The Role of Frontal Beta Oscillations in Learning and Memory

by

Azadeh Haji Hosseini
BSc, University of Tehran, 2002
MSc, Amirkabir University of Technology, 2005
MA, University of Barcelona, 2008

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Psychology

Supervisory Committee

The Role of Frontal Beta Oscillations in Learning and Memory

by

Azadeh Haji Hosseini
BSc, University of Tehran, 2002
MSc, Amirkabir University of Technology, 2005
MA, University of Barcelona, 2008

Supervisory Committee

Dr. Clay B. Holroyd, Department of Psychology
Supervisor

Dr. Adam Krawitz, Department of Psychology
Departmental Member

Dr. Roderick Edwards, Department of Mathematics and Statistics
Outside Member

Abstract

Supervisory Committee
Dr. Clay B. Holroyd, Department of Psychology
Supervisor
Dr. Adam Krawitz, Department of Psychology
Departmental Member
Dr. Roderick Edwards, Department of Mathematics and Statistics
Outside Member

Several theoretical frameworks have implicated the dorsolateral prefrontal cortex (DLPFC) and the anterior cingulate cortex (ACC) in cognitive functions related to learning, decision making, and working memory. However, the particulars about how these functions are carried out are still ambiguous. This thesis investigates whether beta oscillations, a ~20-30 Hz signal in the human electroencephalogram distributed over frontal areas of the scalp, can provide insight into these neurocognitive functions. Increased beta power has been consistently observed following presentation of reward feedback stimuli in a variety of reinforcement learning paradigms such as gambling, probabilistic learning, and time-estimation tasks. This beta oscillatory activity has been proposed to underlie learning from rewards through synchronization of prefrontal cortical regions or by mediating cross-talk between cognitive processes underlying attention, motivation, and memory, but these ideas have not been empirically tested. I hypothesized that frontal beta reflects the activation of neural ensembles in the DLPFC and ACC related to recent actions or action sequences, that bias the activity of other brain areas that are responsible for task execution in order to enhance task performance. Over trials, this process would result in the transfer of task performance from frontal brain areas to other neural systems that can execute behavior relatively automatically. This hypothesis is based partly on existing theoretical frameworks about the functions of DLPFC and ACC. I tested this hypothesis in a series of four experiments. First, I showed that, in line with my hypothesis, frontal beta oscillations do not code for a reward prediction error signal – an important signal in neural theories of reinforcement learning. Second, I showed that reward feedback stimuli compared to error feedback stimuli elicited greater beta power by the DLPFC. Third, I showed greater beta power elicited by feedback stimuli following a sequence of actions compared to just a single action, and that this contrast was associated with the ACC activity. Fourth, I found that frontal beta to reward feedback is reduced when actions preceding the error feedback are useful for desired task performance i.e., they carry task-related information and further, that higher post-feedback beta power was associated with faster recall of the stimuli associated with feedback irrespective of feedback valence. These findings indicate that frontal beta oscillations reflect a mechanism related to boosting the active representation of information related to actions or sequences of actions preceding the feedback irrespective of feedback valence. Further, they show that a general theory about the functional significance of frontal beta oscillations cannot be explained according to reinforcement learning paradigms, but rather point to a more common account that explains this neural signature through processes governed by the PFC and ACC.

Table of Contents

Supervisory Committee ii
Abstract iii
Table of Contents iv
List of Tables vi
List of Figures vii
Acknowledgments viii
Dedication ix
Chapter 1: General Introduction 1
Theoretical background 3
Neural correlates of reinforcement learning and working memory: theoretical framework 3
Brain oscillations 9
Beta oscillations 13
BG and motor cortex 15
Hippocampus 15
PFC 17
Research question 24
Theoretical and clinical importance 31
Chapter 2: Experiment One 32
Introduction 33
Methods 34
Results 38
Discussion 41
Chapter 3: Experiment Two 44
Introduction 45
Methods 46
Results 51
Discussion 53
Chapter 4: Experiment Three 56
Introduction 58
Methods 61
Results 70
Discussion 78
Chapter 5: Experiment Four 83
Introduction 84
Methods 86
Results 93
Discussion 99
Chapter 6: General Discussion 104
Summary of the current findings 105
Existing accounts 108
Learning from feedback 108

Interaction between cognitive functions 109
Frontal beta, PFC, ACC, learning, and memory 109
Other considerations 111
Future directions 113
Concluding remarks 117
Bibliography 119

List of Tables

Table 1: Electroencephalogram studies of frontal beta oscillations 110

List of Figures

Figure 1: Proposed hierarchical reinforcement learning (HRL) framework by Holroyd and Yeung (2012) 4
Figure 2: A 2-second single trial EEG signal band-pass filtered within classical frequency bands 10
Figure 3: The oscillatory model for working memory posits that the mechanism underlying working memory is cross-frequency coupling between faster and slower oscillations. 16
Figure 4: Time-frequency maps showing the burst of beta activity. 19
Figure 5: Time-frequency data for the time-estimation task 39
Figure 6: Time-frequency data at channel F4. 40
Figure 7: Virtual T-maze task. 47
Figure 8: Time-frequency and sLORETA source localization results 52
Figure 9: Time lines for example trials in the A) R-maze and B) H-maze 63
Figure 10: Time-frequency maps of EEG power from 1-40 Hz associated with channel FC3 72
Figure 11: Principal Component Analysis (PCA) results 73
Figure 12: Frontal beta power at virtual channel 1 and behavior 74
Figure 13: SLORETA maps of beta power sources 75
Figure 14: Cross-frequency coupling index (phase amplitude coupling). 76
Figure 15: Event related potentials (ERPs) to reward and error feedback stimuli, and the associated reward positivites, for the R-maze and H-maze at FCz 77
Figure 16: Task timeline. 87
Figure 17: Reaction time at recall phase 93
Figure 18: Scalp distribution of beta power 95
Figure 19: Time-frequency maps of EEG power recorded at channel F3 96
Figure 20: Beta power averaged across the three frontal virtual channels (VC1, VC3, and VC4) 98
Figure 21: Relationship between beta power and speed of recall 98

Acknowledgments

First and foremost, I would like to extend my most sincere gratitude to my supervisor, Dr. Clay Holroyd. Clay, you are an exceptional combination of intelligence, wisdom, diligence, and modesty, all of which make you an incredible mentor to learn from in many ways. Thank you for supporting me in every step of this work and showing me the way. Every time I left your office, I was filled with not only excitement for research but also admiration for human knowledge. Thank you for patiently teaching me everything. I would also like to thank Drs. Adam Krawitz and Roderick Edwards for their guidance during serving on my committee. I also thank my external examiner, Dr. James Cavanagh for his very thoughtful questions and comments on this dissertation. To Drs. Michael Hunter and Michael Masson, I thank you for all your help and advice with my statistical analyses and for constantly illuminating the world of statistics to me. To Dr. Steve Lindsay and other faculty in the CaBS program, thank you very much for providing a friendly and dynamic learning environment. I also thank Drs. Josep Marco-Pallares and Antoni Rodriguez-Fornells in the University of Barcelona who first equipped me with the tools for research in Cognitive Neuroscience— I will always be grateful to you. I would also like to thank Mr. Chris Darby for his help with all sorts of technical issues— you have been my savior many times. I also thank my fellow graduate students and the research assistants in the Learning and Cognitive Control Laboratory and other residents of the basement of the Cornett building with whom I have shared space and memories. Akina Umemoto, I am sure that I will think with a big smile on my face of all the days and nights we have spent in the Cornett basement— thank you for being a friend. I am also grateful to my friends around the world who received me every time I visited, recharged my batteries with love and friendship and sent me back to work. In particular, I thank Leila Azodi Ghajar who has been my best friend since our first day of primary school— I simply could not imagine doing any of this without your friendship and emotional support. And last but not the least, I would like to thank my good brothers, Amirhossein and Arash, for their continuous support, love, and kindness and for setting good examples to their younger sister.

Dedication

To my father, whose love and trust are the pillars of my existence.

Baba, you were with me every moment of this. Thank you for leaving me memories that have never ceased to warm my heart and light my way.

and

To my mother, who pushed me forward every time I was too hesitant.

Maman, this is all because you generously let me fly. Thank you.

Chapter 1: General Introduction

Neural oscillations that are recorded from the surface of the scalp reflect synchronization of neuronal assemblies in the brain, especially in the cortex. Theoretical and empirical considerations indicate that this synchronization is due to correlations between the spatiotemporal patterns of neuronal activity producing rhythmicity despite the stochastic nature of spiking in cortical neurons (Wang, 2010). Although neurons show a high tendency to engage in oscillatory behavior, the role of these oscillations in cognition is only recently becoming clear. What type of interaction among this immense number of neurons gives rise to higher cognitive functions? How do these oscillatory interactions provide information about cognitive and motor function and dysfunction?

The topic of this dissertation is the association of beta oscillations (20-30 Hz) with feedback-guided learning. Learning from positive and negative performance feedback, usually presented in form of rewards and punishments, or *reinforcement learning* is a crucial topic in cognitive control that underlies many aspects of behavior, and is often disrupted in psychiatric and neurological disorders. In recent years, it has been consistently observed that reward feedback stimuli in tasks that provide reward or non-reward performance feedback elicit a burst of beta power over frontal areas of the scalp in humans that can be detected by recording the electroencephalogram (EEG). The underlying mechanisms of this oscillatory signature and its specific relation to reward feedback have been the topic of a few hypotheses, but these ideas are rapidly evolving. Currently, ideas regarding the function of these oscillations can be divided into two main lines: a) reward-related beta oscillations underlie learning from rewards, in contrast to

theta (4-8 Hz) oscillations that have been attributed to learning from errors (Cohen, Wilmes, & van de Vijver, 2011), and b) reward-related beta oscillations represent an interplay between neural systems underlying attention, memory, and motivation (Marco-Pallarés, Münte, & Rodríguez-Fornells, 2015). These proposals are mostly speculative about the neural source of frontal beta, pointing to several frontal brain areas including fronto-polar cortex, orbitofrontal cortex (OFC), dorsolateral prefrontal cortex (DLPFC), and anterior cingulate cortex (ACC) as their possible source. Support for these theories is found in multiple studies that reported reward-related beta power in gambling, probabilistic learning, task-switching, and time-estimation paradigms with positive and negative performance feedback (Cohen, Elger, & Ranganath, 2007; Cunillera et al., 2012; Marco-Pallares et al., 2008; van de Vijver, Ridderinkhof, & Cohen, 2011). However, few studies have directly tested theories about the source and function of frontal beta oscillations (Luft, 2014; Marco-Pallarés et al., 2015).

In this dissertation, I address the question of the source and function of frontal beta oscillations with four experiments, the details of which are described in the following chapters. These studies help to clarify the role of human EEG oscillatory behavior as it relates to theoretical frameworks of reinforcement learning, and in particular highlight the interplay between cognitive functions involved in reinforcement learning and working memory. Importantly, the non-invasive nature of the EEG technique makes the findings amenable to further study using non-invasive stimulation techniques for therapeutic purposes.

In this chapter, I introduce the problem and the theoretical framework that guides my hypotheses, predictions, and interpretations throughout this dissertation.

Theoretical background

Cognitive control and decision making in goal-directed behavior are mediated by a complex system of brain areas. Several theoretical frameworks have elucidated the functional significance of the network of areas that comprise this system, including the ACC, prefrontal cortex (PFC), basal ganglia (BG), and the midbrain dopamine system (e.g., Botvinick, Niv, & Barto, 2009; Holroyd & Coles, 2002; Miller & Cohen, 2001). Here, I utilize a recent theoretical framework developed by Holroyd and Yeung (2012) to provide context for understanding frontal beta oscillations. Although the theory is specifically about the function of the ACC, it leverages several existing ideas about the functions of the PFC, BG, dopamine, and the related neural systems. In doing so, I will review several influential ideas on the role of PFC in guiding working memory processes for the purpose of cognitive control and decision making.

Neural correlates of reinforcement learning and working memory: theoretical framework

As described above, my experimental predictions relate to dopamine activity and a network of neural regions including the PFC, BG, ACC, and other brain areas. Figure 1 illustrates the functions and interactions of these brain areas as proposed by a recent theory of ACC (HRL-ACC theory; Holroyd & Yeung, 2012). In what follows I will describe each of the components of this system in turn, devoting the most attention to the brain areas most relevant to this thesis. I will later situate the literature about frontal beta oscillations within this theoretical framework.

Dorsal striatum. The striatum is a part of the collection of subcortical nuclei called the basal ganglia (BG) that are primarily associated with motor control. The dorsal part of the striatum is suggested to have the role of 'actor' in reinforcement learning models (Botvinick et al., 2009; O'Doherty et al., 2004). An actor module selects actions based on reward signals received from a 'critic' in the classic 'actor-critic' model structure (Sutton & Barto, 1998). In the HRL-ACC theory, dorsal striatum is said to select the 'primitive actions' that comprise the high-level extended action sequences mediated by the ACC (Botvinick et al., 2009; Holroyd & Yeung, 2012).

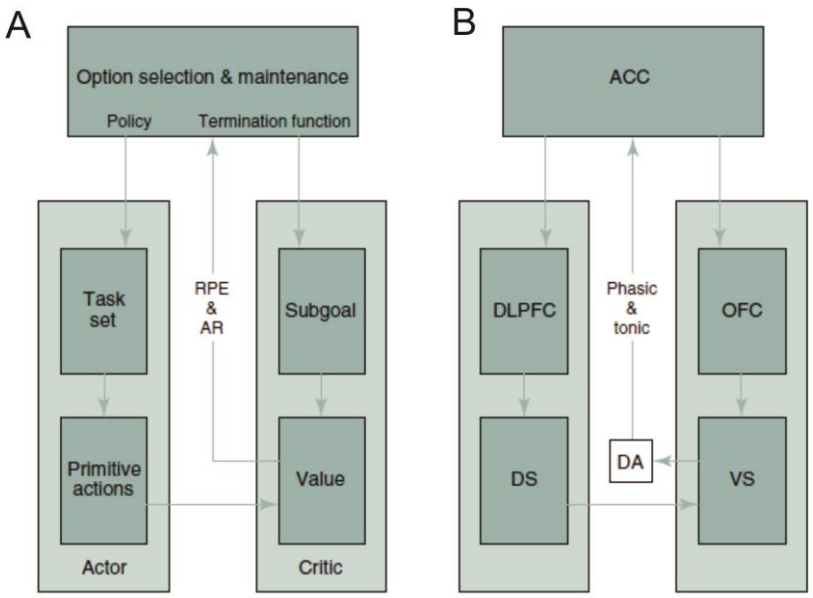

Figure 1: Proposed hierarchical reinforcement learning (HRL) framework by Holroyd and Yeung (2012) A) Abstract functions of the model components based on an actor-critic structure (Sutton & Barto, 1998) where selection of an option (action sequence) is at the top of the hierarchy. Execution of single (primitive) actions is governed by the actor in lower levels of the hierarchy and based on the reward prediction error (RPE) signal and the average reward (AR) calculated by the critic. B) Proposed neural implementation of the hierarchy where ACC is in charge of selection and maintenance of action sequences and DLPFC in charge of controlling their execution by the dorsal striatum (DS). Ventral striatum (VS) together with the OFC evaluate the current state based on predicting the future rewards. Dopamine (DA) signal from the midbrain (not shown) carries the RPE to all areas in the network (not shown except for the projection to ACC) From (Holroyd & Yeung, 2012).

Ventral striatum. The ventral part of the striatum is suggested to have the role of 'critic' in reinforcement learning models (Botvinick et al., 2009; O'Doherty et al., 2004). The critic module learns to calculate the value of current states based on the predictions about future reward (Sutton & Barto, 1998). According to a common account, the ventral striatum is part of the critic that evaluates ongoing events in order to predict future reward or punishment.

OFC. This area is highly implicated in processing of different types of rewards (Kringelbach, 2005). Holroyd & Yeung (2012) have proposed that this area constitutes part of the critic. However, I will not discuss the OFC further as it is not the focus of this thesis.

DLPFC. The DLPFC is strongly implicated in neurocognitive processes related to working memory, especially the active maintenance of action representations, and the update and implementation of task sets. One of the pioneer computational theories proposed for the function of PFC posits that cognitive control in the brain is governed by the PFC through active maintenance of task-related information necessary for goal-directed behavior. This is accomplished by neural pathways from PFC that send signals to other areas (including the dorsal striatum; Figure 1) that map inputs and internal states to outputs in a "top-down" manner (Miller & Cohen, 2001). Top-down control occurs when cognitive processes are accomplished based on the internal state of the system rather than driven by the external inputs to it. Therefore top-down control is highly associated with high-level cognitive processes such as cognitive control, decision making, and working memory. Supported by evidence from human and monkey studies, Miller and Cohen (2001) proposed that neurons in the PFC are the most probable

candidates for sending top-down control signals to other brain areas due to their widely distributed anatomical connections across the brain (Pandya & Yeterian, 1990; Pandya & Barnes, 1987), their ability to maintain the activation of rule representations through sustained activity (Fuster & Alexander, 1971), and their flexibility for fast updating of these representations (perhaps by way of a gating mechanism governed by phasic activity of dopamine neurons, see below; Braver & Cohen, 2000). Strong evidence for this theory is provided by direct recordings from non-human primate PFC and neuroimaging studies in humans (Sakai, 2008). Several studies have identified task-specific ensembles of neurons in different regions of the PFC including the DLPFC (Buschman, Denovellis, Diogo, Bullock, & Miller, 2012; Hoshi, Shima, & Tanji, 2000; Sakai & Passingham, 2003).

ACC. In the system of brain areas that are proposed to govern cognitive control, the function of ACC is highly controversial (Holroyd & Yeung, 2011). A recent proposal holds that the ACC is responsible for motivating the execution of extended sequential behaviors according to principles of hierarchical reinforcement learning (HRL-ACC theory; Holroyd & Yeung, 2012). Hierarchical reinforcement learning (HRL) approaches propose that learning from positive and negative outcomes can be explained based on hierarchical organization of behavior by letting the agent learn in a hierarchical manner by creating sub-goals for action sequences called ‘options’ (Botvinick et al., 2009). This process is governed by a system of brain areas whose function links to different levels of this hierarchy (e.g., Figure 1A and B). In reinforcement learning, options are defined as a sequence of actions (or a ‘policy’) that is executed if the agent is in a valid state for the initiation of that option and is terminated when a valid termination state for that option is

reached (Sutton, Precup, & Singh, 1999). The HRL-ACC theory suggests this area is responsible for selecting options to be executed by the actor and is constantly updated by the critic about the current state of the system (Holroyd & Yeung, 2012). If option execution by the DLPFC and the striatum goes wrong, the critic communicates a signal to the ACC which is in turn directed to the DLPFC to return to the appropriate option. Of note is that the theories of PFC that were mentioned in the previous section suggest that the PFC provides top-down control over execution of task sets, but it is not specified how task sets are learned and selected. The HRL-ACC theory provides an answer to this problem.

Also, ACC has been observed to be involved in the maintenance of task sets by showing sustained task-related activity across multiple, highly dissimilar tasks (Dosenbach et al., 2006). In sum, representation, updating, and implementation of task sets are governed by different regions in the PFC including the DLPFC and the ACC. In this thesis, I use 'task set' and 'option' interchangeably.

