars being above the red bars (middle). Single subject's example of the alpha-gamma phase-amplitude coupling (right panel). Normalized gamma (upper part of the right panel) is non-uniformly distributed across the phase of alpha oscillation (lower part). Time on x-axis is relative to the alpha peak. Color indicates percentage change of the baseline.

6.4.4 Phase synchronization between the prefrontal cortex and the ventral visual stream.

Next, I tested the gating-by-inhibition hypothesis [Jensen and Mazaheri, 2010] in the ventral visual stream. I investigated whether the network connectivity between the prefrontal cortex and the ventral visual stream is enhanced during the condition which shows a decreased alpha-band power within the ventral stream (i.e. the identity condition). Again, I calculated phase synchronization in the theta range as a proxy for network connectivity. Indeed, five out of six subjects who had electrodes in both regions show increased synchrony during maintenance of identity (see Figure 6.17) compared to the remaining conditions. This result further corroborates predictions of the gating-by-inhibition model in the ventral visual stream by showing that theta phase synchronization between this region and the prefrontal cortex is increased specifically while that area undergoes alpha-band disinhibition.

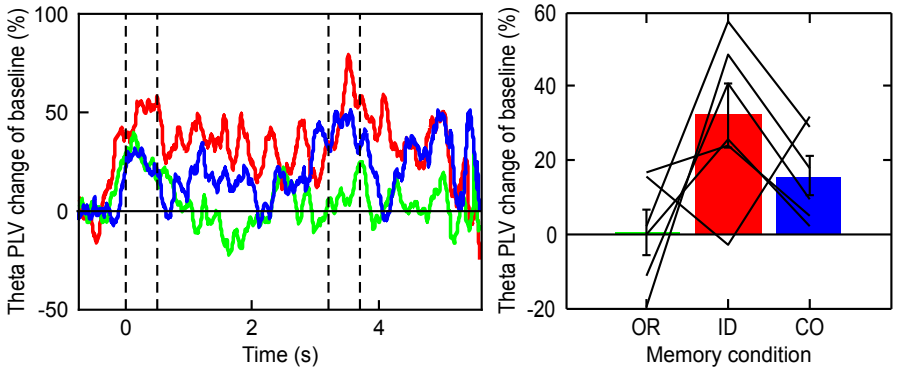

FIGURE 6.17: Phase synchronization between the prefrontal cortex and the ventral visual stream. The average across all patients having electrodes in both regions (left; N=6). Red, green and blue reflect maintenance of orientation, identity and control, respectively. The vertical dashed lines indicate presentation of the memory stimulus and the probe. The right panel shows phase synchronization in the theta range averaged across the maintenance interval. Bars present average across patients and lines reflect single subject. Error bars show the SEM.

6.4.5 Inter-regional CFC between the phase of PFC and the amplitude of the ventral visual stream.

Finally, I tested if the inter-regional CFC between the prefrontal alpha-band phase and the ventral gamma amplitude is also increased above chance. I used the same non-parametric, rank-based statistics with reference distribution obtained by shuffling trial labels and permuting 100 times. For all subjects with electrodes both in the prefrontal cortex and the ventral visual stream (N = 6) the empirical value exceeded the 95^{th} percentile (corresponding to $p < 0.05$; see Figure 6.18). The mean modulated frequency was 70 ± 11.8Hz (mean ± standard deviation). These results support the idea that the ventral gamma is modulated by the phase of prefrontal alpha.

Together, these results suggest that neural oscillations in the alpha band are modulated in feature specific networks in the ventral and dorsal stream during working memory maintenance. These findings suggest further that the decreased alpha power contribute to WM by selective disinhibition of areas that represent currently relevant features (in the ventral and the dorsal visual stream). Importantly, the current results extend the functional inhibition role of alpha into the ventral visual stream. Furthermore, the gamma activity in these regions is modulated be the phase of alpha resulting in the alpha-gamma phase amplitude coupling. Finally, the strength of functional network connectivity follows the changes of alpha power. The strength of connectivity was increased between the prefrontal cortex and a brain network showing alpha band disinhibition (dorsal and ventral during maintenance of orientation and identity, respectively). This result supports the gating-by-inhibition model by showing that the local increase in alpha band oscillations may block task redundant pathways and therefore disconnect these cortical regions.

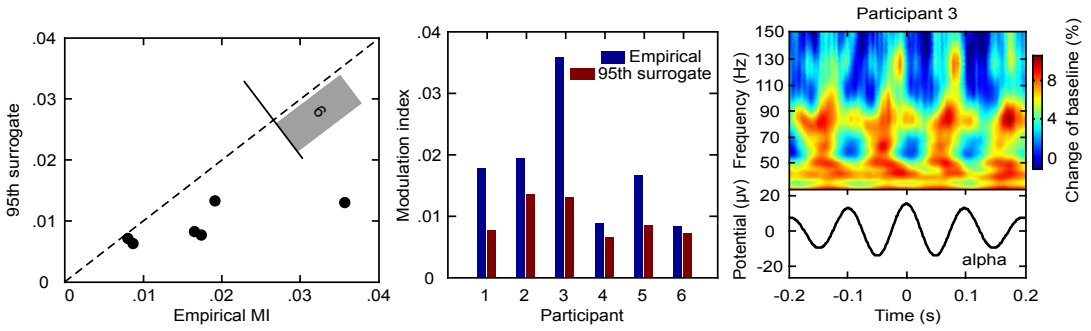

FIGURE 6.18: Inter-regional CFC between the PFC and the ventral visual stream. All six patients (each represented by a single dot) show increased empirical CFC (x-axis) compared to the 95^{th} surrogate (y-axis; left). The alpha-gamma coupling between the frontal and the ventral visual stream is reliably stronger than chance which is indicated by all dots being below the diagonal (left) and by all blue bars being above the red bars (middle). Single subject's example of the alpha-gamma phase-amplitude coupling (right panel). Normalized ventral gamma (upper part of the right panel) is non-uniformly distributed across phase of the frontal alpha oscillation (lower part). Time on x-axis is relative to the alpha peak. Color indicates percentage change of the baseline.

6.5 In search for the WM code in the vlPFC.

In the third part of the thesis I explored LFP data recordings from the ventrolateral prefrontal cortex (vlPFC) of two macaque monkeys performing multi- and single-item WM. I investigated the representational properties coded in the potentials of the vlPFC. First, I used two methods to search for any load-dependent power changes in the time-frequency spectrum. To this end, I used linear regression and multidimensional pattern classification to search for load-dependent effects during the maintenance interval. Since the PFC has been also suggested to represent abstract features I tested if such an abstract representation may be decoded from the local field potential recordings of the vlPFC. To this end, I used the response mapping (left button press for match and right for non-match response) which may be considered as such an abstract representation. Finally, I tested one other important prediction of the alpha-inhibition model. In particular, the duration of a duty cycle should relate to the release from neural inhibition [Jensen et al., 2014, Spaak et al., 2012]. Consequently, the length of gamma burst should be longer durating larger duty cycle. This effect has never been measured outside of the

primary visual area.

6.5.1 Behavioral results.

First, I calculated the accuracy (percentage of correct trials) as a function of memory load. Note, that both monkeys were trained to perform load 1 version of the task. The current experiment is the first time when both experienced loads 2 and 3. To quantify the effect of memory load on performance I calculated independent samples one-way ANOVA separately for each of the monkeys followed by post-hoc independent samples t-tests. I observed that the accuracy decreased across memory loads ($F(2,819) = 3.51$, $p < 0.05$). Performance in load 1 was higher than in load 2 (77.86% vs. 69.34%; $t(552) = 2.2808$, $p < 0.05$). Accuracy in load 2 was higher compared to load 3 (69.34% vs. 59.17%; $t(561) = 2.5252$, $p < 0.01$; both for monkey L). There was no significant difference in the RT across memory loads (456ms; 484ms; 498ms; $F(2,510) = 0.39$, $p = 0.6761$). The performance of the second monkey also decreased across memory loads ($F(2,945) = 6.66$, $p < 0.001$). Post-hoc independent samples t-tests revealed that performance in load 1 was higher than in load 2 (86.08% vs. 76.06%; $t(644) = 3.2626$, $p < 0.001$) and accuracy in load 2 was higher compared to load 3 (76.06% vs. 68.48%; $t(658) = 2.1784$, $p < 0.05$; both for monkey J). There was also a difference in reaction time observed in behavior of this monkey ($F(2,675) = 5.28$, $p < 0.01$). In details, the RT in memory load 1 was shorter than in load 2 (501ms vs. 531ms; $t(521) = 2.6447$, $p < 0.01$). There was no difference between loads 2 and 3 (531ms vs. 539ms; $t(475) = 0.687$, $p = 0.492$). The lack of consistent difference in RT is likely due to the so called the floor effect. The overall response times are already very fast which is likely due to extensive training. Therefore, a further modulation of RT by the current task is unlikely to be observed (see Figure 6.19).

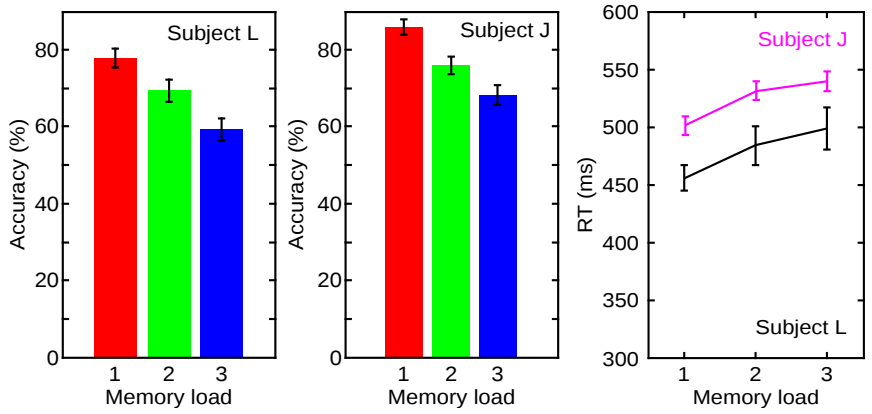

FIGURE 6.19: Behavioral results (Experiment 4). Accuracy scores for monkey L (left) and J (middle) as a function of memory load. The right panel presents reaction times across memory loads for monkey J and L (magenta, black, respectively). Error bars show the SEM.

6.5.2 Time-frequency patterns of activity.

I performed time-frequency analysis and searched for any linear changes in power spectrum (4:150Hz) during maintenance interval (see Figure 5.20). I calculated t-statistics for regression coefficient regressing trial specific spectro-temporal data on predictor variable which was memory load. I applied non-parametric surrogate statistics with cluster correction for multiple comparisons [Maris and Oostenveld, 2007]. No significant cluster was observed (all cluster candidates $p > 0.2$). There was also no difference when I looked at specific frequency bands (i.e. when analysis was limited to the alpha frequency band) where I expected to see the load-dependent difference to be the strongest. Several possibilities explaining this null effect are described in the discussion.

6.5.3 Decoding of memory load conditions.

The approach used in the previous paragraph which quantifies an overall signal intensity change as a function of memory load although robust in many fields, has recently been criticized for its limited sensitivity and specificity [Postle, 2015]. An alternative method which benefits from the complexity of the considered data has been proposed and successfully applied to the fMRI [for review see Postle, 2015].

In order to gain sensitivity I used similar multivariate approach to the current dataset. In particular, I investigated if such a method indeed finds any evidence for load-dependent changes in the time-frequency spectrum. To this end, I trained a support vector machine classifiers on the LFP power data across frequency bands ranging from 4Hz to 200Hz. I trained and tested 46 separate classifiers on consecutive data bins during the whole encoding and maintenance epoch. Note, that each classifier was trained and tested on the data from the same time bin. I performed 7-folds cross-validation decoding with three classes (load 1, 2, 3). The chance performance was at 33%.

Indeed, the decoding approach revealed time intervals where classification was significant (above chance and above the 95^{th} percentile of the surrogate distribution; see Figure 6.20B). Importantly, the significant intervals were mostly limited to the interval of stimulus presentation and shortly after. During most of the 3s maintenance period classifiers were not able to decode different loads above the 95^{th} percentile of the surrogate data. The accuracy of classification was at 33%. The results were comparable for both monkeys. These results suggest that although there are load-dependent power effects observed in the vlPFC but so far I have been able to observe them mostly during the encoding part of the task and not during the maintenance.

6.5.4 Decoding of an abstract rule in the vlPFC (match vs. non-match).

More recent models describing a role of the prefrontal cortex in WM suggest that rather than individual items the PFC represents abstract rules. In the context of WM a common form of such an abstract representation is a response mapping. In the current experiment monkeys were responding with left button press if the probed stimulus was presented in memory set (match trial) and with the right button press if the probe was not presented in the memory set (non-match). I used the support vector machine (7-folds, two-class classification) to train a classifier in decoding which of these two categories a trial belongs to. Indeed, I observed that

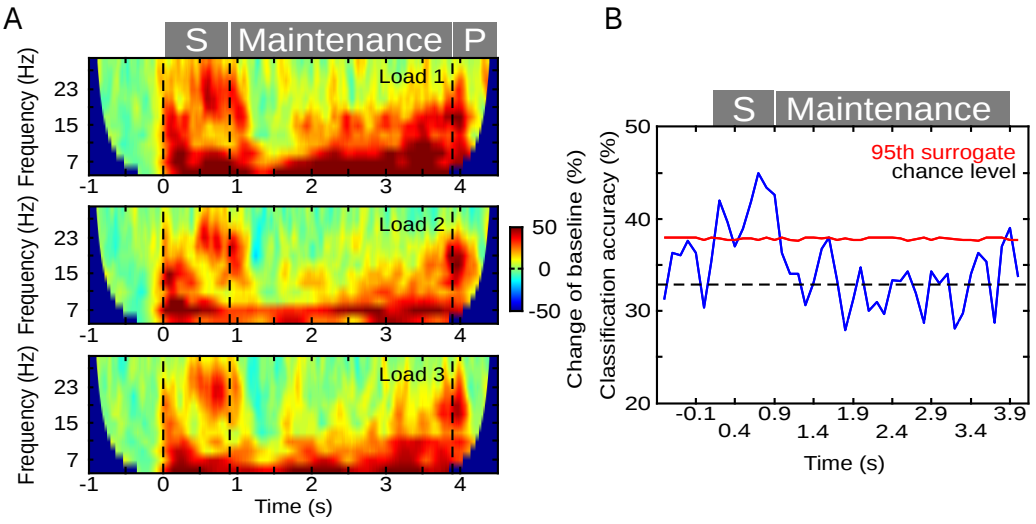

FIGURE 6.20: In search for the load-dependent effects in the vlPFC. (A) Time-frequency patterns of activity presented separately for each memory load (monkey L). The grey bar above spectrograms indicates trial structure (S – memory stimulus, P – memory probe). The vertical dashed lines indicate time points of these events. Color indicates percentage change of baseline. (B) The result of three-class multivariate pattern classification analysis. Blue trace presents accuracy of consecutive classifiers trained and tested on the respective time bins. The black dashed line indicates the level of classification accuracy at chance (33%). The red line presents significance threshold (corresponding to p=0.05) being the results of the 95^{th} percentile of the surrogate distribution. The time axis is relative to the memory stimulus onset.

a classifier was able to learn the trial category and generalize it to new trials with high accuracy (76.0% and 69.9% for L and J, respectively). This results are above chance level (50%) and significantly differ from the surrogate distribution ($z = 4.7$, $z = 3.7$, for monkey L and J, respectively; see Figure 6.21). These results support the concept of the vlPFC involvement in representing abstract rules. Previous studies with single unit recordings have shown that such information is indeed coded by neurons in the PFC [Meyers et al., 2012]. The current results provide evidence that such information may be also decoded from the LFP signal in the vlPFC.

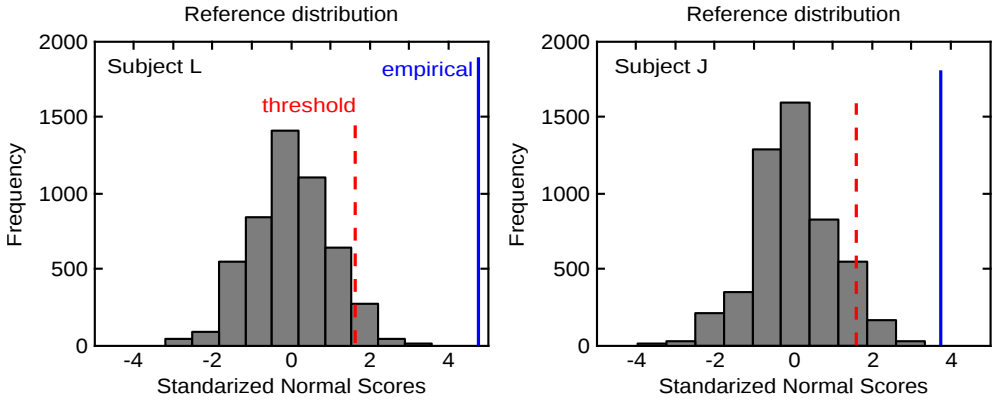

FIGURE 6.21: Classification of the response mapping in the vlPFC. The result of two-class multivariate pattern classification analysis. The histogram presents the standardized normal scores (z-scores) of the reference surrogate distribution under the null hypothesis that classification was not different from chance. The reference distribution was obtained by shuffling repeatedly (N=5000) the class labels of the test data. The red dashed line indicates significance threshold (corresponding to p = 0.05) being the 95^{th} percentile of this surrogate distribution. The blue line presents empirical z-scored result of classification accuracy on the test data. The results for monkey L (left panel) and J (right panel) are presented.

6.5.5 The length of gamma burst depends on the duty cycle in the vlPFC.

Finally, I intended to test yet another prediction from the alpha-inhibition model which posits that the size of a duty cycle constraints the neural activity [Jensen et al., 2014, Spaak et al., 2012]. The duty cycle is the phase of alpha oscillation during which the neural activity (e.g. gamma bursts) is released from inhibition. The increasing window of the duty cycle (as indexed with the alpha power) results in longer disinhibition period during which neural activity may be observed. Here, I tested if the increasing duration of the duty cycle relates to the length of the gamma burst. In details, I expected to see that the increased duration of the duty cycle is associated with prolonged gamma bursts. To this end, I filtered LFPs in the alpha-band, during the whole 3s maintenance interval, sorted alpha peaks depending on duration of the duty cycle (following Spaak et al. [2012] the magnitude of power was used as an index for duration of the duty cycle) and looked at the length of gamma burst separately for short and long duty cycles.

Both monkeys showed consistently increased length of the gamma burst for long duty cycles (gamma during long vs. short duty cycle; 37ms vs. 35ms; $t(1998) = 3.557$, $p < 0.0005$ monkey L and 35ms vs. 32ms; $t(1998) = 3.181$, $p < 0.001$ monkey J). This finding is very interesting because the effect of duration of the duty cycle on the length of gamma burst has been so far only observed in the primary sensory cortex [Spaak et al., 2012]. The current result extends the relation between the alpha duty cycle and the length of gamma burst beyond primary visual cortex.

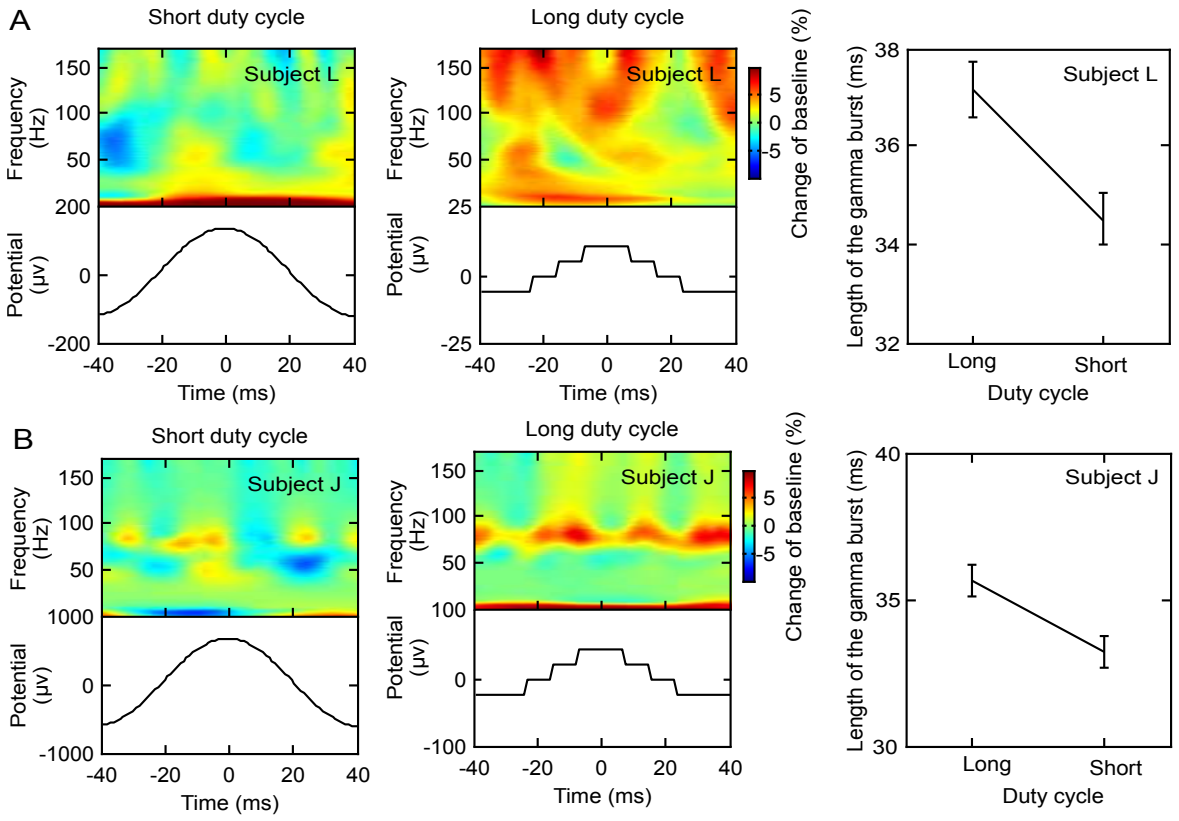

FIGURE 6.22: Alpha duty cycle modulates the length of gamma burst in the vlPFC. (A, B) Normalized gamma activity (upper part) plotted across the alpha cycle (lower part) during short (left panel) and long (middle panel) duty cycle separately for monkey L (A) and J (B). The color shows percentage change from the baseline. Time is plotted relative to the alpha peak. The line plot (right panel) indicates the duration of gamma burst during long and short duty cycle. Error bars show the SEM.