Dopamine. Dopamine is a neurotransmitter produced in the substantia nigra pars compacta and ventral tegmental area in the midbrain and projected widely to other brain areas including the BG, PFC, and ACC. Dopamine is tightly linked with the current theories of reinforcement learning and working memory, and the functions of the PFC and ACC in relation to these cognitive processes. The phasic and tonic activity of dopamine neurons have been implicated in motor and cognitive functions. In the context of reinforcement learning, phasic (i.e., brief) dopamine activity is believed to index a reward prediction error (RPE) signal. The RPE is an important component of reinforcement learning models that indicates whether ongoing events are better or worse

than expected: Rewards or events that are better than expected elicit a positive RPE whereas punishments or events that are worse than expected elicit a negative RPE (Sutton & Barto, 1998). Crucially, single unit recordings from monkey dopamine neurons have shown that outcomes that are better or worse than expected are coded by an increase or a decrease in the activity of midbrain dopamine neurons respectively, therefore indexing an RPE (Schultz, Dayan, & Montague, 1997). Considerable evidence indicates that phasic dopamine signals contribute to reinforcement learning in the striatum (Schultz, 2002). For example, dopamine dysfunction in Parkinson's disease and schizophrenia is associated with impairments in reinforcement learning (Frank, Seeberger, & O'reilly, 2004; Gold, Waltz, Prentice, Morris, & Heerey, 2008). On the other hand, the tonic (i.e., sustained) activity of dopamine neurons is suggested to control action vigor by indexing the average reward received (Niv, Daw, Joel, & Dayan, 2007). In the HRL-ACC theory, the ACC learns the overall option values by the phasic dopamine RPE signals and maintains the options in the working memory by tonic dopamine signals representing average reward as specified by a recent computational model (Holroyd & McClure, 2015) (Figure 1A).

Importantly dopamine is also largely implicated in working memory processes by regulating the "breadth" of information held in working memory "buffers" (Seamans & Yang, 2004) or by a gating mechanism that regulates access of new information to the PFC (Braver & Cohen, 2000). To be specific, the activity of D1 and D2 dopamine receptors in the PFC have been associated with robust maintenance of working memory information and response flexibility is switching between states, respectively (Durstewitz & Seamans, 2008) . Also, dopamine dysfunction in disorders such as Parkinson's disease

and schizophrenia has been reported to impair working memory (Beato et al., 2008; Lee et al., 2010; Lewis, Slabosz, Robbins, Barker, & Owen, 2005).

Of note is that the specific interactions between the PFC and BG are also proposed to contribute to working memory processes. Inspired by the theories for the PFC-BG interactions in motor control, and by the role of dopamine in gating as proposed by Braver and Cohen (2000), it was proposed that BG acts as a more powerful and indirect gating mechanism for updating working memory representations in the PFC (Frank, Loughry, & O'Reilly, 2001). This gating mechanism is proposed to be triggered by dopamine release that consequently relieves the disinhibition from the BG on the PFC. Further, it is suggested that a successful model for performance in reinforcement learning and working memory tasks must combine both features, a model that suggests that reinforcement learning and working memory processes are governed by the BG and PFC, respectively (Collins & Frank, 2012).

Brain oscillations

The recent advancements in computational techniques for processing of electrophysiological data have given impetus to the study of brain oscillations in association with cognitive processes. The influential theories reviewed above and their instantiations in computational models are based on classical ideas related to synaptic connections across neural networks, but are silent about how brain oscillations facilitate communication across these networks. In this thesis I hope to help bridge that divide by elucidating the contribution of frontal beta oscillations to these neurocognitive functions. But first, to put these ideas in context I begin by reviewing the literature on the functional

role of different ranges of brain oscillations related to a variety of such processes. It is important to note that the relationship between neuronal spiking activity, neural network oscillations, and the EEG is a substantial topic that is outside the scope of this thesis. Here, I focus on the neural oscillations at the network and EEG levels.

The physiological properties of different types of neurons, the architecture of

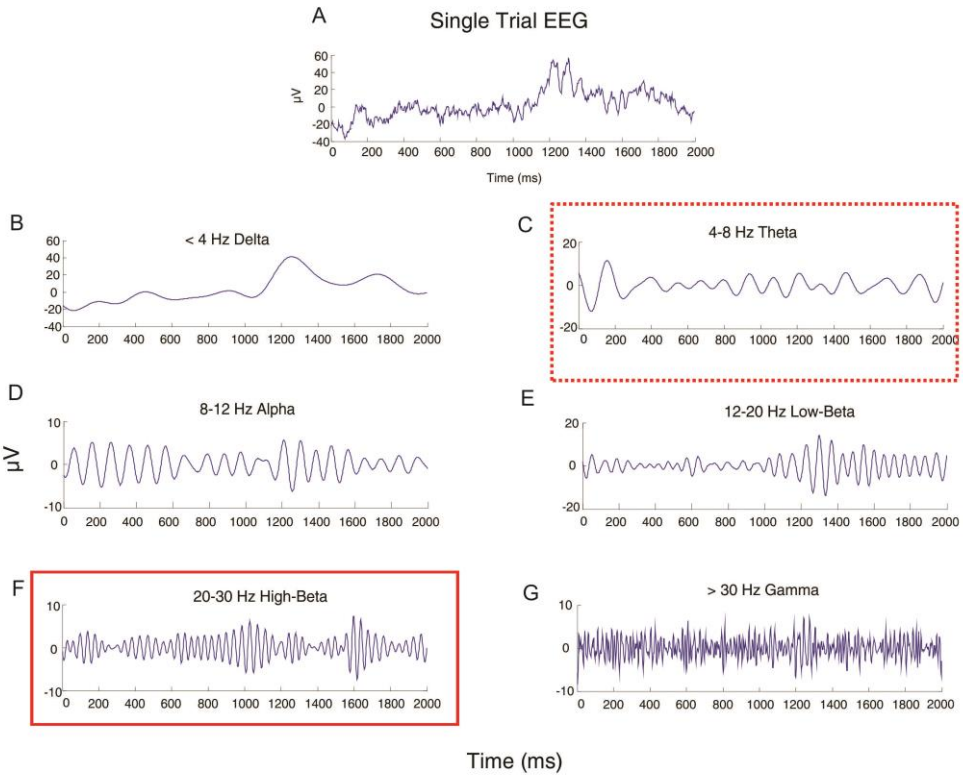

Figure 2: A 2-second single trial EEG signal band-pass filtered within classical frequency bands. A) A single trial of EEG data where a stimulus was presented at 1000 ms. Panels B-G show that there is activity in all frequency bands: B) delta, C) theta, D) alpha, E) low-beta, F) high-beta, and G) gamma before and after the stimulus presentation. The red solid and dotted boxes in F and C show the frequency bands of interest in this dissertation: High-beta is the main focus and theta is studied in terms of its interaction with beta. Data is selected from a random trial in a single subject performing a task where visual stimuli are presented on the screen.

neuronal networks and the speed of neuronal conduction contribute to the production of oscillations in different ranges of frequencies in these networks (Buzsáki & Draguhn,

2004). Higher frequencies (faster oscillations) are mostly observed in local networks whereas lower frequencies (slower oscillations) can exist in larger networks (Von Stein & Sarnthein, 2000). Computational feasibility and empirical evidence suggest that wide-spread slower oscillations can act as a neural carrier to modulate faster oscillations; in fact this interplay between the faster and slower oscillations is proposed to provide a mechanism to facilitate successful performance of brain operations at multiple spatial and temporal scales (Buzsáki & Draguhn, 2004; Chrobak & Buzsáki, 1998; Jensen & Colgin, 2007; Lisman & Idiart, 1995). These ranges of frequencies have been roughly named as delta (<4 Hz), theta (4-8 Hz), alpha (9-12 Hz), beta (13-30 Hz), and gamma (>30 Hz) rhythms (Figure 2) and have been studied using multiple invasive and non-invasive techniques including single neuron recordings, local field potentials (LFP), EEG, and intracranial EEG (iEEG), giving rise to numerous studies showing that the neural activity in these frequency bands is associated with mechanisms that underlie different cognitive functions (Başar, Başar-Eroglu, Karakaş, & Schürmann, 2000; Buzsáki & Draguhn, 2004; Sauseng & Klimesch, 2008; Wang, 2010).

For example, delta band oscillations (Figure 2B) are found to be associated with non- rapid eye movement sleep (Dang-Vu et al., 2005), memory consolidation during sleep (Gais & Born, 2004; Genzel, Kroes, Dresler, & Battaglia, 2014), feedback stimuli indicating rewards or correct responses (Bernat, Nelson, Holroyd, Gehring, & Patrick, 2008), and predicting behavioral adaptation (Cavanagh, 2015). Theta oscillations (Figure 2C) are largely associated with spatial navigation (O'Keefe & Recce, 1993) and memory processes driven by the hippocampus (Klimesch, 1999; Lisman & Idiart, 1995). Moreover, frontal midline theta oscillations are associated with working memory (Jensen

& Tesche, 2002), episodic memory (Hsieh & Ranganath, 2014), unexpectedness and uncertainty (Cavanagh, Figueroa, Cohen, & Frank, 2012), motivation, task difficulty, and mental effort (Cavanagh & Frank, 2014; Cavanagh, Zambrano-Vazquez, & Allen, 2012; Mitchell, McNaughton, Flanagan, & Kirk, 2008). Alpha band oscillations (Figure 2D) are found to be associated with inhibition of irrelevant information both in task-switching (Buschman et al., 2012), and working memory processes (Sauseng et al., 2009), inhibition of task-irrelevant regions (Jensen & Mazaheri, 2010), and internal mental processes as opposed to the processing of external stimuli (Knyazev, Slobodskoj-Plusnin, Bocharov, & Pylkova, 2011). Gamma band oscillations (Figure 2G) are detected in many brain regions and have been associated with the integration of visual features into unified percepts, also called 'binding' (Buzsáki, 2006; Tallon-Baudry & Bertrand, 1999), with the integration of sensory information (Fries, Nikolić, & Singer, 2007; Singer et al., 1997), and with attention and memory (Jensen, Kaiser, & Lachaux, 2007). Beta oscillations (Figure 2E and F) will be discussed in detail in the next section.

Brain oscillations may provide a timing mechanism for neural communication across the brain; disruption of such a timing mechanism has been postulated to occur in several neurological disorders associated with cognitive and motor dysfunctions including Alzheimer's disease, Parkinson's disease, schizophrenia, and epilepsy (Uhlhaas & Singer, 2006). As reviewed above, the classical frequency bands have been found to be associated with multiple cognitive and motor functions. Yet it is known that the same frequency band can be produced in multiple ways in different brain structures (Kopell, 2013). Therefore there is not a one-to-one mapping between classical frequency bands

and functions or brain areas. However there is converging evidence regarding the association of certain frequency bands in humans with specific cognitive functions.

In particular, investigating the association of brain oscillations with cognitive processes through time-frequency analysis of electrophysiological data has been an active area of research in recent years. In the EEG literature, these investigations have led to valuable information on the association of power and phase dynamics of these oscillations with cognitive functions (Başar, Başar-Eroglu, Karakaş, & Schürmann, 1999; Engel, Fries, & Singer, 2001; Fries, 2005). Importantly, these findings are even more significant when supported by previous findings using classical neurocognitive measures like event related potentials (ERPs) and behavioral data. Contemporary computational modeling efforts are also advancing this research by linking together the triangle of behavioral, electrophysiological, and computational evidence for the hypotheses under investigation (e.g., Cavanagh, Eisenberg, Guitart-Masip, Huys, & Frank, 2013; Collins, Cavanagh, & Frank, 2014; Mas-Herrero & Marco-Pallarés, 2014).

Beta oscillations

Beta band oscillations (12-30 Hz; Figure 2E and F) are associated with movement preparation and inhibitory control in the motor system; beta power decreases at the onset of movement execution and increases when a response is withheld (Brittain & Brown, 2014; Wang, 2010). Beta is also associated with sensory motor integration and top-down signaling (Wang, 2010). Motor-related beta activity usually engages a full range of beta with a peak at 20 Hz (Davis, Tomlinson, & Morgan, 2012) whereas beta oscillations elicited by reward-feedback stimuli (Cohen, Elger, & Ranganath, 2007; HajiHosseini et

al., 2012; Marco-Pallares et al., 2008) and those associated with task-related neural ensembles in the monkey DLPFC (Antzoulatos & Miller, 2014; Buschman et al., 2012) are above 20 Hz (Figure 2F). The higher range of beta oscillations (~20-35 Hz) are also associated with working memory according to the oscillatory model for working memory, as will be discussed below (Lisman & Idiart, 1995). It is also suggested that slower beta oscillations occur in the deeper layers of the cortex compared to faster gamma oscillations that occur in more superficial layers (Buffalo, Fries, Landman, Buschman, & Desimone, 2011).

Beta oscillations are relatively less well understood in terms of their functionality, perhaps because of their excessive activity in Parkinson's disease and other disorders, suggesting that that they are a sign of neurocognitive dysfunction, and because of contradictory results regarding the effect of inducing beta oscillations in motor cortex with brain stimulation techniques (Feurra et al., 2011; Pogosyan, Gaynor, Eusebio, & Brown, 2009). These ambiguous findings might explain why a recent article on augmenting cognition by inducing oscillations has not addressed beta oscillations at all (Horschig, Zumer, & Bahramisharif, 2014). Also, some hypotheses associate beta oscillations with seemingly contradictory neurocognitive mechanisms such as "maintenance of the status quo" (Engel & Fries, 2010) versus indexing the likelihood of the need for a novel action (Jenkinson & Brown, 2011); I will review these ideas later in this chapter. However, the most common idea across all the interpretations about the functionality of beta oscillations is that beta oscillations provide an ideal mechanism for neural communication across distributed brain areas. Brain regions where the most prominent effects on beta oscillations have been observed are the following.

BG and motor cortex

Beta oscillations are prominently seen in the BG (Jenkinson & Brown, 2011). In fact, observations of pathological beta oscillations in the motor cortex and BG in Parkinson's disease may have given rise to the belief that synchrony in beta frequency band is a sign of dysfunctional neural network. Also, based on evidence that beta activity in the BG is altered by drugs that manipulate dopamine and its receptors, it has been suggested that amount of beta in the BG indicates net dopamine levels at the sites of cortical inputs to the BG (Hammond, Bergman, & Brown, 2007; Jenkinson & Brown, 2011). Overcoming this excessive beta synchrony in the BG is the basis of deep brain stimulation (DBS) techniques for treating Parkinson's disease. However, the non-frequency selective attenuation of LFP activity in beta frequency range by the conventional DBS methods is thought to underlie the common side effects of DBS treatments, including speech and cognitive problems, paradoxical movement deterioration (Chen et al., 2006), and possibly impulsivity (Luigjes et al., 2011; but see Hälbig et al., 2009). This may suggests a functional role for beta in motor and cognitive behavior.

Hippocampus

Hippocampus is a brain structure located under the cerebral cortex in both hemispheres and has been largely associated with memory and spatial navigation (Pinel & Edwards, 2008). Based on recordings from hippocampal place cells in mice (Berke, Hetrick, Breck, & Greene, 2008), it is suggested that beta oscillations provide a mechanism for a fast and stable memory formation when the mice navigate in a new environment (Grossberg, 2009). Also, according to an influential theory, working memory maintenance occurs through the interaction of beta-gamma and theta oscillations

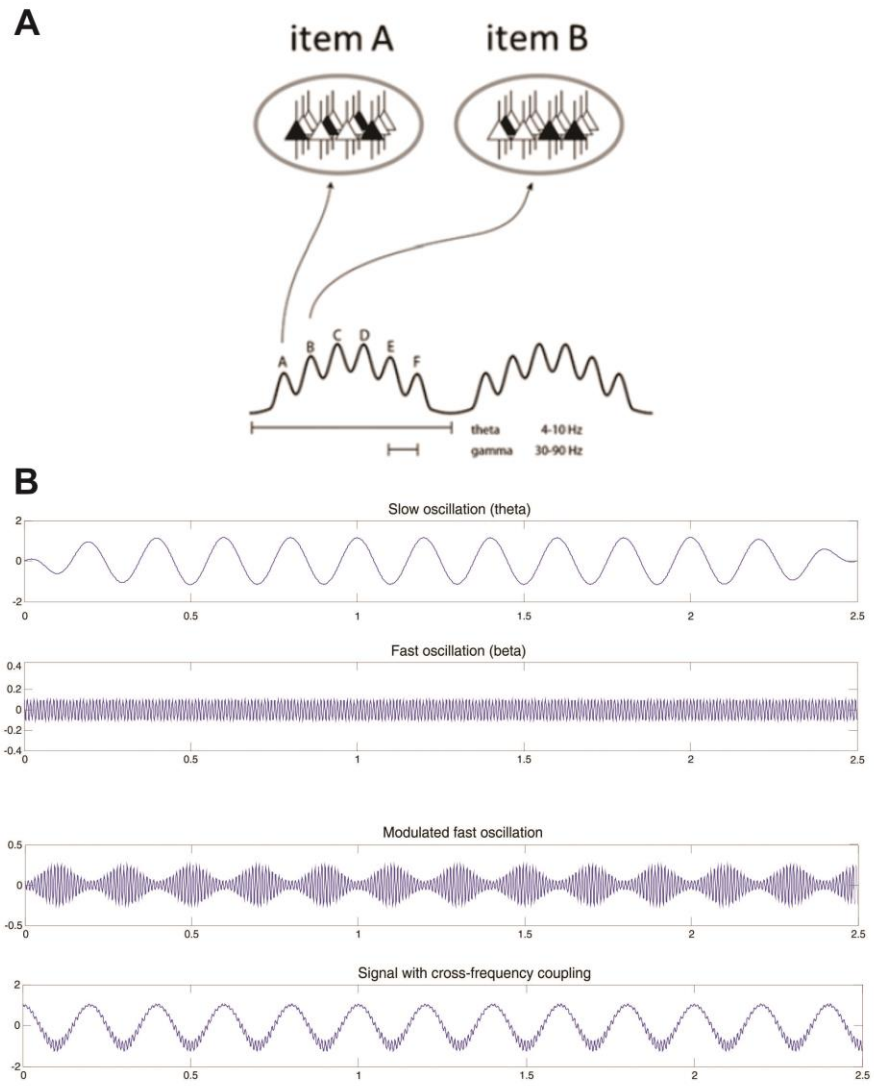

Figure 3: The oscillatory model for working memory posits that the mechanism underlying working memory is cross-frequency coupling between faster and slower oscillations. A) Every single item is stored in a cycle of gamma and contributes to a state of the network which repeats itself through every cycle of theta; adapted from Lisman and Jensen (2013). B) Simulation showing the effect of cross-frequency coupling in creation of a pattern predicted by the theory in A.

in the hippocampus (Jensen & Lisman, 1998; Lisman & Idiart, 1995): Each item of information is stored in a short cycle of higher-frequency beta or gamma oscillations, coupled with a certain phase of lower-frequency oscillations such as theta. On this view, every cycle of theta holds several cycles of repeating beta or gamma oscillations, a

process that allows for maintaining those items of information (Lisman & Idiart, 1995) (Figure 3A). This theory has recently been generalized to account for communication of "multi-item messages" across the brain through coupling between the slow and fast oscillations, which is called the "theta-gamma neural code" (Lisman & Jensen, 2013). Cross-frequency coupling between beta-gamma and theta oscillations, manifested as increased beta power at certain phases of theta oscillations (Figure 3B; Axmacher et al., 2010; Canolty et al., 2006), has been proposed to capture this mechanism: iEEG recordings from the hippocampus of epilepsy patients reveal that the power of beta oscillations increases at the trough of theta cycles and the degree of this coupling increases with working memory load (Axmacher et al., 2010).

Although there are remaining questions regarding this coupling mechanism, it is proposed to provide an ideal candidate for many cognitive and motor functions where multiple brain areas are involved.

PFC

The most consistent observations of beta oscillations in the PFC are from LFP recording in non-human primates (Benchenane, Tiesinga, & Battaglia, 2011). To be specific, recordings from the banks of the principle sulci in monkeys, homologues of the DLPFC area in humans, show that synchrony in the beta frequency band is associated with the formation and selection of task-relevant neural ensembles in a task-switching paradigm (Buschman et al., 2012). Further, another study revealed that beta oscillations are synchronous between the PFC and striatum in monkeys during categorical learning (Antzoulatos & Miller, 2014). These findings provide converging evidence on the role of beta oscillations in bottom-up and top-down interactions between the PFC and the areas

that are responsible for processing of input such as the visual cortex (Benchenane et al., 2011) and task execution such as the striatum (Antzoulatos & Miller, 2014; Feingold, 2011). Buschman and colleagues (2012) also found that this mechanism is driven by interaction between beta and alpha oscillations such that beta synchrony "selects" the rule-relevant neural ensemble and alpha "deselects" the rule-irrelevant neural ensemble in a two-rule task switching paradigm. Besides the oscillatory model for working memory briefly described above, this is another example suggesting that the interplay between multiple frequencies might be a key to explain their function.

In humans, reward feedback stimuli elicit beta oscillations distributed over the frontal areas of the scalp (Cohen et al., 2007; Doñamayor, Marco-Pallarés, Heldmann, Schoenfeld, & Münte, 2011; HajiHosseini et al., 2012; Marco-Pallares et al., 2008). For example, Cohen and colleagues (2007) tested the brain oscillatory responses to positive and negative feedback in a probabilistic learning paradigm in which subjects selected one of two targets with changing probability of wins across blocks, and were presented with feedback at the end of trial. They found an increase in beta power after reward feedback that was larger for low probability wins, although the statistical significance of this effect was measured across all frequency bands, not just beta. They interpreted this finding as enhanced connectivity following gains but not losses.

In a gambling paradigm involving a choice between two possible monetary bets, Marco-Pallares et al. (2008) found that monetary gains elicited an increase in beta power (20-30 Hz) over frontal and central areas of the scalp. This change in beta power was sensitive to the magnitude of gains over frontal areas of the scalp so that larger magnitude gains elicited more beta power. The authors suggested that burst of beta activity elicited

by rewards is related to the synchronization of a broad network of brain areas that process rewards and emotion. Based on the characteristics of the hemodynamic correlates of EEG (Kilner, Mattout, Henson, & Friston, 2005), they also proposed that the power of neural oscillations in beta frequency range are related to an increase in the blood oxygenation level dependant (BOLD) signal, supporting the idea that beta activation is associated with the activation of rostral ACC, superior frontal gyrus, posterior cingulate cortex, and striatum to rewards. Marco-Pallares and colleagues have also proposed that beta is

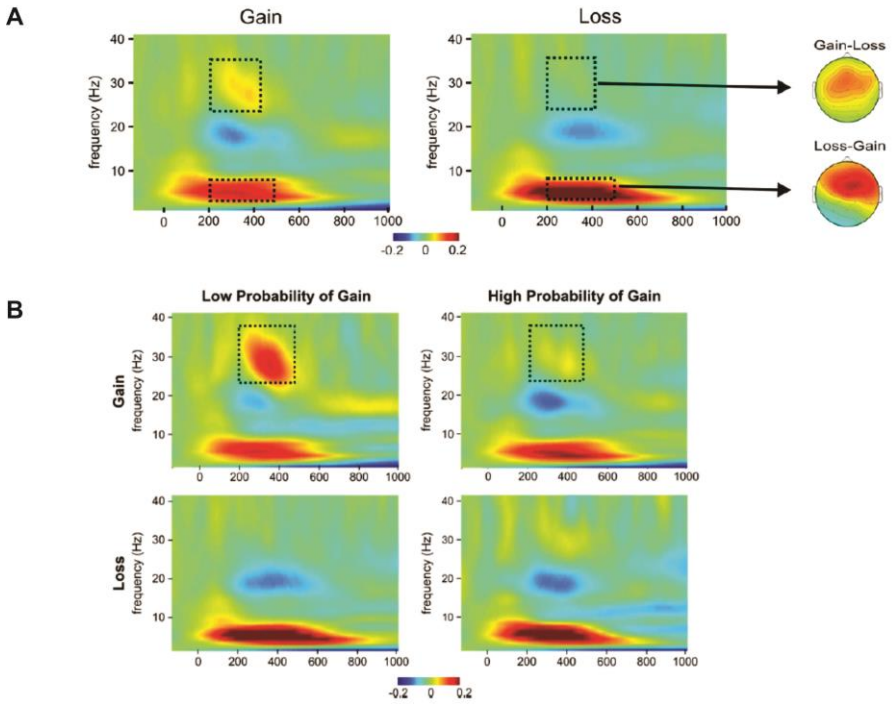

Figure 4: Time-frequency maps showing the burst of beta activity in A) gains vs. losses, and B) unexpected gains vs. losses. Beta contrast is frontally distributed. Adapted from HajiHosseini and colleagues (2012).

associated with increased phasic dopaminergic activity in the striatum due to decreased tonic dopaminergic activity in the PFC (Marco-Pallarés et al., 2009).

Following this, a second study by this group used a different gambling paradigm where magnitude and probability of reward were modified (HajiHosseini et al., 2012).

Here, subjects were informed of the magnitude and probability of the upcoming reward at the start of each trial; therefore they had an accurate estimate of the reward probability, i.e., high or low probability. Subjects were instructed to select one of 3 cards presented on a computer screen. Following their choice, the feedback stimulus revealed whether the selected card was a win or loss as well as the two other non-selected cards. Results showed a main effect of valence and probability on beta power, meaning that rewards elicited more beta power to gain (reward) feedback than to loss feedback (Figure 4A) and that unexpected rewards elicited more beta power than expected rewards (Figure 4B). However, an analysis of RPE revealed a linear relationship with beta power in case of gain (reward) outcomes but not in losses. As explained in the previous sections, RPE is the main element of reinforcement learning models that indexes a signed deviation from the expected outcomes, therefore indexing a negative large value for unexpected losses as opposed to a positive large value for unexpected gains. When outcomes are expected RPE is smaller or zero. Whether or not beta indexes an RPE is crucial in terms of understanding its function in reinforcement learning.

Interestingly, Cunillera and colleagues (Cunillera et al., 2012) also found that the first correct feedback after a change of rule in a modified version of the Wisconsin card sort test (WCST) elicited an increase in the power of beta oscillations distributed over the frontal areas of the scalp. This finding was interpreted by authors as reflecting a neural signature associated with the processing of relevant and novel positive information that is required for learning and behavioral adjustment.

Lastly, using a version of the time estimation task, van de Vijver and colleagues (van de Vijver et al., 2011) showed an increase in beta power to feedback stimuli

indicating correct performance that also predicted the correct performance on the next trial. This beta activation was interpreted as a signal for the continuation of the current neurocognitive state after positive feedback.

The interpretations of the authors in the above studies are based on one or a combination of the following proposals. The common basis of these proposals is that beta oscillations serve to couple neural activity across distant brain areas. This is also supported by computational models that propose that beta oscillations are better than gamma oscillations for long-distance communication (Kopell, Ermentrout, Whittington, & Traub, 2000).

Maintenance of the status quo. Engel and Fries (2010) proposed a unifying hypothesis for the function of beta oscillations. According to this hypothesis, increases or decreases in the activity of beta oscillations are associated with involvement of top-down processes in signaling the need for maintenance of the current motor or cognitive set or the need to change, respectively. They further suggest that this signal may also reflect the contents of such top-down signal depending on the relevance of the new information: In tasks involving strong endogenous top-down control, the level of beta activity increases, whereas in tasks that involve more exogenous bottom-up factors, beta activity decreases. In this proposal, pathological beta activity in movement disorders like Parkinson's disease is interpreted as reflecting pathological engagement of a neural signal that maintains a motor state that is no longer appropriate, giving rise to the movement dysfunction. In sum, the unifying function of beta oscillations has been suggested to be the maintenance of the status quo of the cognitive or motor set (Engel & Fries, 2010).

The function of reward-related beta power has been interpreted according to the

status quo hypothesis as the tendency to remain in the current motor and cognitive state after reward receipt (Marco-Pallarés et al., 2015). However, a valid question would be how a motor or cognitive state is defined, what informational content is provided by such state, and how this might be related to learning from rewards.

In fact, Jenkinson and Brown (2011) suggested that the "status quo theory" has limited heuristic value because it does not explain beta activity in the context of other theories of BG function. Instead, they propose that level of beta activity in the BG-cortical system provides an index of the likelihood of the need for a novel voluntary action as a direct consequence of net dopamine levels at the site of cortical input to BG. According to this account, beta function is related to "anticipatory resourcing" such as motor readiness (Jenkinson & Brown, 2011).

Learning from positive feedback. Cohen and colleagues (2011) proposed a framework for feedback-guided learning based on oscillatory synchronization within and between different parts of the PFC. According to this framework, feedback learning occurs by synaptic modifications across PFC regions that link stimulus processing with response generation in a top-down fashion. This framework predicts a dissociated pattern for learning from negative and positive feedback: Learning from negative feedback engages frontal-midline theta activity whereas learning from positive feedback engages ventromedial (near the border of rostral ACC and medial OFC) beta activity. This theory also predicts the involvement of the DLPFC, fronto-polar regions, and the ACC depending on the task demands.

Therefore reward-related beta oscillations are interpreted as a mechanism for synchronizing activity across the PFC, especially between the OFC and motor cortex.

This framework proposes a general synchronization mechanism governed by reward-related beta oscillations across the frontal cortex but the direct sensitivity of these oscillations to specific task demands remains to be tested.

Interplay across multiple cognitive functions. In a recent review, Marco-Pallares and colleagues (2015) proposed that reward-related beta activity constitutes a brain signature of unexpected rewards reflecting transmission of a "fast motivational signal" to the reward processing network. On this view, the beta signal serves a need for an "interplay of attention mechanisms (novelty processing), reward processing and memory (storing action plans and learning for future episodes)". They further proposed the dorsal ACC or ventromedial PFC as the possible neural generators of beta. Then this fast motivational signal is transmitted to striatal areas and other parts of reward network such as hippocampus and amygdala to create coupling between attention, motivation, and memory systems.

As indicated above, the proposed sources of beta activity have mostly been speculative. In fact, the need for localizing the source of beta oscillations using the existing EEG source-localization techniques, joint analysis of EEG and functional magnetic resonance imaging (fMRI) data (Mas-Herrero, Ripollés, HajiHosseini, Rodríguez-Fornells, & Marco-Pallarés, 2015), and simultaneous EEG and fMRI recordings has been raised in recent reviews about oscillatory mechanisms for learning from feedback (Luft, 2014; Marco-Pallarés et al., 2015). Despite the absence of information about the neural source of frontal beta oscillations, their distribution over the frontal-central and frontal-lateral areas of the scalp that is consistently found across studies suggests the PFC and ACC as possible sources. Also, the elicitation of frontal

beta to reward feedback stimuli suggests that it can provide further insight into the reinforcement learning functions of the PFC and ACC.

Given the characteristics of frontal beta, such as its sensitivity to reward and its distribution over frontal areas of the scalp, the HRL-ACC theory provides a fertile theoretical framework for understanding the function and neural source of beta. Further, the previous literature on the role of beta oscillations in working memory and the role of PFC in this process suggest that frontal beta oscillations can be elucidated with a combination of the HRL framework, the oscillatory model for working memory, and the role of PFC in such processes. Also, the unique properties of brain oscillations in communication within and across brain areas might explain some of the ambiguous neural processes involved in these theoretical frameworks, such as the formation of options governed by the ACC or the neural representation of task demands by the DLPFC. Integration of oscillations into the existing theoretical frameworks could also be a resourceful contribution considering the rich information provided by brain oscillatory behavior that is conveniently captured by different types of electrophysiological recordings.

Research question

Although the studies and proposals that I reviewed above have interpreted reward-related beta activity as a means for synchronizing the activity of different brain areas, the neural genesis and associated functions of frontal beta are still ambiguous. Are beta oscillations indispensable for learning? What is the neural generator of reward-related beta power? Does manipulation of attention, motivation, or memory requirements

of a task affect the function and source of these oscillations? If we know the specific function of beta and its source, can it be used for therapeutic purposes?

In this dissertation, I have focused on one main question: What cognitive function do frontal beta oscillations reflect? My working hypothesis was inspired by the proposed role for the DLPFC and ACC in the HRL-ACC framework (Holroyd & Yeung, 2012), as well as on observations of the role of beta oscillations in forming task-related neuronal ensembles in the DLPFC (Buschman et al., 2012) and their sensitivity to working memory processes (Axmacher et al., 2010; Howard et al., 2003). Further evidence came from an unpublished experiment in which I asked participants to navigate a virtual T-maze task by trial-and-error with random and equiprobable (50%) reward and error feedback probabilities. This study failed to produce an increase in beta oscillations following rewards even though other behavioral and ERP components revealed normal sensitivity to reward vs. error feedback. This null result suggested that beta is not elicited when the mapping of responses to rewards is easy to establish (as in a simple T-maze task) and/or independent of behavior (for example, when subjects see that feedback is random).

Based on these observations, I propose that beta oscillations play a role in forming and updating action representations in the DLPFC that is triggered by rewarding events following action execution. On this view, beta oscillations signal the need for selective updating and maintenance of representations in DLPFC that are associated with successful actions, in other words, by boosting the neural representation of the actions immediately taken before the feedback. When the task rules are learned and neural representations established in other brain areas directly responsible for task performance,

beta is diminished. Further, if beta is related to working memory, then these considerations indicate that it must be also sensitive to working memory load associated with the number of actions immediately preceding feedback delivery. In particular, HRL-ACC theory predicts that the ACC will be involved in maintaining longer sequences of actions in working memory as an option-specific policy. Further, if beta is related to the 'task-related' information conveyed by outcomes on associating preceding actions with desired task performance, it could also be elicited to any type of feedback that provides task-related information worth remembering. For example, when actions associated with error outcomes are important to remember, error feedback stimuli should elicit beta.

Therefore, my main working hypothesis was that *frontal beta reflects the activation of neural ensembles related to the most recent action or sequence of actions. As a consequence of this activation, the frontal system applies a top-down control signal that transfers this information to other brain areas that are responsible for future task execution. Depending on the task requirements, such as the number of actions in the sequence prior to feedback presentation and the need to update task rules, this process would be jointly mediated by the DLPFC and ACC through a working memory maintenance and update mechanism. Feedback indicating useful task-related information associated with the preceding action/action sequence enhances the neural representation of this information in the DLPFC, which in turn applies a top-down control signal that communicates this information to other brain areas such as striatum for future task execution. Further, for tasks that involve a sequence of actions, ACC forms and maintains the representation of these sequences as options or task sets, and communicates these representations to the DLPFC to facilitate control over task*

performance.

To test this main hypothesis and its auxiliary predictions, I conducted a series of experiments where I tried to identify the behavior of frontal beta in response to task manipulations and its neural source. To clarify the role of reward-related beta oscillations in reinforcement learning, the first experiment examined whether beta codes for an RPE signal. Then I proceeded to test the hypothesis stated above in three subsequent experiments.

Experiment One (Chapter 2). *Does frontal beta index an RPE?*

According to my main hypothesis, beta is related to a more general working memory process elicited by rewards as opposed to a reinforcement learning signal. This experiment tests whether frontal beta is a reinforcement learning signal.

Reinforcement learning signals constitute the reinforcement learning terms in reinforcement learning algorithms (Sutton & Barto, 1998). For example, RPE signals relate to the temporal difference error term in temporal difference learning, a method for reinforcement learning, which can be utilized in various ways for performance optimization (Sutton & Barto, 1998). Given the extensive projections of dopamine neurons that carry RPE signals to the PFC and ACC (Holroyd & Coles, 2002), and that I previously found that frontal beta was sensitive to the probability of reward (HajiHosseini et al., 2012), it is important to rule out the possibility that beta reflects the impact of this important teaching signal on frontal brain areas. To investigate the relationship between beta power and RPE signals, I re-analyzed two existing datasets that employed a time-estimation paradigm with two levels of probability (high, low) and two levels of valence (reward, error) (Holroyd & Krigolson, 2007; Holroyd, Pakzad-Vaezi, & Krigolson,

2008). These data had been previously used to show the sensitivity of the reward positivity, an ERP component that is shown to index an RPE (Sambrook & Goslin, 2015), and frontal midline theta oscillations (HajiHosseini & Holroyd, 2013) to outcome probability. Therefore, if beta coded for an RPE, it should be sensitive to an interaction between probability and valence in these data as well.

Experiment Two (Chapter 3). *Source-localization of beta power in a reinforcement learning paradigm*

According to my hypothesis, frontal beta is related to a top-down control process mediated by the DLPFC and ACC. In this experiment, a virtual T-maze task with reinforcement was used with a rich EEG channel arrangement to elicit and source-localize frontal beta power elicited by reward-related feedback stimuli.

The EEG inverse problem refers to identifying the sources of the electrical activity that is recorded from the surface of the scalp (Pascual-marqui, 1999). Any particular distribution of voltages recorded at the scalp can be accounted for by an infinite number of combinations of possible current dipole sources. This arises from the fact that the problem is ill-posed mathematically: There are more unknown variables than equations. Nevertheless, although there is no unique solution to the inverse problem, several heuristic techniques have been developed for source-localizing EEG activity. Most existing algorithms are used for localizing the source of ERP components, so the localization of oscillatory power has been less commonly used. I used standardized low resolution brain electromagnetic tomography (LORETA) for source-localizing beta power (Pascual-Marqui, 2002). As a sanity check, I also localized the effect of probability on theta power, which has been shown in multiple studies to come from the

ACC (Asada, Fukuda, Tsunoda, Yamaguchi, & Tonoike, 1999; Ishii et al., 1999).

According to my hypothesis, reward-related beta oscillations reflect a working memory process mediated by the DLPFC. As it was reviewed in this chapter, observation of beta oscillations in the monkey DLPFC (Buschman et al., 2012), their association with working memory (Lisman & Idiart, 1995), and the theories on the association of working memory with DLPFC activity (Sakai, 2008) indicate a possible source for reward-related frontal beta in the DLPFC. To test this, I used a modified version of the virtual T-maze paradigm (Baker & Holroyd, 2009) with two levels of probability and valence, while recording the EEG using a relatively rich 57 channel arrangement to increase the spatial resolution of the EEG for the purpose of source localization.

Experiment Three (Chapter 4). *Sensitivity of reward-related beta oscillations to task demands on working memory*

According to my hypothesis, reward-related beta oscillations reflect a working memory process where the actions/action sequences immediately prior to a reward feedback are communicated and maintained through a neurocognitive mechanism mediated by the DLPFC and ACC according to task demands. This mechanism will place a higher load on working memory for tasks that require a sequence of actions as opposed to a single action. Based on the HRL-ACC theory, my hypothesis predicts that ACC would contribute to the maintenance of longer sequences. This experiment tests whether beta increases with the number of actions in the sequence preceding the feedback and, if so, localizes the source of this effect.

I tested these predictions by instructing participants to perform two different virtual maze tasks with reinforcement that differed according to the number of responses

required by the subject on each trial before the presentation of the feedback stimulus: One task required one response and the other task a sequence of three responses.

Experiment Four (Chapter 5). *Sensitivity of beta oscillations to informative value of outcomes irrespective of feedback valence*

According to my hypothesis, beta activity elicited to reward feedback stimuli signals a need for updating/maintaining the task-related information content of the actions immediately prior to the recent reward. This experiment tests whether frontal beta is sensitive to this 'task-related' information in general as opposed to only 'reward-related' information.