Chapter 7

Discussion

The findings discussed here support the concept of WM as a process involving multiple cortical areas including the hippocampus, the dorsal, ventral visual stream and the prefrontal cortex. Each of these brain regions contribute differently to WM. The sensory cortices are likely involved in the storage of information. For example, the dorsal visual stream possibly participates in the maintenance of orientation. The ventral visual stream contributes to maintenance of facial identities. Importantly, the hippocampus may represent sequences of items and contribute to the maintenance of some abstract information like relations among the multiple items. In addition, the prefrontal cortex contributes to WM by representing more abstract features like stimulus contingencies or response mappings. It also selectively connects with various feature specific networks likely reflecting a top-down control over these cortices. The current findings support the concept of WM as relying on a distributed network with various nodes (i.e. the hippocampus, ventral, dorsal visual stream, frontal cortex) of this system operating coherently to support maintenance of different contents. The current findings further support the notion of working memory as a dynamic process which depends on local oscillations and long-range coherence as well as a hierarchy of oscillations. I will first discuss the role of the hippocampus in WM maintenance. Subsequently, I would like to describe the contribution of the visual regions to working memory followed by the role of prefrontal cortex.

7.1 Working memory maintenance in the hippocampus.

In the first part of my thesis I investigated patterns of the hippocampal activity during two working memory tasks. In particular, I analysed the iEEG recordings obtained from the human hippocampus to test predictions from the multiplexing model of WM [Lisman and Idiart, 1995, Lisman and Jensen, 2013] and the alpha-inhibition hypothesis [Jensen et al., 2014]. The current results presented chapters 6.1 and 6.2 show that the maintenance of multiple items in WM relies critically on a hierarchy of oscillations in the human hippocampus. More specifically, successful maintenance of multiple items in WM depends on periodic fluctuations between two different oscillatory processes. Periods of *memory activation* were indicated by memory load-dependent reductions in alpha power. Importantly, these periods showed also lower levels of the CFC strength. Furthermore, a relevant CFC parameter (load-dependent changes of the peak modulated frequency) correlated with individual WM capacity. Additionally, the *memory activation* periods were interleaved with intervals showing memory independent high alpha power and increased CFC strength. The two types of intervals – *memory activation* and load independent – were organized by the phase of an endogenous delta oscillation forming a hierarchy of oscillations (an oscillatory phenomenon previously observed in the auditory cortex [Lakatos et al., 2008, 2005]. These results corroborate predictions from both the alpha-inhibition hypothesis [Jensen et al., 2012, Klimesch, 2012] and the multiplexing model of WM [Lisman and Idiart, 1995, Lisman and Jensen, 2013].

On a more general level these results support the hypothesis that hippocampal involvement is important for working memory maintenance. Furthermore, strong alpha power observed in load 1 indicates that the hippocampus is inhibited during maintenance of a single item. Only during maintenance of multiple items the alpha power decreases. This suggest that the hippocampus is particularly relevant for maintenance of multiple items in WM. Together with studies showing impaired multi-item WM performance with hippocampal lesions [Aggleton et al.,

1992, Nichols et al., 2006, Olson et al., 2006] the current data suggest that the hippocampus is relevant for relational processing [Eichenbaum, 2000] or binding of information [Davachi, 2006, Davachi and DuBrow, 2015] at retention intervals as short as 3 seconds.

Another important observation is that the hippocampus was similarly involved both in WM maintenance of novel (i.e. trial unique faces) and highly familiar content (i.e. digits). This is important because it has been previously suggested that maintenance of novel information may indeed involve the hippocampus but maintenance of familiar content likely relies on the neocortical structures [Hasselmo and Stern, 2006]. The current results suggest however, that independent of the type of memory material the hippocampus contributes similarly to WM maintenance. The amount of information (i.e. multiple items) rather than the type of information itself is what seems to determine the involvement of the hippocampus in WM.

7.1.1 The alpha inhibition model.

The alpha-inhibition model posits that neural activity is inhibited by the alpha oscillations in networks that are functionally irrelevant for a given task [Jensen et al., 2012, Klimesch, 2012]. This is important because it suggests that the alpha power is a marker of neural inhibition. Furthermore, it offers the number of testable predictions. This account predicts for example that increased involvement of a brain network should be associated with attenuation of the alpha rhythm. Abundant evidence supports this functional inhibition mechanism [Bonnefond and Jensen, 2012, Haegens et al., 2011, Jensen et al., 2012, Jokisch and Jensen, 2007, Klimesch, 2012, Sauseng et al., 2009, Spaak et al., 2012, Thut et al., 2006]. The majority of these results however describe effects observed in the sensory cortices (i.e. the parieto-occipital regions or primary sensory cortices). The results reported here show that the functional inhibition property of the alpha oscillations generalizes to the human hippocampus. In particular, the low-frequency activity decreased

linearly with WM load during the maintenance interval. This decrease was organized into periods that were equally spaced across the whole retention interval and were interleaved with periods showing no load-dependent disinhibition. Importantly, the decrease of power was absent during incorrect trials. Since there is a negative correlation observed between alpha power and either the action potential rates [Haegens et al., 2011] or high-frequency activity [Spaak et al., 2012], the results showing increased alpha power have been interpreted as reflecting active functional inhibition. In turn, release from alpha-related inhibition is needed to process relevant information [Jensen et al., 2012, Klimesch, 2012, Klimesch et al., 2007]. This suggests that alpha activity in the human hippocampus might play a similar inhibitory role like it does in the parietal and occipital regions. Furthermore, hippocampal alpha disinhibition during maintenance of multiple items is critical for successful performance.

Since no prior data on the alpha-related inhibition in the human hippocampus was available, the current study did not formulate a priori assumptions regarding the precise frequency where the effect of inhibition might be observed. To this end, I considered a broader frequency spectrum ranging from theta to alpha and beta (<30Hz) with a conservative correction for multiple comparisons (see Methods). The reason for implementing such a broad window was based on previous findings that both theta [Fell et al., 2011, Huang et al., 2014, Jensen et al., 2015, Mehta et al., 2002, van Kerkoerle et al., 2014] and beta [Bauer et al., 2014, Jensen et al., 2015, van Kerkoerle et al., 2014, Waldhauser et al., 2012] oscillations may show similar functional effects to alpha oscillations. Indeed, load-dependent decrease was observed to spread broader across frequency range in both datasets (7-23Hz, 4-20Hz in the Experiment 1 and 2, respectively). Therefore, the current results suggest that the hippocampus may show functionally similar oscillations across a broader frequency range.

7.1.2 The multiplexing model of WM.

The results discussed here corroborate several predictions from the multiplexing model of WM [Lisman and Idiart, 1995, Lisman and Jensen, 2013]. First, the model posits that non-uniform distribution of high-frequency activity across the phase of low-frequency oscillations (which results in the CFC) is a mechanism of WM maintenance. Indeed, the present results show increased CFC between the phase of low-frequency oscillations and the power of high-frequency activity during maintenance of multiple items which extends previous empirical findings [Axmacher et al., 2010b, Sauseng et al., 2009, Siegel et al., 2009]. This is also consistent with the literature on the role of the hippocampus in multi-item WM [Aggleton et al., 1992, Axmacher et al., 2010b, Nichols et al., 2006, van Vugt et al., 2010]. Second, the model predicts a linear relation between parameters of the CFC (the CFC ratio as well as the peak modulating low-frequency oscillations and peak modulated high-frequency activity) and individual WM capacity limit. Such correlations were observed in both datasets (Experiments 1 and 2). Load-dependent increase in the peak modulated high-frequency activity results in higher individual WM capacity limit. Importantly, this effect was independent of the remebered stimuli materials (i.e. similar correlation was observed for both digits and faces). An interesting finding which was not predicted by the model, is that the CFC strength varies across the maintenance interval. In particular, the CFC strength is reduced during the *memory activation* periods as compared to the interleaving periods. This result may appear surprising because the CFC has been related to maintenance information in WM [Lisman and Idiart, 1995, Lisman and Jensen, 2013]. However, weaker CFC strength reflects broader distribution of modulated high-frequency activity across the phase of low-frequency oscillation. Hence, the current finding of decreased CFC during load-dependent alpha disinhibition is compatible with the concept of increased involvement of the hippocampus during maintenance of longer sequences. This conclusion is further supported by increased CFC strength during incorrect as compared with correct trials.

7.1.3 Rhythmic working memory fluctuations.

Two different oscillatory modes which were observed as periodic fluctuations of both the power of low-frequency and the CFC strength were organized by an endogenous delta oscillation. In particular, the *memory activation* periods were locked to a very specific delta phase resulting in a hierarchy of oscillations [Lakatos et al., 2008, 2005]. There is a fascinating yet still speculative convergence to rodent literature. These two oscillatory states might reflect a similar process of switching between two modes as observed in the firing patterns of the entorhinal grid cells [De Almeida et al., 2012] and the CA1 local field potential [Colgin and Moser, 2009]. The entorhinal grid cells have been observed to operate in two distinct modes. (1) In predictive mode they represent a location ahead of animal and (2) in short-term memory mode they represent a location just passed by the animal. Interestingly, these two modes are structured by the theta and gamma oscillations [De Almeida et al., 2012]. It is worth to note that the human delta oscillations have been suggested to reflect the rodent theta rhythm [Jacobs, 2014, Lega et al., 2012, Watrous et al., 2013]. Thus, the current results might be a homologue mechanism of periodic switching between distinct modes in the human hippocampus. Colgin and Moser [2009] observed that the gamma in the CA1 similarly alternate between two different modes. Fast (65-140Hz) and slow (25-50Hz) gamma oscillations in the CA1 synchronize with fast and slow gamma oscillations in the medial entorhinal cortex (MEC) and the CA3, respectively. Therefore, information routing between the CA1 and the MEC/CA3 periodically fluctuates with the theta phase likely providing a mechanism for temporal separation of information routed from distinct sources. The current results of two processing modes organized by the delta phase might also reflect a homologue mechanism of such a temporal separation of potentially interfering information in the human hippocampus (i.e. currently maintained information which might be a homologue of the short-term memory mode and the incoming information – a homologue of predictive mode).