On view of my hypothesis, there is nothing special about reward feedback per se; for typical reinforcement learning tasks it simply carries more information about desired task performance than does error feedback, hence inducing greater beta power. By contrast, error feedback stimuli elicit less beta activity because it is less efficient to allocate neural resources for updating and maintaining information that is unrelated to desired task performance. To test this prediction, I manipulated the informative value of error feedback by instructing subjects to remember the actions that preceded error feedback vs. correct feedback for future recall to see whether these instructions would disrupt reward-related beta. Participants engaged in a card choice task involving a choice phase that was followed by a recall phase, creating a dual-task reinforcement learning/working memory paradigm. In one condition subjects were asked to recall actions that were followed with 5 cent reward feedback, and in another condition they were asked to recall actions followed by 0 cent error feedback. I predicted to find an association between post-feedback beta power and the task-related value of the outcomes

irrespective of their valence in each condition. I also predicted that post-feedback beta power would be correlated with the speed of recall.

Theoretical and clinical importance

This line of research contributes to the literature on the oscillatory signatures of reinforcement learning, cognitive control, and working memory. I empirically tested my prediction on the association of reward-related beta power with working memory, which in so doing tests some of the predictions made previously by others about this oscillatory phenomenon. The localization of reward-related beta activity and its compliance with a role in working memory also contributes to the current theories of the ACC, PFC, reinforcement learning, and working memory. Finally, given the pathological form of beta oscillations observed in neuropsychiatric disorders such as Parkinson's disease and schizophrenia, the results of the experiments presented here can inform the development of novel therapeutic treatments using brain stimulation techniques.

Chapter 2: Experiment One

Abstract

Reward feedback elicits a brief increase in power in the high-beta frequency range of the human EEG over frontal areas of the scalp, but the functional role of this oscillatory activity remains unclear. An observed sensitivity to reward expectation (HajiHosseini, Rodríguez-Fornells, & Marco-Pallarés, 2012) suggests that reward-related beta may index a reward prediction error (RPE) signal for reinforcement learning. To investigate this possibility I reanalyzed EEG data from two prior experiments that revealed RPEs in the human event-related brain potential (Holroyd & Krigolson, 2007; Holroyd, Pakzad-Vaezi, & Krigolson, 2008). I found that feedback stimuli that indicated reward, when compared to feedback stimuli that indicated no-reward, elicited relatively more beta power (20-30 Hz) over a frontal area of the scalp. However, beta power was not sensitive to feedback probability. These results indicate that reward-related beta does not index an RPE but rather relates to a different reward processing function.

Introduction

Several studies have reported that presentation of reward-related feedback stimuli enhances power in the high-beta frequency range of the human EEG (Cohen, Elger, & Ranganath, 2007; HajiHosseini, Rodríguez-Fornells, & Marco-Pallarés, 2012; Marco-Pallares et al., 2008) and MEG (Doñamayor et al., 2011) over frontal areas of the scalp. Although recent proposals have suggested that reward-related beta activity might reflect coupling between neurocognitive processes involved in motivation, attention, and memory (Marco-Pallarés et al., 2015), or neural synchronization that facilitates learning from feedback (Luft, 2014), there is a paucity of data addressing this question. Of note, in one recent study unexpected gains in a gambling paradigm elicited relatively more beta power compared to expected gains (HajiHosseini et al., 2012). This sensitivity of beta power to reward expectancy suggests that beta oscillations might index an RPE, an important training signal in computational theories of reinforcement learning that indicates whether ongoing events are "better" or "worse" than expected (Caplin & Dean, 2008; Sutton & Barto, 1998). Consistent with this possibility, substantial evidence indicates that RPEs are carried by the midbrain dopamine system to their neural targets (Schultz et al., 1997) including frontal areas of cortex (Holroyd & Coles, 2002).

A neural signal that encodes an RPE must be sensitive to a specific interaction between the valence and probability of the eliciting outcome (Sambrook & Goslin, 2015). If beta reflected an RPE signal, then I would expect relatively more beta following unexpected rewards compared to expected rewards, and relatively less beta following unexpected errors compared to expected errors. To investigate whether beta has these properties, I reanalyzed data from two previous EEG experiments that revealed an RPE

signal in the human ERP (Holroyd & Krigolson, 2007; Holroyd et al., 2008). In these experiments, subjects engaged in a time-estimation task in which the correct (rewarded) and incorrect (not rewarded) responses occurred with high or low probability, as determined by a staircase procedure that adjusted task difficulty from trial to trial (see Method). I reasoned that if reward-related beta reflects an RPE signal, then that property should be observed in a task already known to elicit an RPE signal in the ERP.

Methods

EEG datasets were reanalyzed from two previous studies: Dataset 1 (D1) from Holroyd and Krigolson (2007) and Dataset 2 (D2) from Holroyd and colleagues (2008). In both studies the EEG was recorded from participants while they performed a time-estimation task. Because the studies were carried out using nearly identical protocols, the data were reanalysed together (with *dataset* as a between-subject factor) to increase statistical power. The studies were conducted in accordance with the ethical standards prescribed in the Declaration of Helsinki and were approved by the human subjects review board at the University of Victoria. Informed written consent was obtained from all participants prior to the experiment.

Participants

D1 and D2 included the data of seventeen (8 men; 19.6 ± 2.8 years old) and twelve (6 men, 26.7 ± 10.5 years old) participants, respectively, who were undergraduate students at the University of Victoria receiving extra course credits for their participation, or who were paid volunteers. The data of two of the participants associated with D1 were eliminated from the analysis because of an insufficient number of trials following artifact

rejection. Therefore I analyzed the data of a total of twenty-seven participants across both datasets.

Task

In both studies, on each trial participants were required to press a left mouse button when they estimated that 1 s had elapsed following presentation of an auditory cue. At the end of each trial, a visual feedback stimulus indicated whether the response was correct or incorrect. The feedback stimuli consisted of a white plus sign and a white zero (3^{o}, 1000 ms) presented on a high contrast black background. The response was initially evaluated as correct if it was produced within a time window spanning 900-1100 ms following cue onset, and was evaluated as incorrect otherwise. The width of the time window varied from trial to trial by condition according to the following stair case procedure. In the *control* condition, the window size increased by 10 ms following every error response and decreased by 10 ms following every correct response. In the *hard* condition, the window size increased by 3 ms after every error response and decreased by 12 ms after every correct response, and in the *easy* condition, the window size increased by 12 ms after every error response and decreased by 3 ms after every correct response. Participants in D1 completed six blocks of 75 trials: two in the control, two in the hard, and two in the easy conditions. Participants in D2 did five blocks of 100 trials: one in the control, two in the hard, and two in the easy conditions. Participants were told at the start of the experiment that some conditions would be harder than others. Because the order of the control condition was not counterbalanced with the other conditions in either study (the control condition always occurred first), the reanalysis included only trials associated with the easy and hard conditions (300 trials in D1 and 400 trials in D2). Note that the

order of the easy and hard conditions was counterbalanced across subjects for both datasets. For a complete description of the task, please refer to Holroyd and Krigolson (2007).

Data Acquisition

The EEG was recorded from 41 electrode locations in D1 (Holroyd & Krigolson, 2007) and 64 electrode locations in D2 (Holroyd et al., 2008) using BrainVision Recorder software (Brainproducts, Munich, Germany). Electrodes were arranged according to the standard 10-20 layout (Jasper, 1958) and were referenced to the average voltage across the channels. The two electrode montages were overlapping but not identical. Vertical and horizontal ocular movements were recorded by an electrode placed under the right eye (re-referenced offline to FP2), and two on the outer canthi of the right and left eyes (re-referenced offline to each other), respectively. Electrode impedances were kept under 10 kΩ. Data were sampled at 250 Hz and band pass filtered by the amplifiers at 0.017 Hz–67.5 Hz.

Data Analysis

Data pre-processing was performed in BrainVision Analyzer 2. A band-pass filter (0.1-40 Hz) was applied to the EEG data and epochs of EEG activity were selected from 1 s before to 1 s after the onset of feedback stimuli. Data were subsequently re-referenced to the average value recorded at the mastoids. Ocular correction was performed using the Gratton, Coles, & Donchin (1983) algorithm as implemented in the Analyzer software. Feedback segments were baseline-corrected by subtracting, for each channel, subject, and electrode, the average voltage values during the 100 ms prior to the feedback stimulus from the subsequent voltages in the epoch. EEG artifacts were identified and rejected

according to the following criteria: Any abrupt change of voltage greater than 35 μV from one time sample to the next, any difference between the negative and positive peaks in a 200 ms interval that exceeded 150 μV, and any activity that was consistently smaller than 0.5 μV in a 100 ms interval were considered artifacts and the corresponding segment was rejected for all channels. On average, 25% of data were rejected[1]. Data were exported to MATLAB for time-frequency analyses. Topographical scalp maps were plotted with EEGLAB (Delorme & Makeig, 2004).

To extract time-frequency information, a 2 s epoch centered on the feedback presentation time was convoluted with a seven-cycle complex Morlet wavelet that was linearly scaled based on the frequency range of 1-40 Hz. The time-frequency power was extracted relative to a 100 ms baseline before the feedback on each trial as (trial power-baseline power)/baseline power. Then the power values were averaged separately for each condition and subject. I expected the beta power contrast to occur between 20-30 Hz at about 250 ms to 500 ms following the feedback stimulus according to previous studies (Cunillera et al., 2012; HajiHosseini et al., 2012; Marco-Pallarés et al., 2008). Therefore, the grand average time-frequency power maps were inspected according to this a priori assumption (Cohen, 2014) and the time window when beta power reached maximum was determined, from 350-500 ms post-feedback. Beta power was therefore measured as the average value within this interval for all channels.

For the purpose of statistical analysis the beta power values were submitted to ANOVA using dataset as a between-subject factor. The data were evaluated at a subset of nine representative electrode locations that were common to the two datasets (see

[1] This rejection rate is higher than those reported in the original articles due to application of a higher filter cutoff (40 Hz), which retained high-frequency noise as well as signal.

asterisks in Figure 5 A and B). To be specific, I conducted a 2×9×2×2 mixed ANOVA on beta power with a between-subject factor of dataset (D1, D2) and within-subject factors of channel (F3, FZ, F4, C3, CZ, C4, P3, PZ, and P4), feedback valence (reward, error), and feedback probability (high, low). For the purpose of display, the scalp distributions of beta power were also calculated separately for both datasets using the full electrode configurations. Last, beta power was pooled across datasets for the channel that showed the maximal difference in beta power across the reward and no-reward conditions (see below). The Greenhouse-Geisser correction was applied as appropriate to correct violations of the assumption of sphericity (Picton et al., 2000).

Results

A 2×9×2×2 mixed ANOVA on beta power with a between-subject factor of dataset (D1, D2) and within-subject factors of channel (F3, FZ, F4, C3, CZ, C4, P3, PZ, and P4), feedback valence (reward, error), and feedback probability (high, low) revealed a significant main effect of valence ($F(1,25) = 5.82$, $p = 0.023$) and a significant interaction between valence and channel ($F(8,200) = 2.52$, $p = 0.05$); no other main effects or interactions were statistically significant. Within the set of nine electrode locations under consideration, the difference in valence was largest at the right-frontal channel F4. Post-hoc t-tests on the reward vs. error conditions (collapsed across probability and pooled across both experiments) showed greater power at channel F4 compared to P4 (t (26) = 2.64, $p = 0.014$) and channel C4 compared to P4 (t (26) = 2.40, $p = 0.024$). Differences between left and right lateral channels (F3 and F4, C3 and C4, P3 and P4) were not significant ($p > 0.05$). Figure 5 A and B illustrate the scalp distributions

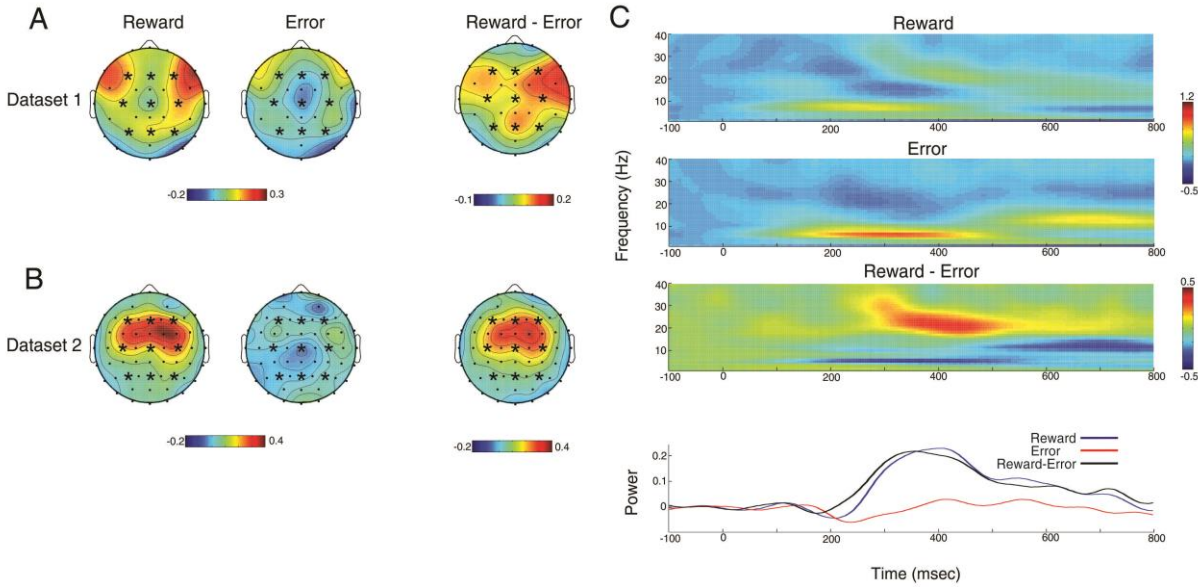

Figure 5: Time-frequency data for the time-estimation task. Scalp distribution of beta (20-30 Hz) power for the reward (left), error (middle), and reward-error (right) conditions for A) dataset 1 (D1) and B) dataset 2 (D2). Asterisks show the subset of channel locations that were statistically analyzed. C) From top to bottom: time-frequency maps of power for the reward, error, reward-error conditions, and time-course of average beta power (20-30 Hz) in reward (blue), error (red), and reward-error (black) conditions, pooled across the two datasets at channel F4. Note that power is unit-less because it is calculated as a proportional increase/decrease relative to average baseline power.

of beta power for both datasets, separately for the reward and error conditions (averaged across the high and low probability conditions) and the difference in beta power across the two conditions. As can be seen by inspection, the difference in power reached a maximum over right-frontal areas of the scalp for both datasets, at channel FC6 for D1 and channel FC2 for D2. For the purpose of illustration, I pooled the data across experiments at channel F4, which was the channel nearest to FC6 and FC2 that was common across D1 and D2. Figure 5C illustrates the time-frequency response at channel F4 pooled across the two datasets; panels from top to bottom show the reward, error,

reward - error time-frequency maps, and the time course of beta power for the reward, error, and reward - error conditions.

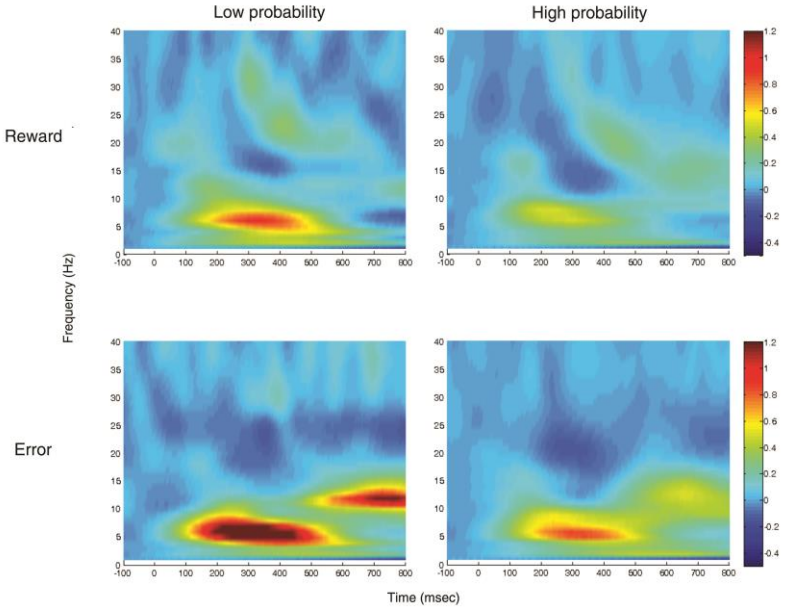

Figure 6: Time-frequency data at channel F4 for low probability reward, high probability reward, low probability error, and high probability error conditions.

Note that no effect of probability or interaction with probability was found. I was concerned that the failure to detect such an effect might be due to low statistical power. Therefore, as a check I collapsed the data across the two datasets and across valence and tested beta power for the low probability condition against the high probability condition independently for each of the 9 channels of interest. This exploratory analysis failed to reveal any effect of probability ($p > 0.05$). Further, a series of 2×2 ANOVAs with probability (high, low) and valence (reward, error) on beta power, conducted separately at each of the 9 channels and collapsed across datasets did not show any interaction of valence and probability at any channel ($p > 0.05$)). These exploratory analyses confirm the absence of strong probability effects on beta power in these datasets. (See also Figure

6 for time-frequency maps of beta power associated with channel F4 for all four feedback conditions).[2]

Discussion

Reward feedback in gambling or guessing tasks has been observed to elicit beta power over frontal areas of the scalp (Marco-Pallarés et al., 2008) that is sensitive to outcome probability (HajiHosseini et al., 2012). Suggestively, these EEG oscillations, which occur from about 250-500 ms post-feedback, immediately follow the production of a frontal-centrally distributed ERP component called the reward positivity, which occurs from about 200-350 ms following feedback onset (Miltner, Braun, & Coles, 1997; Proudfit, 2015). The reward positivity is said to index an RPE (Sambrook & Goslin, 2015), a computationally powerful reinforcement learning signal that encodes an interaction between outcome valence and probability (Sutton & Barto, 1998). These observations raise the question of whether or not frontal beta is also related to RPEs. Consistent with the previous studies, I found that reward feedback stimuli compared to no-reward feedback stimuli elicited greater power in the beta frequency range over frontal regions of the scalp at about 350-500 post-feedback. However, beta power was not modulated by feedback probability, nor was it sensitive to the interaction of valence of and probability, indicating that beta power does not code for an RPE. The fact that this task clearly elicits 1) a valence effect on beta power, 2) no effect of probability on beta

[2] As reported previously, in D1 participants made more errors in the hard condition (76%) than in the easy condition (23%) ($t(16) = -41.39$, $p < .001$), consistent with the mean size of the response window, which was smaller in the hard condition (52 ms) than in the easy condition (160 ms), ($t(16) = -11.39$, $p < .001$) (Holroyd & Krigolson, 2007). In D2, participants made more errors in the hard condition (75.4%) than in the easy condition (24.5%), ($t(11) = -35.4$, $p < .001$), consistent with the mean size of the performance window, which was smaller in the hard condition (128 ms) than in the easy condition (334 ms), ($t(11) = -8.5$, $p < .001$) (Holroyd et al., 2008). See also (HajiHosseini & Holroyd, 2013) for an analysis of theta oscillations related to D2. Here I only report results related to beta oscillations.

power, even when testing each channel independently, 3) an RPE in the time domain (Holroyd & Krigolson, 2007; Holroyd et al., 2008), and 4) a probability effect in the theta frequency range (HajiHosseini & Holroyd, 2013), indicate that the task produces electrophysiological effects related to feedback probability and that the failure to observe these properties in the beta frequency range does not result from low statistical power. This result is especially peculiar given that many neural processes (such as the classic "orienting response") show expectancy effects of one kind or another (e.g., Nieuwenhuis, De Geus, & Aston-Jones, 2011); it remains to be determined whether beta power can be modulated by probability in other task paradigms.

Speaking to this issue, I previously observed that feedback probability in fact modulated beta power (HajiHosseini et al., 2012). In that experiment, on each trial participants were presented with an initial cue that indicated the magnitude and probability of a potential reward at the end of the trial. Post-feedback beta power was sensitive to valence (greater for gains than losses), probability (greater for unexpected outcomes than expected outcomes) and the interaction between valence and probability (HajiHosseini et al., 2012). In this study greater beta power was elicited by unexpected gains relative to expected gains. This discrepancy between the studies could be due to differences in task design that may have induced different subjective perceptions of outcome probability. In particular, on each trial in the task used in the previous study (HajiHosseini et al., 2012), an initial cue informed participants about the exact probability of the upcoming outcome. By contrast, in the time-estimation task the probabilities were learned by exposure to the feedback contingencies across blocks of trials, separately for the easy and hard conditions. The feedback stimuli in the study by

HajiHosseini and colleagues (2012) also provided information about the non-selected choices, possibly eliciting "fictive learning signals" (Lohrenz, McCabe, Camerer, & Montague, 2007) – i.e., information about the appropriateness of the unselected action. Beta power might therefore be more sensitive to valence in trial-and-error learning tasks that require ongoing adjustments of behavior (such as the time-estimation task), and more sensitive to probability in gambling tasks where stimulus cues explicitly indicate the reward probabilities (such as in HajiHosseini et al., 2012).

Finally, in both datasets the maximum power difference occurred over right frontal-central areas of the scalp (Figure 5A and B). Given that the task design was essentially identical across the two experiments, it is somewhat surprising that for the two studies the valence effect was largest at slightly different right-frontal scalp locations. Importantly, this right-frontal distribution is replicated in Experiment Two (Chapter 3) -- suggesting that effect is real and may result from a generator in right prefrontal cortex (see also Cunillera et al., 2012). I therefore suspect that the difference in the scalp distributions observed across experiments here might be due to the differences in the channel arrangements, to variability across population samples, or to statistical noise.

Taken together, these results indicate that reward-related beta does not code for an RPE signal, a process often associated with the function of medial frontal cortex (Holroyd & Coles, 2002). Further, the right-frontal scalp distribution is suggestive of genesis in right frontal cortex and is therefore not directly related to the reward processing functions of medial frontal cortex (Holroyd & McClure, 2015). It remains to be determined whether the signal reflects a different aspect of reinforcement learning mediated by lateral frontal cortex.

Chapter 3: Experiment Two

Abstract

Reward-related feedback stimuli have been observed to elicit a burst of power in the beta frequency range over frontal areas of the human scalp. Recent discussions have suggested possible neural sources for this activity but there is a paucity of empirical evidence on the question. In this chapter, I recorded EEG from participants while they navigated a virtual T-maze to find monetary rewards. Consistent with previous studies, I found that the reward feedback stimuli elicited an increase in beta power (20-30 Hz) over a right-frontal area of the scalp. Source analysis indicated that this signal was produced in right DLPFC. These findings align with previous observations of reward-related beta oscillations in the DLPFC in non-human primates. I speculate that increased power in the beta frequency range following reward receipt reflects the activation of task-related neural assemblies that encode the stimulus-response mapping in working memory.

Introduction

Neural oscillations in the ongoing EEG are believed to reflect the synchronous activity of distributed neuronal cell assemblies that encode distinct neurocognitive functions (Buzsáki & Draguhn, 2004; Wang, 2010). In particular, several studies have reported that reward-related feedback stimuli elicit increased power in the high-beta frequency range (20-35 Hz) in the human EEG and MEG over frontal areas of the scalp (Cohen, Elger, & Ranganath, 2007; Doñamayor, Marco-Pallarés, Heldmann, Schoenfeld, & Münte, 2011; Marco-Pallares et al., 2008). Enhanced frontal beta power is also elicited by unexpected reward feedback stimuli compared to expected reward feedback stimuli (HajiHosseini et al., 2012) and by the first positive feedback stimulus compared to subsequent positive feedback in the WCST (Cunillera et al., 2012). Consistent with these findings, it has recently been proposed that reward-related beta serves to couple attentional and emotional systems associated with novelty and reward processing (Marco-Pallarés et al., 2015), and that beta oscillations play a role in synchronizing neural activity to promote learning from positive feedback (Cohen, Wilmes, & van de Vijver, 2011; Luft, 2014). However, the specific role of these oscillations in reward processing is still poorly understood.

Insight into the functionality of these oscillations could be derived from identifying their neural origin (Luft, 2014; Marco-Pallarés et al., 2015). In Chapter 2, I showed that reward-related feedback stimuli elicit an increase in beta power over right-frontal areas of the human scalp, and speculated that this signal could be produced by right DLPFC (HajiHosseini & Holroyd, 2015b). Consistent with this possibility, studies in non-human primates have revealed beta oscillations in the principal sulcus, a

homologue of human DLPFC that is associated with rule implementation and category learning (Antzoulatos & Miller, 2014; Buschman et al., 2012; Hoshi et al., 2000). However, whether human DLPFC produces beta oscillations is unknown.

Here I recorded the EEG from participants engaged in a reinforcement learning task in which they navigated a virtual T-maze to find monetary rewards, and applied a source localization technique to investigate possible generators of reward-related beta oscillations. I predicted that rewards compared to errors would elicit higher beta power over frontal areas of the scalp and that this contrast would be localized to the DLPFC.

Methods

Participants

Twenty-six undergraduate students (7 men, 20.3 ± 3.8 years old) at the University of Victoria participated in the experiment. Subjects acquired extra course credits for participation and were also paid a monetary bonus that depended on task performance. The study was conducted in accordance with the ethical standards prescribed in the Declaration of Helsinki and was approved by the human subjects review board at the University of Victoria.

Task

Participants performed a version of a virtual T-Maze task used previously to investigate reward-related electrophysiological activity (Baker & Holroyd, 2009), modified according to probabilistic stimulus-reward contingencies derived from Holroyd and colleagues (2009). Note that a previous experiment in which participant responses were rewarded at random on 50% of the trials failed to produce reward-related beta

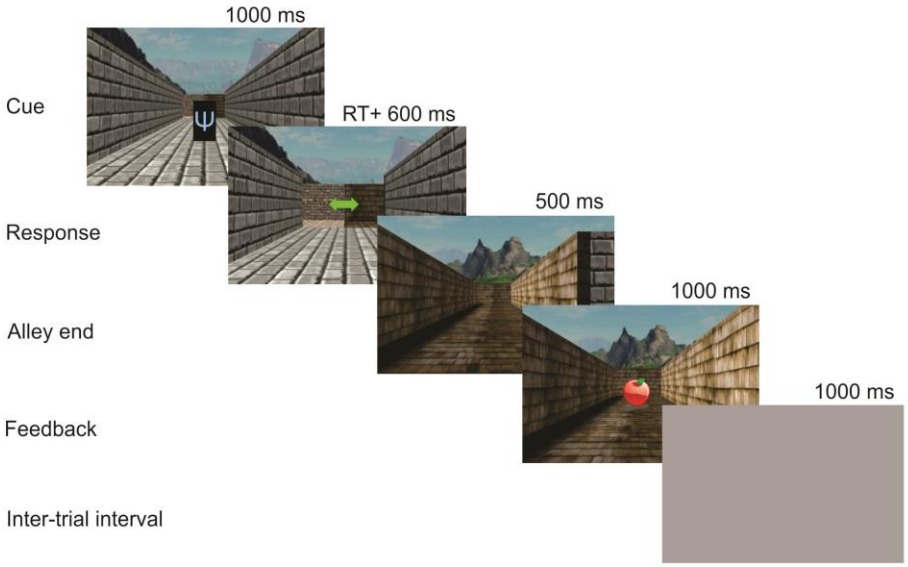

Figure 7: Virtual T-maze task. Participants were instructed to navigate a virtual maze by choosing a left or right response according to visual cues presented at the start of each trial. The cue was presented over an image of the stem alley for 1000 ms ("Cue"). Then, an image of a double arrow appeared on the screen and remained on the screen until 600 ms after a response was selected ("Response"). A view of the selected alley was then presented for 500 ms ("Alley end"), followed by a closer view of the end of the alley with an image of the feedback stimulus (apple or orange) overlaid at central fixation ("Feedback"), indicating that participants earned either 5 cents (reward) or 0 cents (error). A blank screen was presented for1000 ms between trials ("Inter-trial interval"). See methods for the probability mappings between cues, responses and feedback stimuli.

oscillations. I therefore modified the task such that the feedback depended probabilistically on prior stimuli and responses, in the expectation that beta power would be enhanced by an increase in perceived control over the trial outcomes. Subjects were instructed to navigate the virtual T-maze according to visual cues presented at the start of each trial. Figure 7 illustrates the event timing for an example trial in the task. At the beginning of each trial, a visual cue belonging to one of several categories (described below) was presented over an image of the stem alley. To convey a sense of movement, 1000 ms later the stimuli were replaced by an image that showed a closer view of the end

of the alley superimposed by a double-arrow (Figure 7). Upon seeing the arrow participants were instructed to select the right or left alley by pressing the corresponding arrow key on the keyboard. To limit the overall duration of the experiment (as opposed to pressing the participants for speed), responses that exceeded 1s were penalized with a 25 cents loss. Participants were not informed about the specific deadline but were instructed that slow responses would result in the loss. 600 ms after the response, an image of the chosen alley was presented for 500 ms, followed by a closer view of the end of the alley, overlaid with an image of the feedback stimulus (5.5^{o} of visual angle) at central fixation, presented for 1000 ms. Participants were told that that if they found an apple (orange) then they gained 5 cents on that trial and if they found an orange (apple) then they gained 0 cents on that trial; the assignation of reward values to feedback stimuli was counterbalanced across participants.

The task consisted of three blocks of trials, each characterized by a different set of four possible shapes for the initial cue. These three stimulus sets consisted of four geometrical shapes (square, triangle, circle, and trapezoid), four black squares depicting letters from the Greek alphabet (β, π, ψ, and Σ), and four cartoon sky-related shapes (sun, moon, star, and cloud). On each trial the cue was randomly chosen without replacement from the set of four. To prevent against the development of irrelevant stimulus associations, the stimulus colors differed across stimuli both within and across blocks. Within each block, each of the four shapes corresponded to a specific alley-probability combination, determined at random: 70% reward probability for right alley choices, 70% reward probability for left alley choices, 30% reward probability for right alley choices, and 30% reward probability for left alley choices. The opposite alley in all four stimulus

conditions was never rewarded (0% probability of reward). Thus, for each cue only one alley was rewarded and the probability of reward on that alley was either low or high, resulting in a two-by-two task design with levels for valence (reward, error) and probability (low, high).

Data acquisition

The EEG was recorded from 51 electrode locations using BrainVision Recorder software. Electrodes were arranged according to the standard 10-20 layout (Jasper, 1958) and were referenced online to the average voltage across the channels. Vertical and horizontal ocular movements were recorded by an electrode placed under the right eye (re-referenced offline to FP2), and two on the outer canthi of the right and left eyes (re-referenced offline to each other) respectively. Electrode impedances were kept under 10 kΩ. Data were sampled at 500 Hz and high pass filtered online at 0.017 Hz.

Data analysis

Data pre-processing was performed in BrainVision Analyzer 2. A band-pass filter (0.1-100 Hz) was applied to the EEG data and epochs of EEG activity were selected from 1s before to 1 s after the onset of feedback stimuli. Data were subsequently re-referenced to the average value recorded at the mastoids. Ocular correction was performed using the Gratton, Coles, and Donchin (1983) algorithm as implemented in the Analyzer software. Feedback segments were baseline-corrected by subtracting the average voltage values during the 100 ms prior to the feedback stimulus from the value of each sample in the epoch, for each channel, subject, and electrode. EEG artifacts were identified and rejected according to the following criteria: Any abrupt change of voltage greater than 35 μV across consecutive samples, any difference between the negative and positive peaks in a

200 ms interval that exceeded 150 μV, and any activity that was consistently smaller than 0.5 μV in a 100 ms interval were considered artifacts and the corresponding segment was rejected for all channels. On average, 7% of data were discarded. Data were then exported to MATLAB for the ERP and time-frequency analyses. Topographical scalp maps were plotted with EEGLAB (Delorme & Makeig, 2004).

To extract time-frequency information, for each subject, trial and channel, a two-second epoch centered on the time of feedback presentation was convoluted with a seven-cycle complex Morlet wavelet. The wavelet was linearly scaled based on the frequency range of 1-50 Hz and the power for each frequency band was evaluated relative to the 100 ms baseline before feedback onset as 10*log10 (trial power/average baseline power). Power values were averaged across trials for every channel, condition and subject. In line with Chapter 2, I investigated the effect of valence and probability on beta power and the distribution of these effects over the scalp for a subset of 9 representative electrode locations. To be specific, a 9×2×2 ANOVA was applied to beta power averaged over the 250-450 ms post-feedback interval with channel (F5, FZ, F6, C5, CZ, C6, P5, PZ, P6), valence (reward, error), and probability (high, low) as factors. Based on visual inspection, beta power was averaged within the 20-30 Hz range.

Source localization was performed with sLORETA (Pascual-Marqui, 2002). For each subject, channel, and trial, a 2-second data segment spanning 1 s before feedback onset to 1 s after feedback onset was analyzed for time-varying cross-spectra in sLORETA with a 72 sample-long Gaussian window for the beta frequency range (20-30 Hz). Note that source localization cannot be conducted directly on power values, which are related to the square of the voltage values. Therefore, sLORETA brain maps were

determined by recalculating the time-varying cross-spectral power values in the mid-frequency, 25 Hz, for each subject and condition. Statistically significant differences in beta power values were identified for each contrast by conducting paired t-tests for each voxel; the voxels with the largest t-values are reported. Randomization via statistical non-parametric mapping was applied in sLORETA to correct for multiple comparisons.
All error terms reported for the behavioral data constitute standard deviations.

Results

Behavioral analysis

Participants selected the rewarding arm of the maze on 68.2 ±0.1% of the trials overall, on 74.3 ±0.1% of the trials in the easy condition (high probability of reward, low probability of error), and on 62.0% ±0.1% of the trials in the hard condition (low probability of reward, high probability of error). The number of visits to rewarding arms was significantly higher for the easy condition relative to the hard condition ($t(25) = 6.20$, $p < 0.001$). Average reaction time was 275 ±85 ms across conditions, with no statistically significant difference between the conditions. Note that the participants were not permitted to respond until the appearance of the response cue, 1000 ms following the onset of the stimulus cue, which likely accounts for the uniformity of the reaction times across the stimulus conditions.

Time-frequency analysis

A 9×2×2 ANOVA on beta power with channel (F5, FZ, F6, C5, CZ, C6, P5, PZ, P6), valence (reward, error), and probability (high, low) as factors revealed a significant effect of valence ($F(1,25) = 18.25$, $p < 0.001$), a significant interaction of channel and

valence ($F(8,200) = 2.91$, $p = 0.023$), and no other main effects or interactions. Figure 8A illustrates the scalp distribution of beta power for the reward and error conditions and for the difference between the two conditions; note that the difference was distributed over right-frontal areas of the scalp, reaching a maximum value at channel F6. A 2x2 ANOVA on beta power associated with channel F6 confirmed the effect of valence ($F(1,25) = 32.51$, $p < .001$) and no effect of probability or interactions with probability and valence at that channel. Figure 8B presents time-frequency maps of power for reward and error conditions, and their difference, associated with channel F6.

Source localization

Source analysis was applied to the observed valence effect on beta (see methods). Figure 8C illustrates the location of the maximum t-value for the valence-related effect of beta power ($t = 5.84$, $p = 0.001$; X = 35, Y = 25, Z = 40, MNI coordinates),

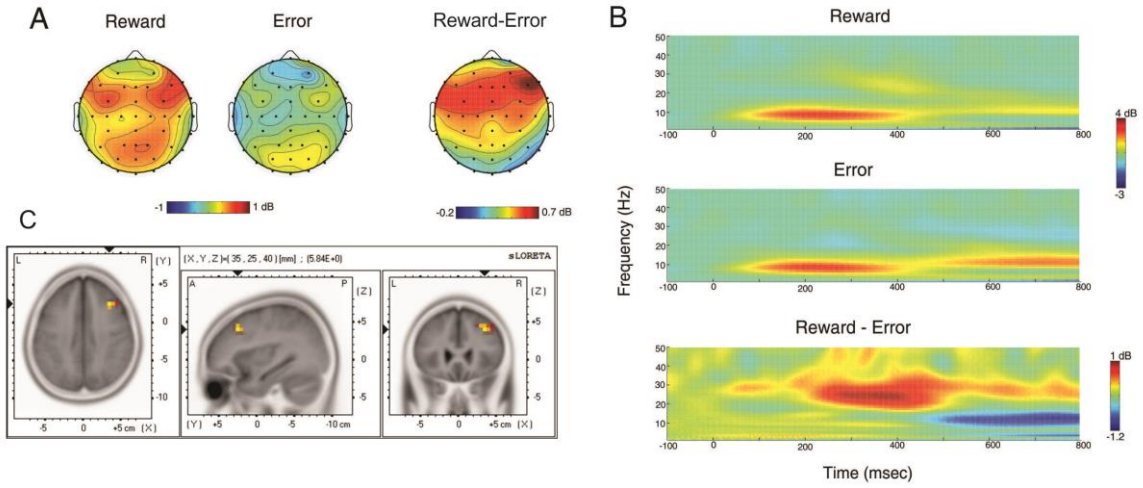

Figure 8: Time-frequency and sLORETA source localization results. A) scalp distribution of beta (20-30 Hz) power in reward (left), error (middle), and reward-error conditions. B) time-frequency maps of power for the reward (top), error (middle), and reward-error (bottom) conditions at channel F6. C) Right dorsolateral prefrontal cortex was revealed as the source of beta (25 Hz) contrast, reward vs. error, averaged over 250-450 ms post-feedback.

corresponding to Brodmann area 9 in the middle frontal gyrus of right DLPFC[3].

Discussion

Frontal beta oscillations are elicited by reward-related feedback stimuli, but not by non-reward or error-related feedback stimuli, over frontal areas of the scalp (Cohen et al., 2007; HajiHosseini et al., 2012; Marco-Pallares et al., 2008). Current proposals suggest that reward-related beta is generated within dorsal anterior cingulate cortex (Luft, 2014; Marco-Pallarés et al., 2015) or ventromedial prefrontal cortex (Marco-Pallarés et al., 2015). However, recent investigations have indicated a focus over right prefrontal areas (Cunillera et al., 2012; HajiHosseini & Holroyd, 2015b; HajiHosseini et al., 2012), suggesting that reward beta might originate from right prefrontal cortex. Here I verified this supposition with what is to my knowledge the first empirical investigation on source-localization of reward-related beta oscillations in humans. The results replicated the previously observed sensitivity of frontal beta to valence and further implicate right DLPFC as the neural generator, as predicted.