7.1.4 Periodic gating by a hierarchy of oscillations in the hippocampus.

Oscillations in low frequencies have been suggested as a gating mechanism during working memory [Raghavachari et al., 2001, 2006]. Gating refers to continuous sustained and stationary change in the signal intensity which starts at the beginning of a working memory task and continuous until the end of the task. It has been suggested as a potential mechanism for WM storage (see chapter 1). The current results show that the activity in the hippocampus is being gated during working memory by a periodic rather than a stationary process. In details it is gated by a hierarchy of oscillations which periodically fluctuates with the phase of delta rhythm. I can only speculate as to what may be the benefit of such a rhythmic mode. One possibility is that such periodicity helps to avoid interference between memory activation and some other process like encoding of potentially upcoming information. Indeed, switching between maintenance of already encoded content and encoding of new potentially relevant information is an important process for successful performance in a WM task [Kessler and Oberauer, 2014]. The storage of information (presumably related to the current *memory activation* intervals) might be temporally separated with an encoding mode needed for "updating" of a memory set. The representation of a memorized sequence may benefit the general strengthening during release from inhibition intervals. The same representation may be updated for potentially new upcoming information during intervals with increased hippocampal alpha-inhibition. One may further speculate that during these encoding or updating intervals information transfer should be increased with other brain regions (like relevant sensory cortices or prefrontal cortex). Such an interpretation has been suggested in other domains of cognitive research for example, a similar mechanism of periodic switching between strengthening and inhibition of neural responses have been observed to optimize the use of limited capacities in visual [Busch and VanRullen, 2010] and auditory [Morillon et al., 2014] perception. The observation of periodic switching between two functional modes in the human hippocampus challenges the existing models of WM. Although the new concepts

in working memory do assume that memory representation dynamically changes over time [see chapter 1 and 2 Stokes, 2015] none of them account for the possibility that WM representation is gated by a periodic hierarchy of oscillations. It is important to note that these conclusions are limited to the human hippocampus. In particular, it does not exclude the possibility that there are other brain areas showing sustained and continuous rather than periodic processing mode. In fact, the current results from the ventral and dorsal stream suggest that indeed there are some other brain areas which do not show such a periodic mode. Therefore, it is very likely that the current periodicity is specific to the hippocampus.

7.1.5 Correlations with behavior.

One very prominent prediction from the multiplexing model of WM is that individual limit in WM capacity is constrained by parameters of the CFC. In line with this prediction, the current results show that the central measure of CFC (change in the peak modulated frequency) inter-individually correlates with WM capacity limits. In particular, patients with higher load-dependent increases of the peak frequency-for-amplitude had higher WM capacity limits and were able to memorize longer sequences. Note that these results were subsequently replicated in an independent group of patients performing another WM task. The same factor (change in the peak modulated frequency) limited the individual WM in both data sets. Since, the frequency-for-phase was constant across memory loads the increasing peak frequency-for-amplitude for participants with higher WM capacity may be associated with more distinct representations at increasing memory load. This was further examined looking at the m:n phase-phase coupling. This measure reflects how well consecutive cycles of the high-frequency modulated activity are locked to distinct phases of the low-frequency oscillations. Again, inter-individual positive correlation was observed. Higher WM capacity limits were observed in participants with higher coupling factors [for more detailed description see Chaieb et al., 2015]. These suggest that there are two important factors correlating with individual WM capacity limits. First, the load-dependent change in the peak

frequency-for-amplitude. In particular, participants with higher peak frequencies are capable maintaining longer sequences in memory. Second, precise locking of these increasing in frequency cycles is also important. In other words it is not only the increase in frequency which matters for high WM capacity but also precise locking of the gamma cycles to the consecutive theta phases. One interesting follow-up question is whether this mechanism is specific to the hippocampus or whether it generalizes to other brain networks that are involved in visual working memory such as prefrontal, parietal, or occipital cortices.

7.2 Working memory in the visual brain.

In the second part of my thesis I investigated predictions from the alpha-inhibition [Jensen et al., 2014] and the gating-by-inhibition [Jensen and Mazaheri, 2010] models in the context of working memory. To this end, I used the iEEG recordings from the ventral and the dorsal visual stream, as well as from the prefrontal cortex of patients performing a single item WM task. This task was designed to selectively engage either the dorsal or ventral visual stream during maintenance of distinct visual features. Based on previous studies, I anticipated alpha power in the dorsal stream to be attenuated in a task of a spatial nature [i.e. maintenance of the orientation Harrison and Tong, 2009, Jokisch and Jensen, 2007, Pasternak and Greenlee, 2005]. In contrast, I hypothesized that the alpha power in the ventral visual stream would decreased in the identity task [Allison et al., 1999, Druzgal and D'esposito, 2003, Jokisch and Jensen, 2007, McCarthy et al., 1999, Postle et al., 2003, Ranganath et al., 2004]. Indeed, I observed that the power of alpha oscillations double dissociates between feature specific networks in the dorsal and ventral visual streams. In these brain regions gamma activity was locked to the phase of alpha. Furthermore, I found that the phase synchronization (a proxy for network connectivity) also shows a double dissociation between the prefrontal cortex and the feature specific networks. Changes in phase synchronization matched these from the alpha power. The prefrontal cortex was selectively connecting with networks in the ventral and dorsal visual stream depending on the

task and the amount of alpha power in a given brain region. Additionally, gamma activity in the ventral and the dorsal visual stream was organized inter-regionally by the phase of alpha oscillations from the prefrontal cortex. This might reflect a top-down signal originating from the prefrontal cortex which modulates local networks in the visual brain.

7.2.1 The alpha activity double dissociates between feature specific networks in the dorsal and ventral visual stream.

The results presented here show a decreased alpha power in the dorsal visual stream during maintenance of orientation as compared to identity and control condition. In contrast, the alpha power in the ventral visual stream was attenuated while participants were maintaining the identity of faces as compared to orientation and control. The current findings replicates previous results obtained with MEG [Jokisch and Jensen, 2007] and extends the functional role of alpha into the ventral visual stream. There are several important features making these results a particularly strong support for the alpha-inhibition model.

First, during all three conditions the set of stimuli presented to participants was exactly same. Conditions differed only with respect to the instructions which asked to maintain distinct features (orientation, identity or passively view a face). Therefore, participants looking at the same images of faces needed to maintain distinct feature properties of these stimuli. This is important as the low level visual features cannot explain the observed difference. It further suggests that participant's goal (i.e. to maintain a certain feature) might modulate oscillations in the dorsal and ventral visual stream.

The overall difficulty of the task also cannot explain the current results (no difference in accuracy across the two memory conditions).

Finally, the results from the dorsal and ventral visual stream together with the hippocampal alpha results described above suggest that the functional inhibition property of the alpha oscillations generalize across multiple brain regions [see also

Haegens et al., 2015]. This is important because previous results focused mostly on the visuospatial attention and working memory in the primary sensory cortex [Bonnefond and Jensen, 2012, Jensen et al., 2012, Klimesch, 2012, Sauseng et al., 2009, Spaak et al., 2012, Thut et al., 2006]. Therefore, these findings significantly extend the alpha-inhibition model.

7.2.2 Intraregional cross-frequency coupling: the multiplexing model of WM in the neocortex.

The multiplexing model of WM [Lisman and Idiart, 1995, Lisman and Jensen, 2013] was initially proposed as a neural code for representing sequences of multiple items in working memory. Roux and Uhlhaas [2014] proposed that the theta-gamma and alpha-gamma phase-amplitude code might represent distinct information. In details, the authors suggested that the alpha-gamma coupling may underlie neocortical mechanism for inhibition of task irrelevant activity [Roux and Uhlhaas, 2014]. Such an extended multiplexing model is supported by the results reported here which shows an increased CFC between the phase of alpha and the power of gamma both in the ventral and dorsal visual stream. In both of these brain regions the high frequency activity was modulated by the phase of local alpha oscillations. The model would further predict that the strength of the CFC will differ depending on the relevance of memorized material. However, in the current data I observed no no obvious modulation of the CFC strength by the cognitive task. One possible reason for the null effect might be a limited sensitivity of the available method. Although it is difficult to completely exclude such an explanation, I conducted an additional analyses by calculating the CFC with an alternative method [Tort et al., 2010]. The results were qualitatively similar with no modulation across memory conditions. This does not exclude the possibility that both methods lack sensitivity. Another possibility is also that the CFC increases in a feature unspecific manner just by the fact of visual stimulation and reflects a more generic property of the neural oscillations. One other possibility is that the effect of feature specific inhibition might be measured with other paramters of the

modulated signal.

I observed no clear indication of the univariate gamma power change across conditions as previously observed by Jokisch and Jensen [2007]. The gamma power effects usually originate from synchronized neural activity in relatively small and local networks. Although the iEEG signal is spatially comparable to the one obtained with MEG [Buzsáki et al., 2012], the spatial coverage in the dorsal visual stream used in the current study is much sparser compared to a typical MEG montage. Thus, the current study may lack the spatial sensitivity over the dorsal visual stream to replicate the univariate gamma power change across conditions.

7.2.3 Gating of feature specific activity in the dorsal and ventral visual stream.

In this part of my thesis, I tested predictions derived from the gating-by-inhibition hypothesis. Jensen and Mazaheri [2010] suggested that local oscillations in the alpha range may be a mechanism of network connectivity which coud operate by blocking task irrelevant pathways. For example, the model predicts that information will be rather routed to these networks which show decreased neural inhibition (i.e. lower levels of alpha power). Indeed, I observed that the strength of phase synchronization between the feature specific networks and the prefrontal cortex varied depending on the level of alpha inhibition and memory condition. The phase coherence was higher between the prefrontal cortex and the dorsal visual stream during maintenance of orientation as compared to both identity and control. Importantly, these changes in condition specific communication patterns followed similar changes in the alpha power. Again, these connectivity patterns double dissociate between the ventral and the dorsal visual stream. These results corroborate predictions from the gating-by-inhibition model suggesting that information is likely to be gated to these networks which show alpha-band disinhibition.

To further explore prefrontal interactions with feature specific networks in the posterior brain regions I calculated the inter-regional phase-amplitude CFC. The

rationale for this analysis was to investigate whether the phase of prefrontal cortex modulates amplitude in the ventral and the dorsal visual stream. Such a modulation signal may provide a top-down control [Liebe et al., 2012, Zanto et al., 2011] signal to bias the competition between representations in sensory brain areas. Indeed, I observed that the gamma activity in both the dorsal and the ventral visual stream was reliably locked to phases of the prefrontal alpha oscillations.

7.3 Searching for the WM code in the vlPFC.

In the third part of my thesis I investigated patterns of the ventrolateral prefrontal cortex activity during a single- and muliti-item working memory task. To do this, I used the LFP recordings obtained from nonhuman primates to test predictions regarding prefrontal contribution to WM. To this end, I investigated the type of information coded in the vlPFC signal (load-dependent vs. abstract rule-dependent). Additionally, I tested prediction of the alpha inhibition hypothesis [Jensen et al., 2014] that the size of the alpha duty cycle modulates the duration of the gamma burst. These results are described in chapter 6.5.

7.3.1 No sign of load-dependent effect.

Despite several attempts I observed no evidence of sustained load-dependent oscillatory power modulations during maintenance in the vlPFC. I did however observe linear power changes across memory loads during encoding and retrieval (not presented in the current thesis). There may be several reasons for this null result. First, because of the extensive training that both monkeys received on a single-item version of the task, it may be that the task does not rely on the PFC anymore but rather on striatum or some other brain structures which are involved in habitual processes [Graybiel and Grafton, 2015, Shohamy, 2011]. Although I cannot fully exclude such a possibility, I have looked at the data from various periods of training and the results look very similar. This makes such an explanation rather

unlikely. Another possibility is the lack of statistical power. However, having over 200 repetitions for each condition this alternative seems also implausible. A further potential explanation is that the variance of recordings is too high. The large sample of trials may be in fact detrimental to statistical power. The investigated effect might show a substantial variability across such long recording intervals. Having around 800 correct trials (for all conditions together) the data is likely to be recorded over several hours. It is possible that there are some interfering fluctuations (likely related to vigilance or attention) in performance and the LFP signal which increase the variance or even change the way a task is performed. Indeed, while exploring the data I noticed that focusing on a subsample of experimental trials reduces the variance and leads to increased classification accuracy. This is a very interesting direction for the follow-up analyses of the current data. Finally, it is also possible that the vlPFC does not show any load-dependent changes during WM maintenance neither in a form of univariate signal intensity change nor as a multivariate pattern change.

7.3.2 Response-related information is coded in the vlPFC.

Since the prefrontal cortex has been suggested as a control hub [Curtis and D'Esposito, 2003, D'Esposito, 2007, Sreenivasan et al., 2014, Stokes, 2015] it is likely that rather than representing single items it represents more abstract features. I used pattern classification approach to decode a response mappings which may be considered an example of such an abstract representation. To this end, I trained a classifier to recognize the responses that a monkey performed in each trial. Indeed, the current results show that responses are coded in the LFP signal of the vlPFC. This is evident from classification accuracy high above chance level and significant given the reference distribution of the surrogate data. To summarize, looking at LFP data from the vlPFC a classifier was able to learn the difference between two responses monkey had to perform (match, non-match). Importantly, the classifier was able to generalize and predict the category of a new, untrained trial with a high accuracy. These results extend previous findings from studies of single unit

activity which have shown that the response mappings were indeed coded in the population of units recorded in the PFC [Meyers et al., 2012]. Note, that an observation of the vlPFC coding a response mapping does not exclude a possibility that it is also involved in representing individual items.

There is one important limitation of this finding. The response mapping in the current experiments was constant (monkeys were always pressing the left button to indicate match condition and the right for non-match). Therefore, the current study cannot fully distinguish between an abstract rule of response mapping and the response preparation. Nevertheless, it is still very interesting to observe any response-related information coded in the vlPFC.

7.3.3 The duty cycle in the vlPFC.

I further explored the current data looking at the alpha duty cycle. As described above, the duty cycle reflects the phase of alpha oscillation during which inhibition is weakest and therefore neural processing is released. The alpha-inhibition model predicts that the duration of a duty cycle varies with the alpha-band power. The higher the alpha-band power, the shorter the duty cycle (see Figure 2.3B for the model) and the higher the inhibition. Therefore, one may expect that the high frequency activity which likely reflects multi-unit activity [Mukamel et al., 2005] will vary as a function of the duration a the duty cycle. I indeed observed that the length of gamma burst decreased during short duty cycles as compared to the long duty cycles. A similar effect has been previously observed only in the primary sensory cortex [Spaak et al., 2012]. This is the first evidence that the alpha band duty cycle provides a similar mechanism of gamma inhibition in the outside of the primary sensory cortex.

7.4 General conclusions.

Together, the current results suggest that WM is a complex and highly dynamic process engaging multiple brain parts and relying on coherent, dynamic interactions across these cortical regions. Given the current findings it is increasingly difficult to believe that WM may rely on a single process located within a single neural region or a structure. Working memory rather depends on interactions across multiple levels of neural information processing ranging from slow delta oscillations to high gamma and single unit activity. Performance also depends on coherent interactions acoross multiple brain regions. These different brain networks are involved in representing various features maintained in WM. The dynamic coherence across these regions is required to efficiently attenuate networks tuned to represent redundant (or distracting) information and boost those representing task relevant features. This, however, requires more abstract properties like the stimulus and task contingencies to be also represented and communicated to the feature specific regions. One possible mechanism for such coordination is the long-range phase coherence and an interregional alpha-gamma coupling between the frontal and feature specific regions. Studies on the long-range network connectivity and dynamic interactions across brain regions and their contribution to representing various features in working memory are therefore essential and will significantly improve our understanding of the working memory.

Bibliography

Abdi, H. (2007). *Kendall rank correlation*, pages 508–510. Thousand Oaks (CA): Sage.

Aggleton, J., Keith, A., Rawlins, J., Hunt, P., and Sahgal, A. (1992). Removal of the hippocampus and transection of the fornix produce comparable deficits on delayed non-matching to position by rats. *Behavioural brain research*, 52(1):61–71.

Akyürek, E. G., Leszczyński, M., and Schubö, A. (2010). The temporal locus of the interaction between working memory consolidation and the attentional blink. *Psychophysiology*, 47(6):1134–1141.

Allison, T., Puce, A., Spencer, D. D., and McCarthy, G. (1999). Electrophysiological studies of human face perception. i: Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cerebral cortex*, 9(5):415–430.

Andres, P. (2003). Frontal cortex as the central executive of working memory: time to revise our view. *Cortex*, 39(4-5):871–895.

Astle, D. E., Nobre, A. C., and Scerif, G. (2010). Subliminally presented and stored objects capture spatial attention. *J. Neurosci.*, 30(10):3567–3571.

Astle, D. E., Scerif, G., Kuo, B. C., and Nobre, A. C. (2009). Spatial selection of features within perceived and remembered objects. *Front Hum Neurosci*, 3:6.

Atkinson, R. C. and Shiffrin, R. M. (1968). *Human memory: A proposed system and its control processes*, volume 2, pages 89–195. Academic, New York.

Awh, E., Anllo-Vento, L., and Hillyard, S. A. (2000). The role of spatial selective attention in working memory for locations: evidence from event-related potentials. *J Cogn Neurosci*, 12(5):840–847.

Awh, E. and Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends Cogn. Sci. (Regul. Ed.)*, 5(3):119–126.

Awh, E., Vogel, E. K., and Oh, S. H. (2006). Interactions between attention and working memory. *Neuroscience*, 139(1):201–208.

Axmacher, N., Cohen, M. X., Fell, J., Haupt, S., Dümpelmann, M., Elger, C. E., Schlaepfer, T. E., Lenartz, D., Sturm, V., and Ranganath, C. (2010a). Intracranial eeg correlates of expectancy and memory formation in the human hippocampus and nucleus accumbens. *Neuron*, 65(4):541–549.

Axmacher, N., Henseler, M. M., Jensen, O., Weinreich, I., Elger, C. E., and Fell, J. (2010b). Cross-frequency coupling supports multi-item working memory in the human hippocampus. *Proceedings of the National Academy of Sciences*, 107(7):3228–3233.

Axmacher, N., Lenz, S., Haupt, S., Elger, C. E., and Fell, J. (2010c). Electrophysiological signature of working and long-term memory interaction in the human hippocampus. *Eur. J. Neurosci.*, 31(1):177–188.

Axmacher, N., Mormann, F., Fernández, G., Cohen, M. X., Elger, C. E., and Fell, J. (2007). Sustained neural activity patterns during working memory in the human medial temporal lobe. *the Journal of Neuroscience*, 27(29):7807–7816.

Baddeley, A. (1998). The central executive: a concept and some misconceptions. *J Int Neuropsychol Soc*, 4(5):523–526.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends Cogn. Sci. (Regul. Ed.)*, 4(11):417–423.

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nat. Rev. Neurosci.*, 4(10):829–839.

Baddeley, A. and Hitch, G. J. (1974). *Working memory*, volume 8, pages 47–90. Academic Press.

Baddeley, A., Thomson, N., and Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Behavior*, 14:575–589.

Bahramisharif, A., van Gerven, M. A., Aarnoutse, E. J., Mercier, M. R., Schwartz, T. H., Foxe, J. J., Ramsey, N. F., and Jensen, O. (2013). Propagating neocortical gamma bursts are coordinated by traveling alpha waves. *The Journal of Neuroscience*, 33(48):18849–18854.

Bauer, M., Stenner, M.-P., Friston, K. J., and Dolan, R. J. (2014). Attentional modulation of alpha/beta and gamma oscillations reflect functionally distinct processes. *The Journal of Neuroscience*, 34(48):16117–16125.

Bays, P. M. and Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, 321(5890):851–854.

Berger, H. (1929). Über das elektrenkephalogramm des menschen. *European Archives of Psychiatry and Clinical Neuroscience*, 87(1):527–570.

Bertrand, O. and Tallon-Baudry, C. (2000). Oscillatory gamma activity in humans: a possible role for object representation. *International Journal of Psychophysiology*, 38(3):211–223.

Bigelow, J., Rossi, B., and Poremba, A. (2014). Neural correlates of short-term memory in primate auditory cortex. *Frontiers in neuroscience*, 8.

Bonnefond, M. and Jensen, O. (2012). Alpha oscillations serve to protect working memory maintenance against anticipated distracters. *Curr. Biol.*, 22(20):1969–1974.

Brooks, L. R. (1968). Spatial and verbal components of the act of recall. *Canadian Journal of Psychology*, 22:349–368.

Busch, N. A. and VanRullen, R. (2010). Spontaneous eeg oscillations reveal periodic sampling of visual attention. *Proceedings of the National Academy of Sciences*, 107(37):16048–16053.

Buschman, T. J., Denovellis, E. L., Diogo, C., Bullock, D., and Miller, E. K. (2012). Synchronous oscillatory neural ensembles for rules in the prefrontal cortex. *Neuron*, 76(4):838–846.

Buzsáki, G., Anastassiou, C. A., and Koch, C. (2012). The origin of extracellular fields and currents—eeg, ecog, lfp and spikes. *Nature reviews neuroscience*, 13(6):407–420.

Buzsáki, G. and Draguhn, A. (2004). Neuronal oscillations in cortical networks. *science*, 304(5679):1926–1929.

Buzsáki, G. and Moser, E. I. (2013). Memory, navigation and theta rhythm in the hippocampal-entorhinal system. *Nature neuroscience*, 16(2):130–138.

Canolty, R. T., Edwards, E., Dalal, S. S., Soltani, M., Nagarajan, S. S., Kirsch, H. E., Berger, M. S., Barbaro, N. M., and Knight, R. T. (2006). High gamma power is phase-locked to theta oscillations in human neocortex. *science*, 313(5793):1626–1628.