Current theories of reward-related beta oscillations have variously suggested that the signal might reflect a neural mechanism for learning from feedback (Luft, 2014), for synchronizing neural activity to promote learning from positive feedback (Cohen et al., 2011), and for coupling systems involved in memory, attention, and motivation (Marco-Pallarés et al., 2015). Further, an influential theory of working memory proposes that

[3] As a check, I also calculated the sLORETA values for theta power by contrasting the high- and low-probability conditions. Consistent with previous findings (Asada, Fukuda, Tsunoda, Yamaguchi, & Tonoike, 1999; Ishii et al., 1999), the maximum statistically significant difference (t = -4.97, p = 0.02; X = -10, Y = 30, Z = 30, MNI coordinates) was localized to Brodmann area 32 in the cingulate gyrus of the dorsal ACC. Given that ACC is a known source of theta power (for a review, see Cavanagh & Frank, 2014), these results lend confidence in the accuracy of our localization of the beta signal.

maintenance of individual items in working memory is mediated by interacting beta and theta oscillations (Lisman & Idiart, 1995). This theory has been supported by observations that beta-gamma oscillations in the human frontal cortex and hippocampus scale with working memory load (Howard et al., 2003) and couple with oscillations in theta range as predicted by the theory (Axmacher et al., 2010). In view of the well-known role of human DLPFC in maintaining task-related information in working memory (Barbey, Koenigs, & Grafman, 2013; D'Ardenne et al., 2012; Funahashi, 2006), the current results suggest that beta oscillations mediate a link between DLPFC processes related to reward learning and working memory.

In line with these observations, neurons on the banks of the monkey principal sulcus, a homologue of human DLPFC, are active in a rule-specific manner depending on task requirements (Hoshi et al., 2000), and code for the currently relevant task rule by synchronizing in the beta frequency range (Buschman et al., 2012). Synchrony in the beta frequency range between monkey striatal and PFC neurons also increases during category learning (Antzoulatos & Miller, 2014), suggesting that beta oscillations may facilitate communication between the PFC and striatum during such learning.

In the context of this literature, I speculate that increased power in the beta frequency range following reward receipt reflects enhanced activation of task-related neural assemblies that encode the stimulus-response mapping for that trial (Miller & Cohen, 2001). On this view, the synchronous activity at the beta frequency range of neurons in DLPFC and the striatum would facilitate the transfer of rewarded action sequences to other brain areas (Antzoulatos & Miller, 2014; Frank et al., 2001; Frank, 2005). Once learned, these sequences could be executed automatically, reducing the need

for communication of task demands placed on the DLPFC (Cohen, Dunbar, & McClelland, 1990; Cunillera et al., 2012; Todd, Niv, & Cohen, 2008), a process that would complement other proposed mechanisms for integrating working memory with reinforcement learning (Collins & Frank, 2012; Todd et al., 2008). This hypothesis could be tested by disrupting or enhancing reward-related beta oscillations in human DLPFC using non-invasive stimulation techniques such as transcranial magnetic stimulation or transcranial direct current stimulation (Knoch et al., 2006; Reinhart & Woodman, 2014).

Chapter 4: Experiment Three

Abstract

Several influential theories have strongly implicated the PFC and ACC in cognitive processes related to cognitive control, decision making, and working memory, but the particulars about how these functions are carried out remain ambiguous. Recently, frontal beta oscillations recorded by the EEG have been observed to reward feedback stimuli in a variety of reinforcement learning paradigms. In Chapter 3, I localized these reward-related beta oscillations to the DLPFC and proposed they reflect a working memory mechanism mediated by the DLPFC for transfer of the recent successful stimulus-response rules to other brain areas that are responsible for task execusion. Here I tested whether frontal beta is sensitive to the working memory aspects of the task such as the number of actions in the sequence immediately preceding the feedback stimuli. I had subjects perfom two virtual maze tasks with reinforcement where one of the tasks required a single response and the other three separate responses before the presentation of feedback stimuli. I found that reward feedback elicited greater beta power compared to error feedback and the maze with a longer sequence of actions elicited greater beta power compared to the maze with a single action while both effects were distributed over the frontal areas. There was also coupling of beta power to theta phase following the reward feedback in the maze with a longer sequence of actions. Further, source analysis revealed the left DLPFC and ACC as generators of the valence and maze contrasts in beta power. These findings show that, in addition to valence, frontal beta power is sensitive to the working memory demands of the task. I propose that beta oscillations in the DLPFC

following the presentation of reward feedback reflect the transfer of the task-related information preceding the feedback to other brain areas for the purpose of strengthening the active representations of this information for adaptively modifying performance. Also, in the context of the HRL-ACC theory (Holroyd & Yeung, 2012), I suggest that ACC-mediated beta oscillations reflect communication of action sequences between the ACC and the DLPFC, perhaps reflecting a mechanism for the formation of options in the ACC.

Introduction

Several influential theoretical frameworks have strongly implicated the PFC and ACC in cognitive processes related to cognitive control, decision making, and working memory, but the particulars about how these functions are carried out remain ambiguous. The PFC, supported by extensive anatomical connections across the brain, has been proposed to communicate "top-down" control signals to other brain areas to facilitate task performance (Miller & Cohen, 2001). The role of PFC in maintaining active representations of task sets through sustained neural activity, and in flexibly updating this information, reflects the primary function of working memory as this term is understood in relation to the frontal system (Fuster, 2008). Although the function of ACC is somewhat more controversial (Holroyd & Yeung, 2011), one recent proposal holds that the ACC selects and supports the execution of hierarchically organized action sequences called "options", and directs the PFC to apply those options (Holroyd & Yeung, 2012). In the context of this thesis, outside of the formal definition of options in reinforcement learning frameworks (Sutton et al., 1999), sequences of actions that are functionally linked together in a certain order can form options.

Recently, neural oscillations in the beta frequency range recorded from the PFC have provided insight into the function of this brain area. In particular, recent studies have indicated a role for beta oscillations in the DLPFC in representing rule-specific neural ensembles: It was observed that neurons in the monkey DLPFC coded for the currently relevant task rule by synchronizing in the beta frequency range (Buschman et al., 2012). Synchrony in the beta frequency range between monkey striatal and PFC neurons also increases during category learning (Antzoulatos & Miller, 2014), suggesting

that beta oscillations may facilitate communication between the PFC and striatum during such learning. This evidence suggests that better understanding of beta oscillations could elucidate how ACC and PFC communicate during option creation and execution.

Related to this issue, the relatively short period of beta oscillations (~30-50 ms) has been suggested to mediate the storage of single items of information in working memory (Lisman & Idiart, 1995). On this view, multiple beta cycles occurring during a slower oscillation like theta provide a neural code for communicating 'multi-item messages'(Lisman & Jensen, 2013). As evidence for this theory, intracranial recordings from the human hippocamus revealed that power of beta oscillations was coupled with the phase of theta oscillations during the maintenance period in a modified Sternberg paradigm (Axmacher et al., 2010). Also, intracranial and EEG recordings showed that the power of beta oscillations over the frontal cortex was positively correlated with working memory load (Basar-Eroglu et al., 2007; Howard et al., 2003).

Also, several studies involving the human EEG and MEG have reported that reward feedback stimuli compared to no-reward (or error) feedback stimuli in gambling, probabilistic learning, and time-estimation tasks elicit an increase in the power of high-beta oscillations (20-35 Hz) over frontal areas of the scalp (Cohen et al., 2007; Doñamayor et al., 2011; HajiHosseini & Holroyd, 2015a, 2015b; HajiHosseini et al., 2012; Marco-Pallares et al., 2008; van de Vijver et al., 2011). There are two main theories on the function of reward-beta oscillations: First, that beta oscillations are related to learning from positive feedback by facilitating communication across different areas in the PFC or amygdala (Cohen, Wilmes, & van de Vijver, 2011), and second, that beta oscillations elicited by positive feedback reflect coupling between neurocognitive

processes involved in attention, memory, and motivation (Marco-Pallarés et al., 2015). However, these predictions have yet to be empirically tested.

In Chapter 3, I source-localized reward-related beta power to the right DLPFC, a finding that is in line with the apparent role of PFC in actively maintaining action representations (Miller & Cohen, 2001). In light of the interrelationships between beta, working memory, and DLPFC function, and my findings in Chapters 2 and 3, I suggested that reward-related beta reflects enhanced activity in working memory of neural ensembles related to cues, responses and rewards, and the concomitant transfer of this information from DLPFC to other brain areas for improved task performance. This hypothesis suggests that reward-related beta may be sensitive to various sources of important task-related information. In particular, I propose that one aspect of task performance that contributes to beta power in reinforcement learning paradigms is the number of actions maintained in working memory preceding feedback delivery. This proposal is based on several considerations. First, the action-outcome relationship constitutes the fundamental currency of reinforcement learning paradigms, so it is plausible that reward-related beta would be sensitive to this aspect of task performance. Second, as described above, short cycles of beta oscillations are proposed to carry single items of information in working memory (Lisman & Jensen, 2013), consistent with empirical evidence associating beta oscillations with working memory load (Axmacher et al., 2010; Howard et al., 2003). Third, the HRL theory holds that ACC is responsible for selecting and maintaining options that consist of action sequences (Holroyd & Yeung, 2012), which entail maintaining the selected option in memory until its completion (Botvinick et al., 2009).

In this study I investigated this issue by recording the EEG from participants as they navigated two virtual mazes to find rewards: a radial arm maze (R-maze) in which a feedback stimulus was presented on each trial after a single action, and a more complex maze (H-maze) in which a feedback stimulus was presented on each trial after three consecutive actions. On the basis of the above considerations, I predicted that feedback would elicit greater beta power in the ACC in a task involving a sequence of actions before feedback delivery (the H-maze) compared to a task with only a single action before feedback delivery (the R-maze). Further, based on my previous findings I predicted that reward feedback would elicit greater beta power in the DLPFC compared to error feedback, thereby facilitating the encoding of the rewarded action or sequence of actions in brain areas (such as the striatum) underlying task execution. Finally, in line with the "multi-item messaging" framework (Lisman & Jensen, 2013), I predicted that post-feedback beta power would be coupled with theta phase in the H-maze, which depends are remembering multiple actions in working memory, but not in the R-maze, which depends on remembering only a single action.

Methods

Participants

Thirty-one undergraduate students (9 men, 21.7 ± 3.3 years old) at the University of Victoria participated in the experiment. Subjects acquired extra course credits for their participation and were also paid a monetary bonus that depended on task performance. The study was conducted in accordance with the ethical standards prescribed in the Declaration of Helsinki and was approved by the human subjects review board at the

University of Victoria. Formal written consent was acquired from participants prior to the experiment.

Task

Participants performed a virtual R-maze and a virtual H-maze task (Figure 9; cf. Baker & Holroyd, 2013) in a counterbalanced order. They were instructed to navigate the mazes, identify the most rewarding alleys by trial and error, and maximize their earnings. On each trial of the R-maze task, a central fixation cross was presented for 200 ms, followed by an image of the top view of the maze center with 8 alleys extending in the cardinal and primary intercardinal directions, which served as a prompt for subjects to select an alley by moving a joystick (Logitech Extreme 3D Pro) in a corresponding direction (Figure 9A). In order to limit the overall duration of the experiment, responses that exceeded 1 s were penalized with a message on the screen indicating a 10 cent loss and the instruction "Please respond faster!" Participants were not informed about the specific deadline but were told that slow responses would result in the 10 cents loss. 200 ms after the response, an image of the chosen alley was presented for 500 ms, followed by an image of the feedback stimulus (5.5^{o} of visual angle) presented for 1000 ms at central fixation over an image of the alley end. Participants were told that if they found an apple (orange) then they gained 5 cents on that trial and if they found an orange (apple) then they gained 0 cents on that trial; the assignation of reward values to feedback stimuli was counterbalanced across participants. Finally, a gray screen was presented for 2000 ms during the inter-trial interval.

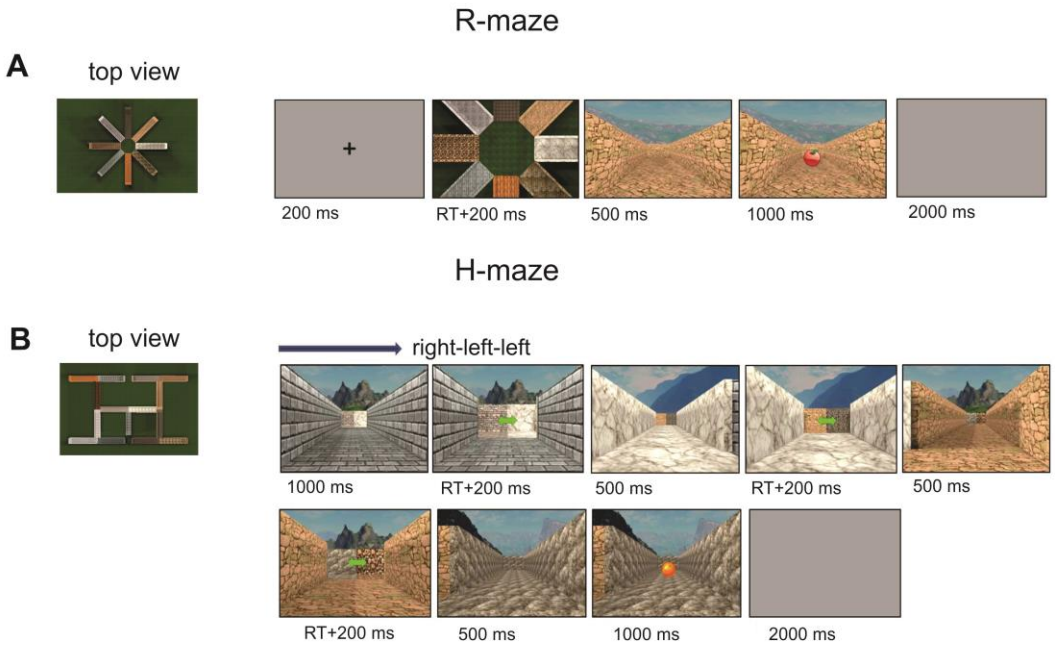

Figure 9: Time lines for example trials in the A) R-maze and B) H-maze. In the R-maze, subjects selected one of 8 alleys and were presented with a feedback stimulus indicating whether or not that choice was rewarded. In the H-maze, subjects made three consecutive responses (e.g., Right-Left-Left) prior to receiving a feedback stimulus that indicated whether or not they were rewarded on that trial. Note that the images labeled "top view" for each maze were never presented to subjects and are shown here for descriptive purposes only.

On each trial of the H-maze task, participants were required to make a series of three Right and Left responses using the joystick. The trial began with presentation of an image of the stem alley for 1000 ms (Figure 9B), followed by an image of the end of the alley overlaid with a double arrow at central fixation, indicating that participants should select either the Right or Left alley by moving the joystick in the corresponding direction. 200 ms after the response, an image of the selected alley was displayed for 500 ms followed by a view of the alley end together with a double arrow, indicating a second right-left choice to the participant. Following this second response, an image of the (second) selected alley was displayed for 500 ms, which in turn was followed by a view of the alley end together with a double arrow that indicated the third right-left choice to

the participant. Finally, 200 ms after the third response an image of the end of the alley for the corresponding location was presented for 500 ms. An image of the feedback stimulus (apple or orange) was then presented at central fixation over the image of the alley end for 1000 ms. The trial ended with a gray screen presented for 2000 ms during the inter-trial interval. If any of the 3 responses exceeded 1 s, then the subject was penalized 10 cents on that trial as described for the R-maze task. Note that the H-maze affords a total of 8 choices between individual three-step sequences. The images used to indicate the reward (5 cents) and the no-reward (0 cents) conditions were consistent across the two tasks. All of the alleys in both mazes could be distinguished by unique wall and floor surfaces. Before starting each task participants carried out 15 practice trials, and their understanding of the correspondence between the feedback stimulus and reward condition was confirmed.

Participants performed 3 blocks of 100 trials in each task. The reward probabilities associated with the eight ends of the maze were non-stationary and modified according to the following algorithm in both tasks. In the start of each block of trials, two of the eight ends were selected at random with equal probability and without replacement across blocks, hereafter called *End 1* and *End 2*. End 1 and End 2 were initially assigned 100% and 20% probabilities of reward, respectively. Then, each visit to End 1 decreased the probability of obtaining reward on future visits to that alley by 4% and increased the probability of obtaining reward at End 2 by 4%. Similarly, each visit to End 2 decreased the probability of obtaining reward on future visits to that alley by 4% and increased the probability of obtaining reward at End 1 by 4%. These reward probabilities were bounded by 100% and 0%. Each of the other six alley ends yielded rewards with a fixed 20%

probability. Therefore, optimal performance on each block of trials for both tasks depended on finding and returning to End 1 until the probabilities of rewards for End1 and End 2 were equalized at 50%, and then switching back and forth between the two ends thereafter.

Upon the completion of the experiment all participants responded to a questionnaire that asked them to 1) reconfirm their understanding of the feedback stimulus, 2) draw a sketch of both mazes, 3) identify any alleys that were more rewarding than the others, 4) report on any strategies utilized for task performance (such as navigating the mazes based on the alley textures or by memorizing sequences of left/right movements), and 4) indicate whether they found either maze tasks to be harder than the other one.

Data acquisition

The EEG was recorded from 51 electrode locations using BrainVision Recorder software (Brainproducts, Munich, Germany). Electrodes were arranged according to the standard 10-20 layout (Jasper, 1958) and were referenced online to the average voltage across the channels. Vertical and horizontal ocular movements were recorded by an electrode placed under the right eye (re-referenced offline to FP2), and two on the outer canthi of the right and left eyes (re-referenced offline to each other) respectively. Electrode impedances were kept under 20 kΩ. Data were sampled at 500 Hz and high pass filtered by the amplifiers at 0.017 Hz.

Data analysis

Data pre-processing was performed using BrainVision Analyzer 2 (Brain Products). A band-pass filter (0.1-100 Hz) was applied to the EEG data and epochs of

EEG activity were selected from 1 s before to 1 s after the onset of feedback stimuli. Data were subsequently re-referenced to the average mastoids. Ocular correction was performed using the Gratton, Coles, & Donchin (1983) algorithm as implemented in the Analyzer software. Feedback segments were baseline-corrected by subtracting, for each channel, subject, and trial, the average voltage values during the 100 ms prior to the feedback stimulus, from the subsequent voltages in the epoch. EEG artifacts were identified and rejected according to the following criteria: Any abrupt change of voltage greater than 35 μV from one time sample to the next, any difference between the negative and positive peaks in a 200 ms interval that exceeded 150 μV, and any activity that was consistently smaller than 0.5 μV in a 100 ms interval were considered artifacts and the corresponding segment was rejected for all channels. On average, 14% of data were rejected in the R-maze task and 7% of data were rejected in the H-maze task. Data were exported to MATLAB for the time-frequency analyses. Topographical scalp maps were plotted with EEGLAB (Delorme & Makeig, 2004).

In order to evaluate whether cognitive factors sensitive to outcome expectancy and task difficulty were similar across the two maze tasks, as a control analysis we measured the reward positivity, a component of the event-related brain potential associated with reward processing (Holroyd, Pakzad-Vaezi, & Krigolson, 2008; Miltner, Braun, & Coles, 1997; see Proudfit, 2014 for a review). For each subject, the pre-processed data associated with channel FCz were averaged across trials separately for the reward and error conditions for both mazes. Reward positivity amplitude was determined by calculating the mean amplitude of the difference between the ERPs to negative and positive feedback within a 200-300 ms window post-feedback, separately for the R-maze

and H-maze. Channel FCz (e.g., Holroyd, Nieuwenhuis, Yeung, & Cohen, 2003; Holroyd et al., 2008) and the time window of interest (e.g., Holroyd, Krigolson, & Lee, 2011) were selected a priori based on previous studies.