Chaieb, L., Leszczynski, M., Axmacher, N., Höhne, M., Elger, C. E., and Fell, J. (2015). Theta-gamma phase-phase coupling during working memory maintenance in the human hippocampus. *Cognitive neuroscience*, 6(4):149–157.

Christophel, T. B., Cichy, R. M., Hebart, M. N., and Haynes, J.-D. (2015). Parietal and early visual cortices encode working memory content across mental transformations. *Neuroimage*, 106:198–206.

Christophel, T. B. and Haynes, J.-D. (2014). Decoding complex flow-field patterns in visual working memory. *Neuroimage*, 91:43–51.

Christophel, T. B., Hebart, M. N., and Haynes, J.-D. (2012). Decoding the contents of visual short-term memory from human visual and parietal cortex. *The Journal of Neuroscience*, 32(38):12983–12989.

Cohen, M. X. (2014). *Analyzing neural time series data: theory and practice*. MIT Press.

Colgin, L. L. and Moser, E. I. (2009). Hippocampal theta rhythms follow the beat of their own drum. *Nature neuroscience*, 12(12):1483–1484.

Corkin, S. (2002). What's new with the amnesic patient hm? *Nature Reviews Neuroscience*, 3(2):153–160.

Cowan, N. (1995). *Attention and Memory: An Integrated Framework*. Oxford Psychology Series. Oxford University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav Brain Sci*, 24(1):87–114.

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Prog. Brain Res.*, 169:323–338.

Crowder, R. G. (1993). Short-term memory: where do we stand? *Mem Cognit*, 21(2):142–145.

Curtis, C. E. and D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends Cogn. Sci. (Regul. Ed.)*, 7(9):415–423.

Curtis, C. E. and D'Esposito, M. (2004). The effects of prefrontal lesions on working memory performance and theory. *Cogn Affect Behav Neurosci*, 4(4):528–539.

Davachi, L. (2006). Item, context and relational episodic encoding in humans. *Current opinion in neurobiology*, 16(6):693–700.

Davachi, L. and DuBrow, S. (2015). How the hippocampus preserves order: the role of prediction and context. *Trends in cognitive sciences*, 19(2):92–99.

De Almeida, L., Idiart, M., Villavicencio, A., and Lisman, J. (2012). Alternating predictive and short-term memory modes of entorhinal grid cells. *Hippocampus*, 22(8):1647–1651.

Dell'Acqua, R., Sessa, P., Toffanin, P., Luria, R., and Jolicoeur, P. (2010). Orienting attention to objects in visual short-term memory. *Neuropsychologia*, 48(2):419–428.

D'Esposito, M. (2007). From cognitive to neural models of working memory. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.*, 362(1481):761–772.

D'Esposito, M., Cooney, J. W., Gazzaley, A., Gibbs, S. E., and Postle, B. R. (2006). Is the prefrontal cortex necessary for delay task performance? Evidence from lesion and FMRI data. *J Int Neuropsychol Soc*, 12(2):248–260.

D'Esposito, M. and Postle, B. R. (2015). The cognitive neuroscience of working memory. *Annu Rev Psychol*, 66:115–142.

Destexhe, A., Bal, T., McCormick, D. A., and Sejnowski, T. J. (1996). Ionic mechanisms underlying synchronized oscillations and propagating waves in a model of ferret thalamic slices. *J. Neurophysiol.*, 76(3):2049–2070.

Druzgal, T. J. and D'esposito, M. (2003). Dissecting contributions of prefrontal cortex and fusiform face area to face working memory. *Cognitive Neuroscience, Journal of*, 15(6):771–784.

Duda, R. O., Hart, P. E., and Stork, D. G. (2012). *Pattern classification*. John Wiley & Sons.

Dvorak, D. and Fenton, A. A. (2014). Toward a proper estimation of phase–amplitude coupling in neural oscillations. *Journal of neuroscience methods*, 225:42–56.

Eichenbaum, H. (2000). A cortical–hippocampal system for declarative memory. *Nature Reviews Neuroscience*, 1(1):41–50.

Einevoll, G. T., Kayser, C., Logothetis, N. K., and Panzeri, S. (2013). Modelling and analysis of local field potentials for studying the function of cortical circuits. *Nature Reviews Neuroscience*, 14(11):770–785.

Eriksson, J., Vogel, E. K., Lansner, A., Bergstrom, F., and Nyberg, L. (2015). Neurocognitive Architecture of Working Memory. *Neuron*, 88(1):33–46.

Ester, E. F., Serences, J. T., and Awh, E. (2009). Spatially global representations in human primary visual cortex during working memory maintenance. *The Journal of Neuroscience*, 29(48):15258–15265.

Ester, E. F., Sprague, T. C., and Serences, J. T. (2015). Parietal and frontal cortex encode stimulus-specific mnemonic representations during visual working memory. *Neuron*, 87(4):893–905.

Fell, J. and Axmacher, N. (2011). The role of phase synchronization in memory processes. *Nature reviews neuroscience*, 12(2):105–118.

Fell, J., Dietl, T., Grunwald, T., Kurthen, M., Klaver, P., Trautner, P., Schaller, C., Elger, C. E., and Fernández, G. (2004). Neural bases of cognitive erps: more than phase reset. *Journal of cognitive neuroscience*, 16(9):1595–1604.

Fell, J., Ludowig, E., Staresina, B. P., Wagner, T., Kranz, T., Elger, C. E., and Axmacher, N. (2011). Medial temporal theta/alpha power enhancement precedes successful memory encoding: evidence based on intracranial eeg. *The Journal of Neuroscience*, 31(14):5392–5397.

Freunberger, R., Werkle-Bergner, M., Griesmayr, B., Lindenberger, U., and Klimesch, W. (2011). Brain oscillatory correlates of working memory constraints. *Brain research*, 1375:93–102.

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in cognitive sciences*, 9(10):474–480.

Fries, P. (2015). Rhythms for cognition: communication through coherence. *Neuron*, 88(1):220–235.

Funahashi, S., Bruce, C. J., and Goldman-Rakic, P. S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.*, 61(2):331–349.

Funahashi, S., Bruce, C. J., and Goldman-Rakic, P. S. (1993). Dorsolateral prefrontal lesions and oculomotor delayed-response performance: evidence for mnemonic "scotomas". *J. Neurosci.*, 13(4):1479–1497.

Fuster, J. M. and Alexander, G. E. (1971). Neuron activity related to short-term memory. *Science*, 173(3997):652–654.

Fuster, J. M. and Bressler, S. L. (2012). Cognit activation: a mechanism enabling temporal integration in working memory. *Trends Cogn. Sci. (Regul. Ed.)*, 16(4):207–218.

Fuster, J. M. and Jervey, J. P. (1982). Neuronal firing in the inferotemporal cortex of the monkey in a visual memory task. *The Journal of Neuroscience*, 2(3):361–375.

Gazzaley, A. and Nobre, A. C. (2012). Top-down modulation: bridging selective attention and working memory. *Trends Cogn. Sci.*, 16(2):129–135.

Gazzaniga, M. S., Ivry, R. B., and Mangun, G. R. (2002). Cognitive neuroscience: The biology of the mind.(2002).

Godefroy, O. and Rousseaux, M. (1996). Divided and focused attention in patients with lesion of the prefrontal cortex. *Brain Cogn*, 30(2):155–174.

Goldman-Rakic, P. S. (1995). Cellular basis of working memory. *Neuron*, 14(3):477–485.

Goodale, M. A. and Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends Neurosci.*, 15(1):20–25.

Gray, C. M. and McCormick, D. A. (1996). Chattering cells: superficial pyramidal neurons contributing to the generation of synchronous oscillations in the visual cortex. *Science*, 274(5284):109–113.

Graybiel, A. M. and Grafton, S. T. (2015). The striatum: where skills and habits meet. *Cold Spring Harbor perspectives in biology*, 7(8):a021691.

Griffin, I. C. and Nobre, A. C. (2003). Orienting attention to locations in internal representations. *J Cogn Neurosci*, 15(8):1176–1194.

Haegens, S., Barczak, A., Musacchia, G., Lipton, M. L., Mehta, A. D., Lakatos, P., and Schroeder, C. E. (2015). Laminar profile and physiology of the α rhythm in

primary visual, auditory, and somatosensory regions of neocortex. *The Journal of Neuroscience*, 35(42):14341–14352.

Haegens, S., Nacher, V., Luna, R., Romo, R., and Jensen, O. (2011). Alpha-Oscillations in the monkey sensorimotor network influence discrimination performance by rhythmical inhibition of neuronal spiking. *Proc. Natl. Acad. Sci. U.S.A.*, 108(48):19377–19382.

Harris, J. A., Miniussi, C., Harris, I. M., and Diamond, M. E. (2002). Transient storage of a tactile memory trace in primary somatosensory cortex. *The Journal of Neuroscience*, 22(19):8720–8725.

Harrison, S. A. and Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, 458(7238):632–635.

Hasselmo, M. E. and Stern, C. E. (2006). Mechanisms underlying working memory for novel information. *Trends Cogn. Sci. (Regul. Ed.)*, 10(11):487–493.

Hayes, J. (1952). Memory span for several vocabularies as a function of vocabulary size. *Quarterly Progress Report*, pages 338–352.

Haynes, J. D. and Rees, G. (2006). Decoding mental states from brain activity in humans. *Nat. Rev. Neurosci.*, 7(7):523–534.

Henri M. Duvernoy, F. C. (2005). *The Human Hippocampus: Functional Anatomy, Vascularization, and Serial Sections with MRI*. Springer-Verlag.

Huang, S., Rossi, S., Hämäläinen, M., and Ahveninen, J. (2014). Auditory conflict resolution correlates with medial–lateral frontal theta/alpha phase synchrony. *PloS one*, 9(10).

J., J., A., R.-L. P., E., S. E., L., A. E. B. L., M., D., J., G., J., L. E., A.L., P., and E., S. (1996). *Verbal and spatial working memory in humans*, volume 35, pages 165–192. New York: Academic Press.

Jacobs, J. (2014). Hippocampal theta oscillations are slower in humans than in rodents: implications for models of spatial navigation and memory. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 369(1635):20130304.

Jensen, O., Bonnefond, M., Marshall, T. R., and Tiesinga, P. (2015). Oscillatory mechanisms of feedforward and feedback visual processing. *Trends in neurosciences*, 38(4):192–194.

Jensen, O., Bonnefond, M., and VanRullen, R. (2012). An oscillatory mechanism for prioritizing salient unattended stimuli. *Trends Cogn. Sci. (Regul. Ed.)*, 16(4):200–206.

Jensen, O. and Colgin, L. L. (2007). Cross-frequency coupling between neuronal oscillations. *Trends in cognitive sciences*, 11(7):267–269.

Jensen, O., Gips, B., Bergmann, T. O., and Bonnefond, M. (2014). Temporal coding organized by coupled alpha and gamma oscillations prioritize visual processing. *Trends Neurosci.*, 37(7):357–369.

Jensen, O., Kaiser, J., and Lachaux, J.-P. (2007). Human gamma-frequency oscillations associated with attention and memory. *Trends in neurosciences*, 30(7):317–324.

Jensen, O. and Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Frontiers in human neuroscience*, 4.

Jensen, O. and Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. *European Journal of Neuroscience*, 15(8):1395–1399.

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

Jokisch, D. and Jensen, O. (2007). Modulation of gamma and alpha activity during a working memory task engaging the dorsal or ventral stream. *J. Neurosci.*, 27(12):3244–3251.

Kahana, M. J. (2006). The cognitive correlates of human brain oscillations. *The Journal of Neuroscience*, 26(6):1669–1672.

Katus, T. and Eimer, M. (2015). Lateralized delay period activity marks the focus of spatial attention in working memory: evidence from somatosensory event-related brain potentials. *The Journal of Neuroscience*, 35(17):6689–6695.

Katus, T., Grubert, A., and Eimer, M. (2014). Electrophysiological evidence for a sensory recruitment model of somatosensory working memory. *Cerebral Cortex*, page bhu153.

Katus, T., Müller, M. M., and Eimer, M. (2015). Sustained maintenance of somatotopic information in brain regions recruited by tactile working memory. *The Journal of Neuroscience*, 35(4):1390–1395.

Kessler, Y. and Oberauer, K. (2014). Working memory updating latency reflects the cost of switching between maintenance and updating modes of operation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(3):738.

Klimesch, W. (1999). Eeg alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain research reviews*, 29(2):169–195.

Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in cognitive sciences*, 16(12):606–617.

Klimesch, W., Doppelmayr, M., Schwaiger, J., Auinger, P., and Winkler, T. (1999). Paradoxical'alpha synchronization in a memory task. *Cognitive Brain Research*, 7(4):493–501.

Klimesch, W., Sauseng, P., and Hanslmayr, S. (2007). Eeg alpha oscillations: the inhibition–timing hypothesis. *Brain research reviews*, 53(1):63–88.

Koepsell, K., Wang, X., Hirsch, J. A., and Sommer, F. T. (2010). Exploring the function of neural oscillations in early sensory systems. *Front Neurosci*, 4:53.

Kubota, K. and Niki, H. (1971). Prefrontal cortical unit activity and delayed alternation performance in monkeys. *J. Neurophysiol.*, 34(3):337–347.

Kumaran, D. and Maguire, E. A. (2006). The dynamics of hippocampal activation during encoding of overlapping sequences. *Neuron*, 49(4):617–629.

Kuo, B. C., Rao, A., Lepsien, J., and Nobre, A. C. (2009). Searching for targets within the spatial layout of visual short-term memory. *J. Neurosci.*, 29(25):8032–8038.

Lachaux, J.-P., Rodriguez, E., Martinerie, J., Varela, F. J., et al. (1999). Measuring phase synchrony in brain signals. *Human brain mapping*, 8(4):194–208.

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., and Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science*, 320(5872):110–113.

Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., and Schroeder, C. E. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *Journal of neurophysiology*, 94(3):1904–1911.

Lee, S.-H., Kravitz, D. J., and Baker, C. I. (2013). Goal-dependent dissociation of visual and prefrontal cortices during working memory. *Nature neuroscience*, 16(8):997–999.

Lega, B. C., Jacobs, J., and Kahana, M. (2012). Human hippocampal theta oscillations and the formation of episodic memories. *Hippocampus*, 22(4):748–761.

Lepsien, J. and Nobre, A. C. (2007). Attentional modulation of object representations in working memory. *Cereb. Cortex*, 17(9):2072–2083.

Leszczyński, M., Fell, J., and Axmacher, N. (2015). Rhythmic working memory activation in the human hippocampus. *Cell reports*, 13(6):1272–1282.

Leszczyński, M., Myers, N. E., Akyürek, E. G., and Schubö, A. (2012). Recoding between two types of STM representation revealed by the dynamics of memory search. *J Cogn Neurosci*, 24(3):653–663.

Leszczyński, M., Wykowska, A., Perez-Osorio, J., and Muller, H. J. (2013). Deployment of spatial attention towards locations in memory representations. An EEG study. *PLoS ONE*, 8(12):e83856.

Levy, R. and Goldman-Rakic, P. S. (2000). Segregation of working memory functions within the dorsolateral prefrontal cortex. *Exp Brain Res*, 133(1):23–32.

Liebe, S., Hoerzer, G. M., Logothetis, N. K., and Rainer, G. (2012). Theta coupling between v4 and prefrontal cortex predicts visual short-term memory performance. *Nature neuroscience*, 15(3):456–462.

Lisman, J. E. and Idiart, M. A. (1995). Storage of 7+/-2 short-term memories in oscillatory subcycles. *Science*, 267(5203):1512–1515.

Lisman, J. E. and Jensen, O. (2013). The theta-gamma neural code. *Neuron*, 77(6):1002–1016.

Logie, R. H., Zucco, G. M., and Baddeley, A. D. (1990). Interference with visual short-term memory. *Acta Psychol (Amst)*, 75(1):55–74.

Lubenov, E. V. and Siapas, A. G. (2009). Hippocampal theta oscillations are travelling waves. *Nature*, 459(7246):534–539.

Luck, S. J. and Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, 31(3):291–308.

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

Luck, S. J. and Vogel, E. K. (2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends Cogn. Sci. (Regul. Ed.)*, 17(8):391–400.

Ma, W. J., Husain, M., and Bays, P. M. (2014). Changing concepts of working memory. *Nat. Neurosci.*, 17(3):347–356.

Maris, E. and Oostenveld, R. (2007). Nonparametric statistical testing of eeg-and meg-data. *Journal of neuroscience methods*, 164(1):177–190.

McCarthy, G., Puce, A., Belger, A., and Allison, T. (1999). Electrophysiological studies of human face perception. ii: Response properties of face-specific potentials generated in occipitotemporal cortex. *Cerebral Cortex*, 9(5):431–444.

Mehta, M. R., Lee, A. K., and Wilson, M. A. (2002). Role of experience and oscillations in transforming a rate code into a temporal code. *Nature*, 417(6890):741–746.

Meyers, E. M., Freedman, D. J., Kreiman, G., Miller, E. K., and Poggio, T. (2008). Dynamic population coding of category information in inferior temporal and prefrontal cortex. *Journal of neurophysiology*, 100(3):1407–1419.

Meyers, E. M., Qi, X. L., and Constantinidis, C. (2012). Incorporation of new information into prefrontal cortical activity after learning working memory tasks. *Proc. Natl. Acad. Sci. U.S.A.*, 109(12):4651–4656.

Miller, E. K. and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.*, 24:167–202.

Miller, E. K., Li, L., and Desimone, R. (1993). Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *The Journal of Neuroscience*, 13(4):1460–1478.

Miller, G. A. (1956). The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychol Rev*, 63(2):81–97.

Miller, G. A., Galanter, E., and Pribram, K. H. (1960). *Plans and the Structure of Behavior*. Holt Rinehart and Winston New York.

Milner, A. D. and Goodale, M. A. (1995). *The visual brain in action*. Oxford: Oxford University Press.

Monto, S., Palva, S., Voipio, J., and Palva, J. M. (2008). Very slow EEG fluctuations predict the dynamics of stimulus detection and oscillation amplitudes in humans. *J. Neurosci.*, 28(33):8268–8272.

Morillon, B., Hackett, T. A., Kajikawa, Y., and Schroeder, C. E. (2015). Predictive motor control of sensory dynamics in auditory active sensing. *Curr. Opin. Neurobiol.*, 31:230–238.

Morillon, B., Schroeder, C. E., and Wyart, V. (2014). Motor contributions to the temporal precision of auditory attention. *Nat Commun*, 5:5255.

Mukamel, R., Gelbard, H., Arieli, A., Hasson, U., Fried, I., and Malach, R. (2005). Coupling between neuronal firing, field potentials, and FMRI in human auditory cortex. *Science*, 309(5736):951–954.

Naidich, T., Duvernoy, H., Delman, B., Sorensen, A., Kollias, S., and Haacke, E. (2007). *Duvernoy's atlas: The human brain stem & cerebellum. High-field MRI: Surface anatomy, internal structure, vascularization & 3D sectional anatomy.* Springer-Verlag, Vienna.

Nichols, E. A., Kao, Y.-C., Verfaellie, M., and Gabrieli, J. D. (2006). Working memory and long-term memory for faces: Evidence from fmri and global amnesia for involvement of the medial temporal lobes. *Hippocampus*, 16(7):604–616.

Norman, D. A. and Shallice, T. (1986). *Attention to action: Willed and automatic control of behaviour*, pages 1–18. Plenum, New York.

Norman, K. A., Polyn, S. M., Detre, G. J., and Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn. Sci. (Regul. Ed.)*, 10(9):424–430.

Oberauer, K. (2001). Removing irrelevant information from working memory: a cognitive aging study with the modified Sternberg task. *J Exp Psychol Learn Mem Cogn*, 27(4):948–957.

Oberauer, K. (2002). Access to information in working memory: exploring the focus of attention. *J Exp Psychol Learn Mem Cogn*, 28(3):411–421.

Oehrn, C. R., Baumann, C., Fell, J., Lee, H., Kessler, H., Habel, U., Hanslmayr, S., and Axmacher, N. (2015). Human hippocampal dynamics during response conflict. *Current Biology*, 25(17):2307–2313.

Olson, I. R., Page, K., Moore, K. S., Chatterjee, A., and Verfaellie, M. (2006). Working memory for conjunctions relies on the medial temporal lobe. *The Journal of Neuroscience*, 26(17):4596–4601.

Owen, A. M., Herrod, N. J., Menon, D. K., Clark, J. C., Downey, S. P., Carpenter, T. A., Minhas, P. S., Turkheimer, F. E., Williams, E. J., Robbins, T. W., Sahakian, B. J., Petrides, M., and Pickard, J. D. (1999). Redefining the functional organization of working memory processes within human lateral prefrontal cortex. *Eur. J. Neurosci.*, 11(2):567–574.

Owen, A. M., Morris, R. G., Sahakian, B. J., Polkey, C. E., and Robbins, T. W. (1996). Double dissociations of memory and executive functions in working memory tasks following frontal lobe excisions, temporal lobe excisions or amygdalo-hippocampectomy in man. *Brain*, 119 (Pt 5):1597–1615.

Passingham, R. (1975). Delayed matching after selective prefrontal lesions in monkeys (Macaca mulatta). *Brain Res.*, 92(1):89–102.

Pasternak, T. and Greenlee, M. W. (2005). Working memory in primate sensory systems. *Nat. Rev. Neurosci.*, 6(2):97–107.

Pearson, D. G., Ball, K., and Smith, D. T. (2014). Oculomotor preparation as a rehearsal mechanism in spatial working memory. *Cognition*, 132(3):416–428.

Penny, W., Duzel, E., Miller, K., and Ojemann, J. (2008). Testing for nested oscillation. *Journal of neuroscience methods*, 174(1):50–61.

Penttonen, M. and Buzsáki, G. (2003). Natural logarithmic relationship between brain oscillators. *Thalamus and Related Systems*, 2(2):145 – 152.