To extract time-frequency information, each two-second epoch of EEG data centered on feedback stimulus onset was convoluted with a Morlet wavelet (family ratio = 6.7), the width of which was linearly scaled based on the frequency range of 1-100 Hz. Time-frequency power was extracted relative to a 100 ms baseline before feedback onset on each trial as 10*log10 (trial power/average baseline power). Then the power values were averaged separately for each condition and subject. Beta power for every subject and condition was calculated as the average power in a window of interest, from 20-30 Hz and 250-450 ms post-feedback (Chapter 2).

To reduce the spatial dimensionality of the data, principal component analysis (PCA) was performed on average beta power values to identify clusters of EEG data that covaried spatially across the scalp. Individual spatial factors were identified by submitting average beta power for every subject and condition to PCA using MATLAB. The first 5 components that together explained over 90% of the total variance, comprising 5 ‘virtual channels’, were retained for Varimax rotation. Spatial factor scores for each virtual channel were obtained by taking the inner product of the factor loadings with the average beta power values for each electrode, for each time point, subject and condition (Spencer, Dien, & Donchin, 2001).

To calculate the coupling between beta power and theta (4-8 Hz) phase, the time course of beta power for individual trials was concatenated across all trials, for each subject and condition. Then, I obtained the factor scores at each time point of every trial

by multiplying the voltage values at each time point and channel with the PCA factor loadings. The factor scores concatenated across trials were then convoluted with a complex Morlet wavelet with family ratio (4.4) (Cohen, 2014). From the resulting complex signal, I extracted the power values for frequencies between 20 and 30 Hz (A_β) and phase values for frequencies between 4 and 8 Hz ($e^{i\varphi_\theta}$) at each time point and each frequency and constructed new complex time series using these high-frequency power and low-frequency phase values: $A_\beta e^{i\varphi_\theta}$. Then I measured the absolute value of the average of these new time series as our phase-amplitude-coupling (PAC) value according to Equation 1, where t is each data point and n is the number of data points in the concatenated signal (Cohen, 2014).

$$PAC = \left| {}^{1}\!/_{n} \sum_{t=1}^{n} A_{\beta t} e^{i\varphi_\theta t} \right| \qquad (1)$$

To correct for random inflations of the PAC values I performed a random permutation technique. I randomly picked a time point in each trial, swapped the high-frequency power values before and after that time point, and calculated the PAC values (PAC_i) based on this randomly constructed power time series and the actual low-frequency phase time series repeatedly for 100 iterations. Then I corrected the PAC values by subtracting the mean of the random PAC values over 100 iterations and dividing the outcome by their standard deviation (PAC_z) according to Equation 2.

$$PAC_z = (PAC - {}^{1}\!/_{n} \sum_{i=1}^{100} PAC_i) \Big/ std(PAC_i) \qquad (2)$$

Then, PAC_z values were averaged across theta (4 to 8 Hz) and beta (20-30) frequency bands to calculate the average PAC index at each condition for each subject between beta power and theta phase.

Finally, the coupling values might be artificially inflated by the temporary burst of beta power and increased inter-trial phase coherence (ITPC) in theta frequency range (Cohen, 2014), which are both expected to occur following the presentation of feedback stimulus. Theta phase coherence reflects the portion of total theta power that is consistent in phase with the eliciting event (i.e., the feedback stimulus) across trials. If theta is phase coherent, the complex values calculated as $A_\beta e^{i\varphi_\theta}$ in Equation 1 are in phase with each other across trials for most time points, therefore inflating the overall vector mean. By contrast, actual coupling with beta power requires an increase in beta power A_β at a certain theta phase $e^{i\varphi_\theta}$ consistently across trials. In other words, the appearance of increased beta-theta coupling could be due to more beta (without increased coupling per se) and/or to increased theta phase coherence (without increased coupling). To address this possibility, the average theta ITPC for 1000 ms post-feedback was calculated according to Equation 3 for each trial, each condition, and each subject.

$$ITPC = \left| {}^{1}\!/_{n} \sum_{t=1}^{n} e^{i\varphi_\theta t} \right| \qquad (3)$$

Then the ITPC values were averaged across trials in the same condition and a value was calculated for each subject and condition. Following the partial regression method I had previously used in HajiHosseini and Holroyd (2013), I partialed out beta power and theta ITPC values from the PAC_z values for each subject by linearly

regressing PAC_z on to beta power and theta ITPC, collapsed across all conditions and subjects, and subsequently analyzing the residuals of the regression.

To identify the sources of beta effects related to feedback valence and maze type, source localization was performed with sLORETA (Pascual-Marqui, 2002). For each subject, channel, and trial, a 2-second data segment spanning one second before feedback onset to one second after feedback onset was analyzed for time-varying cross-spectra in sLORETA with a 72 sample-long Gaussian window for the beta frequency range (20-30 Hz) and for 250-450 ms post feedback. Note that source localization cannot be conducted directly on power values, which are related to the square of the voltage values. Therefore, sLORETA brain maps were determined by recalculating the time-varying cross-spectral power values for each subject and condition. Then statistically significant differences in beta power values were identified for each contrast by conducting paired t-tests for each voxel; the voxels with the largest t-values are reported. Randomization via statistical non-parametric mapping was applied in sLORETA to correct for multiple comparisons. To maximize statistical power, beta power values were collapsed across mazes for the valence contrast and across reward and error conditions for maze task contrast.

All reported measurement errors for the behavioral data constitute standard deviations. Total reported trials exclude “miss” trials when subjects failed to respond within the response deadline, on average 2% of total trials across subjects in both mazes.

Results

Behavioral analysis

Participants visited the rewarding arms in the R-maze (47 ±12% of total trials) significantly more than in the H-maze (41 ±10% of total trials), ($t(30) = 2.16$, $p = 0.04$). Upon completing the experiment, 7 subjects reported the R-maze to be the harder task, 20 reported the H-maze to be harder, and 4 reported the two mazes to be equally difficult. Visual inspection indicated that 19 subjects drew the H-maze correctly and 12 subjects drew it incorrectly.

In the H-maze, subjects' first response in the sequence (321±119 ms) was significantly slower compared to the second response (277±112 ms) ($t(30) = 4.54$, $p < 0.001$) and to the third response (273±122 ms) ($t(30) = 4.47$, $p < 0.001$), whereas the reaction times on the second and third responses were not significantly different ($p > 0.05$). Further, participants responded about equally fast in the R-maze (332 ±80 ms) and for the first left-right choice in the H-maze, whereas R-maze responses were significantly slower than the second ($t(30) = 2.81$, $p = 0.009$) and third ($t(30) = 2.86$, $p = 0.008$) responses in the H-maze.

Time-frequency analysis

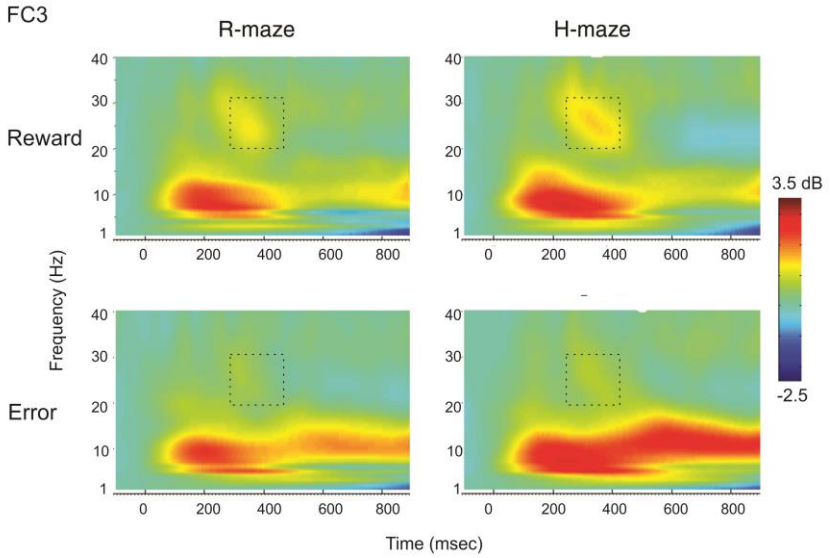

Figure 10: Time-frequency maps of EEG power from 1-40 Hz associated with channel FC3. The panels show EEG power following the reward (top) and error (bottom) feedback in the R-maze (left) and the H-maze (right). The dotted boxes mark the beta frequency range (20-30 Hz) in 250-450 ms post-feedback. Note the relative increase in beta power in the reward conditions in the R-maze (following a single action) and the H-maze (following 3 actions), which is larger for the latter compared to the former.

Power. For the purposes of illustration, Figure 10 provides time-frequency maps of EEG power from 1-40 Hz recorded at channel FC3, which revealed the largest contrast in beta power elicited to reward compared to error feedback. As predicted, the beta burst occurred around 250-450 ms post-feedback from about 20 Hz to 30 Hz. PCA on the beta power values yielded five virtual channels that explained at least 90% of the total variance. As illustrated in Figure 11, three of the virtual channels were associated with frontal distributions (frontal-lateral, VC1; frontal-polar, VC3; and frontal-central, VC4), and the remaining two virtual channels were associated with posterior distributions (occipital, VC2 and temporal-parietal, VC5).

My hypothesis about reward-related beta concerns beta distributed over frontal-lateral areas of the scalp (Cunillera et al., 2012; HajiHosseini & Holroyd, 2015a; HajiHosseini et al., 2012). For this reason, my subsequent analyses focussed on the spatial factor associated with this distribution, VC1 that accounted for 74% of the

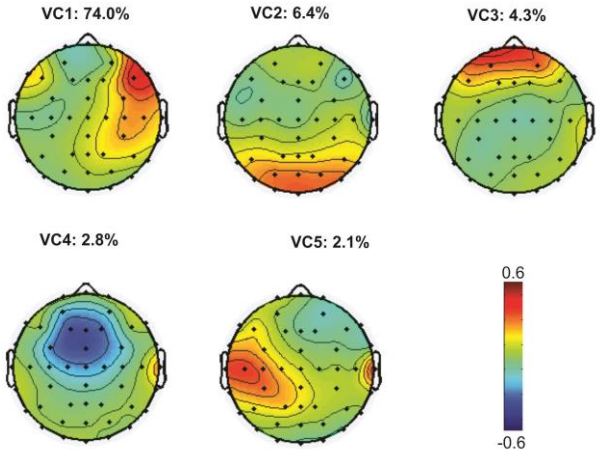

Figure 11: Principal Component Analysis (PCA) results. PCA yielded five virtual channels that together explained more than 90% of the total variance of post-feedback beta power. Each scalp distribution shows the loadings of each factor across the 45 electrode locations, together with the percentage of total variance explained. "VC1" = virtual channel 1, "VC2"=virtual channel 2, etc.

variance in data. A 2×2 ANOVA with maze (R-maze, H-maze) and feedback valence (reward, error) on VC1 revealed a main effect of maze ($F(1,30) = 6.4$, $p = 0.02$) and a main effect of feedback valence ($F(1,30) = 7.3$, $p = 0.01$) (Figure 12A). The number of visits to the good arms in the H-maze was also positively correlated across subjects with beta power in the reward condition ($r = 0.38$, $p = 0.03$) (Figure 12B). This correlation survived the exclusion of an outlier (beta power > 20 dB) ($r = 0.37$, $p = 0.045$).

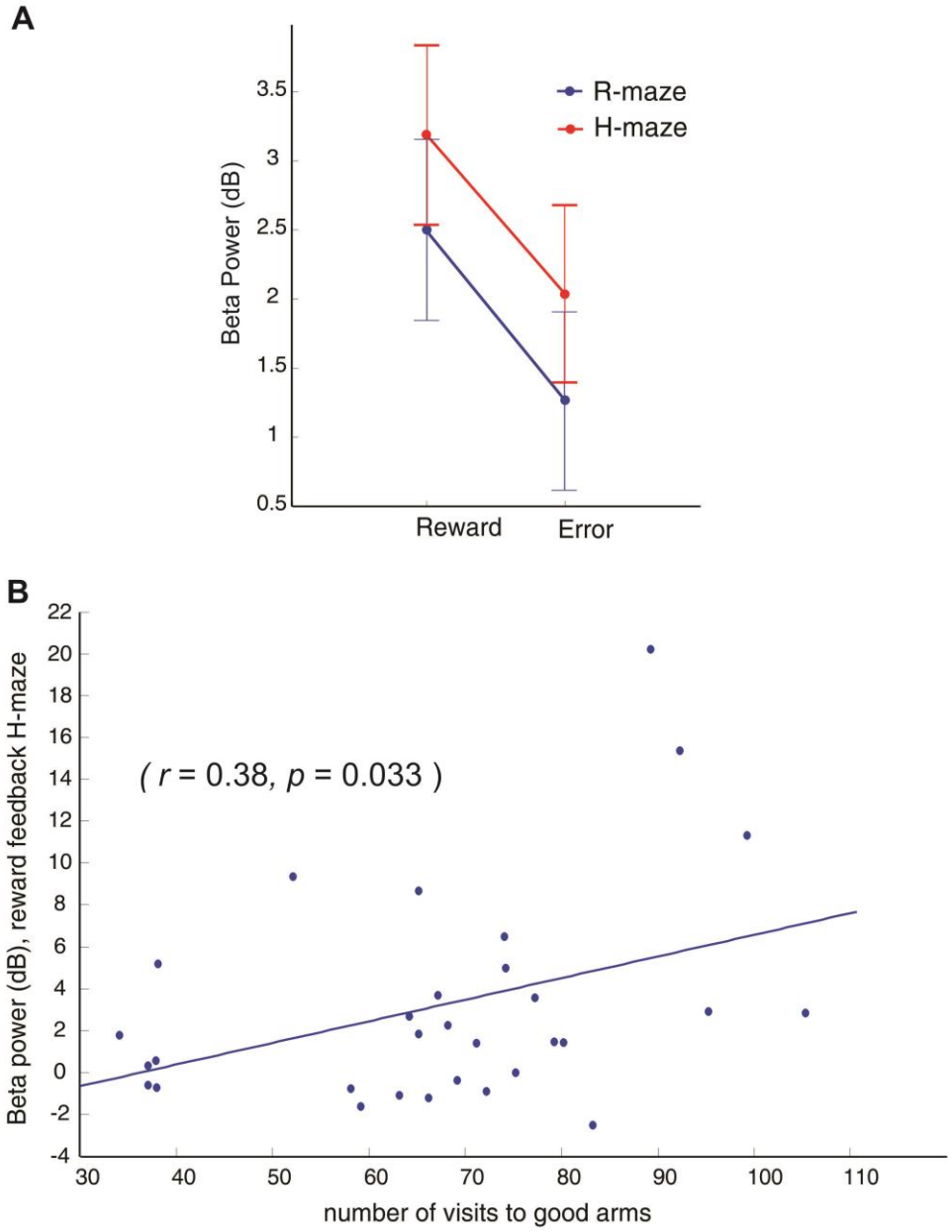

Figure 12: Frontal beta power at virtual channel 1 and behavior. A) Beta power to the reward and error feedback for the R-maze (blue) and the H-maze (red). Error bars indicate pooled estimate of within-subject 95% confidence intervals. B) Correlation between beta power to reward feedback in the H-maze with the number of visits to the good arms. Note that correlation holds when excluding the outlier with beta power > 20 dB (r = 0.33, p = 0.045).

Source localization. I source-localized the valence and maze contrasts on frontal beta power to verify my predictions of its neural source. For the valence contrast, the most positive t-value (t = 3.92, p < .05, X= -20, Y= 35, Z = 45, MNI[4]) was localized on Brodmann area 8 in the left DLPFC (Figure 13A), indicating that the left DLPFC produced more beta to reward feedback than to error feedback. By contrast, the most positive t-value value for the maze contrast was localized on the ACC (Figure 13B) (t = 5.82, p < .001, X= 5, Y= 15, Z = 25, MNI).

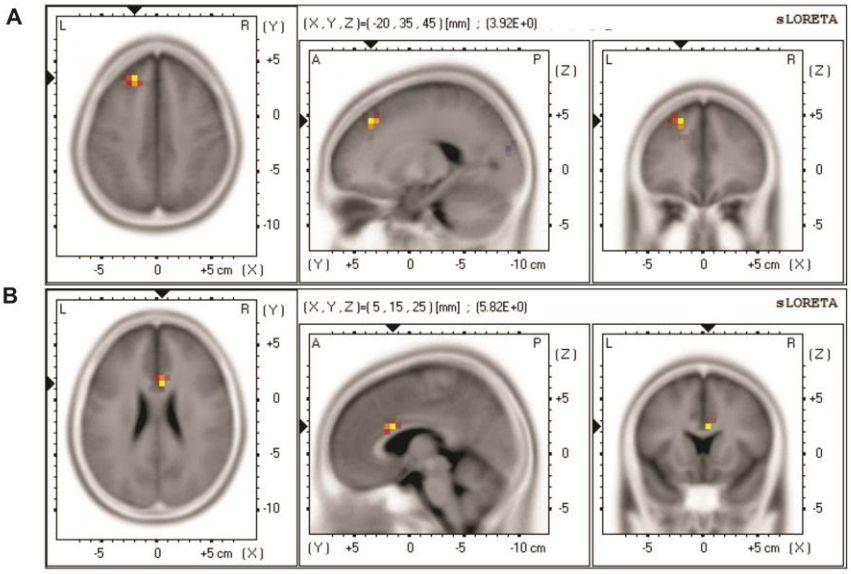

Figure 13: SLORETA maps of beta power sources. Sources for A) the valence contrast (reward vs. error) and B) the maze contrast (H-maze vs. R-maze). Statistically significant differences in beta power values were identified for each contrast by conducting paired t-tests for each voxel and reported t-values are related to the voxels with the largest t-values, as implemented by sLORETA. Randomization via statistical non-parametric mapping was applied in sLORETA to correct for multiple comparisons.

[4] Montreal Neurological Institute brain coordinate system

Cross frequency coupling. Figure 14A indicates the cross-frequency coupling index between beta power at VC1 and the phase of theta oscillations. Coupling at all conditions were different from zero ($p < 0.001$) but 2×2 ANOVA with maze (R-maze, H-maze) and valence (reward, error) as factor did not show any difference across conditions ($p > 0.05$). As explained in the methods section, temporary increases of beta power and theta inter-trial phase coherence following the presentation of the feedback stimuli might independently inflate the PAC index. These temporary increases can separately contribute to the power and phase parameters in Equation 1 in the absence of actual increased coupling, which requires that the increase in beta power coincide at a certain phase of theta on every cycle. Therefore, prior to statistical analysis, partial regression was utilized to remove the contribution of ITPC and beta power (Figure 14B) from the

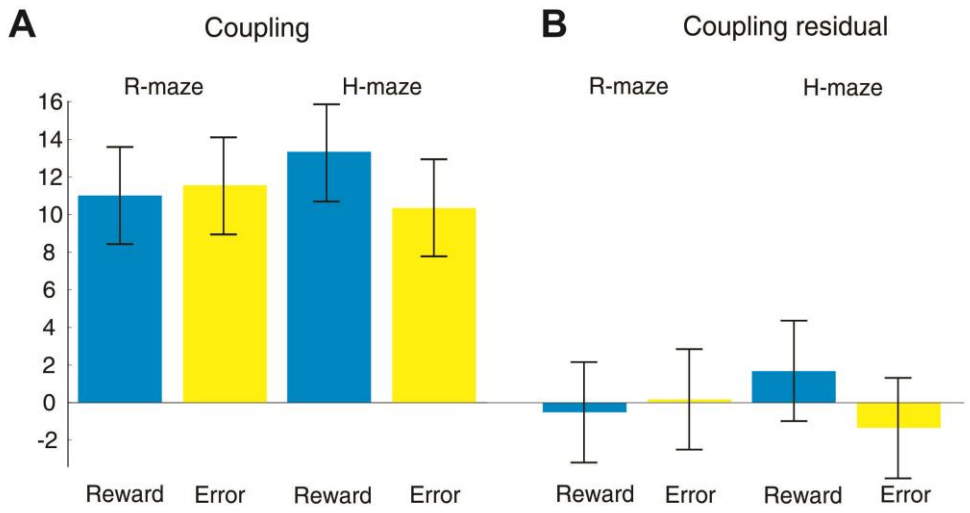

Figure 14: Cross-frequency coupling index (phase amplitude coupling). A) Coupling in the R-maze (left) and H-maze (right) following reward (blue) and error (yellow) feedback. B) Coupling index when theta inter-trial phase coherence and beta power are partialed out for the R-maze (left) and H-maze (right) following the reward (blue) and error (yellow) feedback. Error bars show pooled estimate of within subject 95% confidence intervals (Masson & Loftus, 2003).

normalized phase amplitude index (PAC_z) for all subjects (HajiHosseini & Holroyd, 2013). As illustrated in Figure 14B, after partial regression the coupling index is highest following the reward feedback in the H-maze. However, this value was not statistically significant when compared against zero, perhaps due to lack of statistical power.