Peyrache, A., Lacroix, M. M., Petersen, P. C., and Buzsaki, G. (2015). Internally organized mechanisms of the head direction sense. *Nat. Neurosci.*, 18(4):569–575.

Pisella, L., Berberovic, N., and Mattingley, J. B. (2004). Impaired working memory for location but not for colour or shape in visual neglect: a comparison of parietal and non-parietal lesions. *Cortex*, 40(2):379–390.

Postle, B. (2015). *Activation and information in working memory research*, pages 21–43. Wiley-Blackwell.

Postle, B. R., Druzgal, T. J., and D'Esposito, M. (2003). Seeking the neural substrates of visual working memory storage. *Cortex*, 39(4):927–946.

Raghavachari, S., Kahana, M. J., Rizzuto, D. S., Caplan, J. B., Kirschen, M. P., Bourgeois, B., Madsen, J. R., and Lisman, J. E. (2001). Gating of human theta oscillations by a working memory task. *The journal of Neuroscience*, 21(9):3175–3183.

Raghavachari, S., Lisman, J. E., Tully, M., Madsen, J. R., Bromfield, E., and Kahana, M. J. (2006). Theta oscillations in human cortex during a working-memory task: evidence for local generators. *Journal of Neurophysiology*, 95(3):1630–1638.

Ranganath, C. and Blumenfeld, R. S. (2005). Doubts about double dissociations between short-and long-term memory. *Trends in cognitive sciences*, 9(8):374–380.

Ranganath, C., DeGutis, J., and D'Esposito, M. (2004). Category-specific modulation of inferior temporal activity during working memory encoding and maintenance. *Cognitive Brain Research*, 20(1):37–45.

Ranganath, C. and D'Esposito, M. (2001). Medial temporal lobe activity associated with active maintenance of novel information. *Neuron*, 31(5):865–873.

Riggall, A. C. and Postle, B. R. (2012). The relationship between working memory storage and elevated activity as measured with functional magnetic resonance imaging. *The Journal of neuroscience*, 32(38):12990–12998.

Ritz, R. and Sejnowski, T. J. (1997). Synchronous oscillatory activity in sensory systems: new vistas on mechanisms. *Curr. Opin. Neurobiol.*, 7(4):536–546.

Roux, F. and Uhlhaas, P. J. (2014). Working memory and neural oscillations: alpha–gamma versus theta–gamma codes for distinct wm information? *Trends in cognitive sciences*, 18(1):16–25.

Rowe, J. B., Toni, I., Josephs, O., Frackowiak, R. S., and Passingham, R. E. (2000). The prefrontal cortex: response selection or maintenance within working memory? *Science*, 288(5471):1656–1660.

Ruchkin, D. S., Grafman, J., Cameron, K., and Berndt, R. S. (2003). Working memory retention systems: a state of activated long-term memory. *Behav Brain Sci*, 26(6):709–728.

Rypma, B. and D'Esposito, M. (1999). The roles of prefrontal brain regions in components of working memory: effects of memory load and individual differences. *Proc. Natl. Acad. Sci. U.S.A.*, 96(11):6558–6563.

Sarma, A., Masse, N. Y., Wang, X.-J., and Freedman, D. J. (2016). Task-specific versus generalized mnemonic representations in parietal and prefrontal cortices. *Nature neuroscience*, 19(1):143–149.

Sauseng, P., Griesmayr, B., Freunberger, R., and Klimesch, W. (2010). Control mechanisms in working memory: a possible function of eeg theta oscillations. *Neuroscience and Biobehavioral Reviews*, 34(7):1015–1022.

Sauseng, P., Klimesch, W., Doppelmayr, M., Hanslmayr, S., Schabus, M., and Gruber, W. R. (2004). Theta coupling in the human electroencephalogram during a working memory task. *Neuroscience letters*, 354(2):123–126.

Sauseng, P., Klimesch, W., Heise, K. F., Gruber, W. R., Holz, E., Karim, A. A., Glennon, M., Gerloff, C., Birbaumer, N., and Hummel, F. C. (2009). Brain oscillatory substrates of visual short-term memory capacity. *Curr. Biol.*, 19(21):1846–1852.

Sauseng, P., Klimesch, W., Schabus, M., and Doppelmayr, M. (2005). Fronto-parietal eeg coherence in theta and upper alpha reflect central executive functions of working memory. *International Journal of Psychophysiology*, 57(2):97–103.

Schapiro, A. C., Gregory, E., Landau, B., McCloskey, M., and Turk-Browne, N. B. (2014). The necessity of the medial temporal lobe for statistical learning. *Journal of cognitive neuroscience*, 26(8):1736–1747.

Schapiro, A. C., Kustner, L. V., and Turk-Browne, N. B. (2012). Shaping of object representations in the human medial temporal lobe based on temporal regularities. *Current Biology*, 22(17):1622–1627.

Schroeder, C. E. and Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci.*, 32(1):9–18.

Scott, B. H. and Mishkin, M. (2015). Auditory short-term memory in the primate auditory cortex. *Brain research*.

Scott, B. H., Mishkin, M., and Yin, P. (2014). Neural correlates of auditory short-term memory in rostral superior temporal cortex. *Current Biology*, 24(23):2767–2775.

Scoville, W. B. and Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of neurology, neurosurgery, and psychiatry*, 20(1):11.

Serences, J. T., Ester, E. F., Vogel, E. K., and Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science*, 20(2):207–214.

Shallice, T. and Warrington, E. K. (1970). Independent functioning of verbal memory stores: a neuropsychological study. *Q J Exp Psychol*, 22(2):261–273.

Shohamy, D. (2011). Learning and motivation in the human striatum. *Current opinion in neurobiology*, 21(3):408–414.

Shohamy, D. and Turk-Browne, N. B. (2013). Mechanisms for widespread hippocampal involvement in cognition. *Journal of Experimental Psychology: General*, 142(4):1159.

Siegel, M., Warden, M. R., and Miller, E. K. (2009). Phase-dependent neuronal coding of objects in short-term memory. *Proceedings of the National Academy of Sciences*, 106(50):21341–21346.

Silva, L. R., Amitai, Y., and Connors, B. W. (1991). Intrinsic oscillations of neocortex generated by layer 5 pyramidal neurons. *Science*, 251(4992):432–435.

Smith, E. and Kosslyn, S. (2007). *Cognitive Psychology: Mind and Brain*. Pearson New International Edition. Pearson Education, Limited.

Smith, E. E. and Jonides, J. (1997). Working memory: a view from neuroimaging. *Cogn Psychol*, 33(1):5–42.

Smith, E. E. and Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science*, 283(5408):1657–1661.

Smyth, M. M. and Scholey, K. A. (1994). Interference in immediate spatial memory. *Mem Cognit*, 22(1):1–13.

Spaak, E., Bonnefond, M., Maier, A., Leopold, D. A., and Jensen, O. (2012). Layer-specific entrainment of gamma-band neural activity by the alpha rhythm in monkey visual cortex. *Curr. Biol.*, 22(24):2313–2318.

Squire, L. R. and Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences*, 93(24):13515–13522.

Sreenivasan, K. K., Curtis, C. E., and D'Esposito, M. (2014). Revisiting the role of persistent neural activity during working memory. *Trends Cogn. Sci. (Regul. Ed.)*, 18(2):82–89.

Sternberg, S. (1966). High-speed scanning in human memory. *Science*, 153(3736):652–654.

Sternberg, S. (1969). Memory-scanning: mental processes revealed by reaction-time experiments. *Am. Sci.*, 57(4):421–457.

Stokes, M. G. (2015). 'Activity-silent' working memory in prefrontal cortex: a dynamic coding framework. *Trends Cogn. Sci. (Regul. Ed.)*, 19(7):394–405.

Stokes, M. G., Kusunoki, M., Sigala, N., Nili, H., Gaffan, D., and Duncan, J. (2013). Dynamic coding for cognitive control in prefrontal cortex. *Neuron*, 78(2):364–375.

Tallon-Baudry, C. and Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends in cognitive sciences*, 3(4):151–162.

Tallon-Baudry, C., Bertrand, O., Peronnet, F., and Pernier, J. (1998). Induced gamma-band activity during the delay of a visual short-term memory task in humans. *J. Neurosci.*, 18(11):4244–4254.

Tallon-Baudry, C., Kreiter, A., and Bertrand, O. (1999). Sustained and transient oscillatory responses in the gamma and beta bands in a visual short-term memory task in humans. *Vis. Neurosci.*, 16(3):449–459.

Thut, G., Nietzel, A., Brandt, S. A., and Pascual-Leone, A. (2006). Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *J. Neurosci.*, 26(37):9494–9502.

Tort, A. B., Komorowski, R., Eichenbaum, H., and Kopell, N. (2010). Measuring phase-amplitude coupling between neuronal oscillations of different frequencies. *Journal of neurophysiology*, 104(2):1195–1210.

Trost, S. and Gruber, O. (2012). Evidence for a double dissociation of articulatory rehearsal and non-articulatory maintenance of phonological information in human verbal working memory. *Neuropsychobiology*, 65(3):133–140.

van Kerkoerle, T., Self, M. W., Dagnino, B., Gariel-Mathis, M.-A., Poort, J., van der Togt, C., and Roelfsema, P. R. (2014). Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proceedings of the National Academy of Sciences*, 111(40):14332–14341.

van Vugt, M. K., Schulze-Bonhage, A., Litt, B., Brandt, A., and Kahana, M. J. (2010). Hippocampal gamma oscillations increase with memory load. *J. Neurosci.*, 30(7):2694–2699.

Vanhatalo, S., Palva, J. M., Holmes, M. D., Miller, J. W., Voipio, J., and Kaila, K. (2004). Infraslow oscillations modulate excitability and interictal epileptic activity in the human cortex during sleep. *Proc. Natl. Acad. Sci. U.S.A.*, 101(14):5053–5057.

Vogel, E. K. and Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, 428(6984):748–751.

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

Waldhauser, G. T., Johansson, M., and Hanslmayr, S. (2012). Alpha/beta oscillations indicate inhibition of interfering visual memories. *The Journal of neuroscience*, 32(6):1953–1961.

Wallis, J. D., Anderson, K. C., and Miller, E. K. (2001). Single neurons in prefrontal cortex encode abstract rules. *Nature*, 411(6840):953–956.

Watrous, A. J., Lee, D. J., Izadi, A., Gurkoff, G. G., Shahlaie, K., and Ekstrom, A. D. (2013). A comparative study of human and rat hippocampal low-frequency oscillations during spatial navigation. *Hippocampus*, 23(8):656–661.

Wimmer, G. E. and Shohamy, D. (2012). Preference by association: how memory mechanisms in the hippocampus bias decisions. *Science*, 338(6104):270–273.

Zanto, T. P., Rubens, M. T., Thangavel, A., and Gazzaley, A. (2011). Causal role of the prefrontal cortex in top-down modulation of visual processing and working memory. *Nat. Neurosci.*, 14(5):656–661.

Zhang, W. and Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, 453(7192):233–235.