ERP analysis

To ensure that cognitive processes related to basic aspects of task performance such as task engagement and outcome expectancy were comparable across the two mazes, I evaluated the reward positivity elicited by the feedback stimuli in the R-maze and H-maze (Figure 15). Reward positivity is a robust and extensively studied ERP component that is sensitive to the interaction of outcome probability and outcome valence (Holroyd & Krigolson, 2007; Holroyd, Nieuwenhuis, Yeung, & Cohen, 2003; for reivew see Sambrook & Goslin, 2015) and is proposed to reflect the degree of task engagement

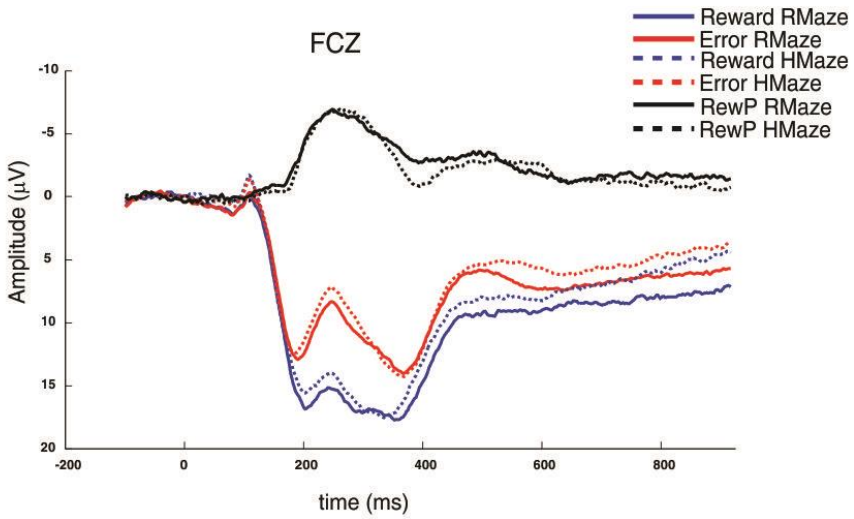

Figure 15: Event related potentials (ERPs) to reward and error feedback stimuli, and the associated reward positivites, for the R-maze and H-maze at FCz. Waveforms indicate the ERPs following the reward (blue) and error (red) feedback in the R-maze (solid) and the H-maze (dotted). The black lines show the reward positivity (error vs. reward difference wave) in the R-maze (solid) and the H-maze (dotted).

mediated by ACC (Holroyd & Yeung, 2012; see also Proudfit, 2014). The reward positivity was present for both conditions ($p < .05$ when tested against zero), but reward positivity amplitude was not significantly different across the two mazes, indicating relatively equivalent task engagement and feedback expectancy ($p > 0.05$).

Discussion

PFC and ACC are understood to work in parallel for the purpose of controlling behavior. On one recent account, these systems are organized hierarchically, with the ACC responsible for choosing which task to execute, and PFC responsible for applying control over the execution of that task by other brain areas such as the dorsal striatum (Holroyd & Yeung, 2012). I proposed in Chapter 3 that increased power in the beta frequency range following reward receipt (Cohen et al., 2007; Marco-Pallares et al., 2008) reflects the latter process, namely, that rewards boost the activation of task-related neural assemblies in the DLPFC that encode the stimulus response mapping for that trial, which in turn facilitates the transfer of that information to the striatum and other brain areas. This suggests that frontal beta oscillations are related to a working memory mechanism that is elicited by reward feedback stimuli. Further, the HRL-ACC theory suggests that the task representations communicated to PFC by ACC are organized as sequences of actions (policies) called options (Holroyd & Yeung, 2012); and an influential theory of working memory proposes that beta oscillations should reflect the number of items in working memory (Lisman & Idiart, 1995), a prediction that has found empirical support (Axmacher et al., 2010; Howard et al., 2003). Taken together, these considerations suggest that frontal beta oscillations produced in the ACC should be

sensitive to the number of actions in the action sequence preceding feedback presentation.

To test these predictions, I had subjects do two versions of a reinforcement learning paradigm, one requiring a single action (R-maze) and the other requiring a sequence of actions (H-maze) prior to presentation of the feedback stimulus. I predicted that reward feedback would elicit greater beta power compared to non-reward feedback, a process mediated by the DLPFC (as observed in Chapter 3), and that feedback following a sequence of actions would elicit greater beta power than feedback following a single action, a process mediated by the ACC. Also, based on the oscillatory model for the communication of multi-item messages, I also predicted greater cross-frequency coupling of beta power to theta phase following the reward feedback in the H-maze compared to the R-maze. Consistent with these predictions, I found that the reward feedback stimuli compared to error feedback stimuli elicited greater beta power over frontal areas of the scalp in both mazes, an observation that was confirmed by spatial PCA and that was source-localized to left DLPFC. Furthermore, beta power over frontal areas was also sensitive to the sequence length, revealing greater power for feedback following several actions compared to feedback following a single action, and source analysis localized this contrast to the ACC. I also found greater phase-amplitude coupling between theta and beta band oscillations that was larger for the H-maze than the R-maze, but this result was not statistically significant.

Taken together and in the context of the previous literature, these findings elucidate how ACC and DLPFC work together to support goal-directed behavior. I suggest that beta oscillations in the DLPFC following the presentation of reward

feedback reflect the transfer of task-related information preceding the feedback to other brain areas for the purpose of strengthening the active representations of this information for adaptively modifying performance. Consistent with this proposal, beta oscillations in the monkey DLPFC reflect the selection of task-specialized neural ensembles in this area and synchronization of DLPFC neurons with striatal activity during learning (Antzoulatos & Miller, 2014; Buschman et al., 2012), and frontal beta oscillations are associated with working memory load in humans (Basar-Eroglu et al., 2007; Howard et al., 2003).

Also, the HRL-ACC theory proposes that the ACC is responsible for selecting and motivating the execution of extended, goal-directed sequences of actions, and communicates this directive to PFC for execution (Holroyd & McClure, 2015; Holroyd & Yeung, 2012). I propose that ACC-mediated beta oscillations reflect communication of these action sequences between the ACC and the DLPFC. This proposal is supported by evidence from working memory studies indicating a role for the ACC in mediating attentional and memory demands for task performance via increased connectivity between the ACC other areas needed to sustain these processes (Lenartowicz & McIntosh, 2005), and evidence that ACC neurons sustain activity to support task-related information in working memory (Petit, Courtney, Ungerleider, & Haxby, 1998).

It is also possible that beta oscillations in the ACC contribute to the formation of options by communication with the DLPFC. Notably, the HRL-ACC theory does not specify how options are developed, but in this experiment the action sequences that comprise the options must be learned from scratch. It is possible that active representations of action sequences maintained by the ACC are boosted by the DLPFC following task-related feedback (such as reward feedback). As learning progresses, the

values of these action sequences are obtained by the ACC as option values. Also, evidence from my unpublished data suggests that frontal beta diminishes when tasks become overlearned. This suggests that when tasks become overlearned, the need for a top-down signal to govern this communication declines.

Note that in Chapter 3 the valence effect was localized to right DLPFC, whereas here it was localized to left DLPFC. Recent studies have shown that right DLPFC supports cognitive processes that require non-verbal abstract problem solving (Barbey et al., 2013; D'Ardenne et al., 2012) whereas left DLPFC supports other processes like letter-number sequencing (Barbey et al., 2013). Therefore, the source of beta oscillations might also depend on specific task demands. Supporting this idea, subject self-reports in the current study showed that most subjects tried to remember the sequence in the H-maze by mental repetition of the sequence (e.g., left, right, right), suggesting that reward-related beta in left DLPFC may be related to verbal maintenance of the action sequences on each trial.

Unexpectedly, I also observed greater beta power in the cuneus to error feedback compared to reward feedback; follow-up source-analyses of the valence contrast, conducted separately for the H-maze and the R-maze, revealed the cuneus as the source of error-related beta power for the R-maze but not the H-maze. Given the well-known role of the cuneus in visual perception (Vanni, Tanskanen, Seppä, Uutela, & Hari, 2001), I speculate that this activation might reflect greater reliance on a visually-based task strategy afforded by the top-down view of the maze, which by presenting all 8 possible choices simultaneously is more informative than the individual images in the H-maze.

However, the specific role of beta oscillations in this process is a question for future research.

In summary, I propose that frontal beta oscillations represent a mechanism governed by the ACC and the DLPFC for maintenance and transfer of information regarding goal-directed actions or action sequences. The ACC maintains the task-related information of action sequences online, whereas the DLPFC signals a need for boosting and transfer of this information to other brain areas responsible for task execution such as the BG. Further, as I discussed in Chapter 1 and 3, when actions or options become automatic, frontal beta diminishes suggesting that these oscillations might have a role in learning the task rules or formation of 'options'—here action sequences—governed by the ACC.

Chapter 5: Experiment Four

Abstract

Several studies have shown that reward feedback stimuli elicit increased beta power (20-30 Hz) over frontal areas of the human scalp. In a previous study, I proposed that frontal beta power reflects the transfer of information about recently rewarded actions from the DLPFC to brain areas responsible for task execution (Chapters 3 and 4). Here, to test whether frontal beta power is specifically sensitive to remembering rewarded actions, I used a novel guessing task that included on each block of trials separate choice and recall phases. Depending on condition, subjects were given instructions to remember the rewarded decks (remember reward: RR) or the decks followed by error feedback (remember error: RE), or were not provided with instructions (no-recall: NR). Beta power to reward feedback was larger in the NR and RR conditions compared to the RE condition. Further, beta power to reward and error feedback in the choice phase was negatively correlated with reaction times for correct recalls in the RR conditions, indicating that high beta power in the choice phase predicted faster correct responses in the recall phase. These results point to an important role for frontal beta oscillations in mediating a working memory process elicited by performance feedback irrespective of its valence. In the context of our previous studies, I suggest that this process is elicited by feedback stimuli for the purpose of transferring information about recent sequences of actions maintained in the PFC to brain areas responsible for task execution.

Introduction

Neural oscillations associated with PFC function have been proposed to reflect neurocognitive processes related to cognitive control, attention, and memory (Benchenane et al., 2011; Cavanagh & Frank, 2014; Hsieh & Ranganath, 2014). The functional role of neural oscillations in relation to these processes is a hot topic in current human and animal research. This interest is partly motivated by the proposal that the PFC supports the active maintenance of task representations that send "top-down" control signals via extensive anatomical projections to regulate the function of other brain areas (Miller & Cohen, 2001). For example, LFP recordings from the monkey DLPFC have provided intriguing evidence that neural oscillations in beta frequency range reflect the activation of specialized neural ensembles underlying the selection and implementation of task sets, in coordination with other brain areas such as striatum and hippocampus (Antzoulatos & Miller, 2014; Brincat & Miller, 2015; Buschman et al., 2012).

Related to this, EEG recordings have consistently revealed an increase in beta oscillatory power distributed on the frontal areas of the scalp that peaks around 350 ms following the delivery of reward-related feedback stimuli in a variety of reinforcement learning tasks (e.g., Cohen, Elger, & Ranganath, 2007; Marco-Pallares et al., 2008). These findings have raised the question of whether or not beta oscillations have a role in feedback-guided learning (Luft, 2014). A recent proposal postulates a role for these oscillations in learning from rewards whereas theta oscillations underlie learning from errors, negative feedback, and punishments (Cohen, Wilmes, & Vijver, 2011; see Cavanagh & Frank, 2014 for a review on frontal midline theta oscillations). According to

this account, beta oscillations facilitate reward learning by synchronizing the activity of multiple brain areas such as the DLPFC, frontopolar cortex, and motor cortex, but the specific neural mechanism mediating this process remains unclear. Reward-related frontal beta has also been proposed to govern the execution of cognitive processes related to motivation, attention, and memory processes (Marco-Pallarés et al., 2015). However, these proposals await empirical verification.

Given the frontal distribution of beta oscillations, the evidence on the role of beta oscillations in selecting task rules, and the great body of literature on the role of the PFC in maintenance, update, and implementation of task sets (Sakai, 2008), it is possible that reward-related beta oscillation represent the role of the PFC in these processes. Consistent with this proposal, in Chapters 3 and 4 I showed a DLPFC source for the valence contrast (reward vs. error) in beta power. Further, in Chapter 4, I elaborated on this finding by showing that beta was sensitive to the number of actions in a sequence preceding feedback delivery irrespective of its valence, and I localized the source of this contrast to the ACC. I proposed that DLPFC-mediated beta reflects a working memory process involved in the activation of a neural ensemble encoding task-related information, and the transfer of that task-related information to other neural systems that are responsible for task execution. In reinforcement learning tasks, this process manifests as greater activity following reward feedback about recent actions in order to transfer information about successful actions to other brain areas. This information facilitates behavioral adaptation on subsequent trials, which in standard reinforcement learning settings entails gaining more rewards (Chapter 3). This process is elicited by reward feedback to boost the active representation (i.e., memory) of the immediately preceding

action sequence, but not by error feedback, because error-related information is less helpful for improving task performance.

Here I tested this hypothesis by experimentally manipulating the 'task-related' value of reward feedback using a dual-task paradigm. In a "No recall" (NR) condition, which served as a control, participants engaged in a relatively standard reinforcement learning task in which I predicted that rewards would elicit frontal beta as previously observed. Further, in a "remember rewards" (RR) condition, participants were required to do the same task as in NR condition while also remembering the choices that led to reward feedback, to be recalled later at the end of each block of trials. I predicted that because the demands of the working memory task were congruent with the demands of the reinforcement learning task in this condition, the reward feedback would also elicit more beta than error feedback. Finally, in a "remember errors" (RE) condition, participants were instructed to remember the choices that led to error feedback and to recall those choices at the end of each block of trials. I predicted that the task demands would reverse the beta effect, causing greater beta activity following error feedback than following reward feedback.

Methods

Participants

Thirty undergraduate students (9 men, 20.0 ± 1.6 years old) at the University of Victoria participated in the experiment. Subjects acquired extra course credits for their participation and were also paid a monetary bonus for their performance. The study was conducted in accordance with the ethical standards prescribed in the Declaration of

Helsinki and was approved by the human subjects review board at the University of Victoria. Informed written consent was obtained from participants prior to the experiment.

Task

Participants performed three memory conditions (NR, RR, and RE) of a card choice task for monetary reward, coded with the Matlab Psychophysics Toolbox extensions (Brainard, 1997), separated into 20 blocks of 10 trials each. During an initial *choice phase* on each block, participants sequentially selected ten cards by trial and error

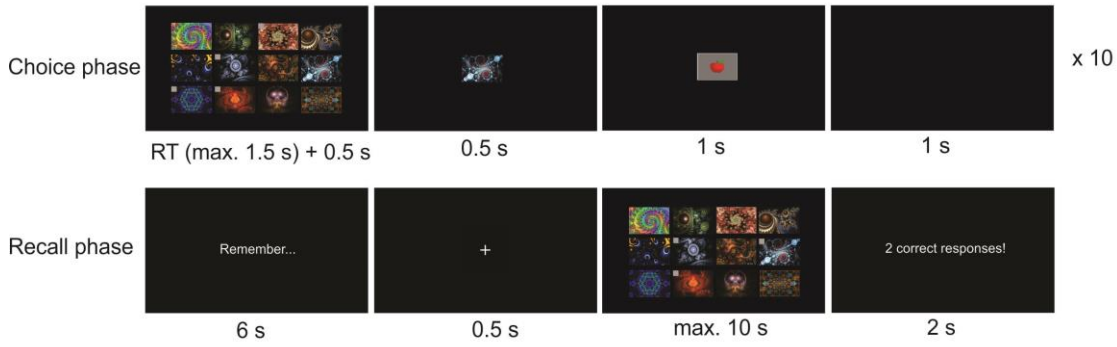

Figure 16: Task timeline. A choice phase on each block consisted of 10 trials in which on each trial participants selected a "deck" of cards. 12 abstract stimuli representing decks were presented at the same locations throughout the block. Every deck could be selected only once in 10 trials, and on each trial the previously selected decks on that block were marked with a small gray square in the corner of the image (top row, far left; three decks have been selected in the previous trials to this example trial). In the RR and RE conditions only, a recall phase occurred after the completion of each block of 10 trials. Participants were given 6 seconds to "remember" and then were presented with the 12 stimuli in the same locations following a fixation cross. They were instructed to select the target decks (reward in the RR and error in the RE conditions); upon selection of all targets or when 10 seconds were elapsed, the number of correct recall responses were presented on the screen.

from 12 "decks" presented on the computer screen, with the goal of maximizing their rewards (Figure 16, top row). For the RR and RE conditions only, the choice phase on

each block was followed by a *recall phase*, during which subjects were required to recall the decks corresponding to the feedback type as indicated by the condition instructions (Figure 16, bottom). In the beginning of the RR and RE conditions subjects were given instructions to remember either the rewarded decks or the decks followed by error feedback, respectively, for later report in the recall phase; no recall instructions were provided in the NR condition. Participants did one block of practice trials (including both phases for the RR and RE conditions) prior to performing the task in each memory condition.

During the choice phase, at the start of each trial 12 fractal images (arranged across 3 rows and 4 columns; 19cm×26.5 cm), described to subjects as representing 12 decks of cards, were presented on a black screen (Figure 16, top row, far left). Participants were instructed to use a left mouse button click to select one deck with the goal of maximizing their rewards. For responses that exceeded 1.5 s, the trial was terminated and the next trial commenced; participants were not informed of this specific deadline but were instructed to respond fast enough to avoid such misses. 500 ms following the response, the selected card was shown in the center of the screen (6^{o} of visual angle) for 500 ms (Figure 16, top row, middle left). Then the feedback was presented as a cartoon fruit image on a gray background, which was described to subjects as appearing on the back of the card, for 1000 ms (Figure 16, top row, middle right). The feedback stimuli consisted of an apple or orange indicating 5 cents (“reward feedback”) or 0 cents (“error feedback”), counterbalanced across participants, and a banana indicating 1 cent (“neutral feedback”) that was not utilized for data analysis and therefore

not counterbalanced (see below). Feedback was followed by a 1000 ms inter-trial interval (Figure 16, top row, far right) and then the next trial began.

Participants were instructed to use the mouse to select one deck by trial and error with the goal of accumulating as many rewards as possible. Participants could select a card from each deck only once; each selected deck was marked with a gray square in the corner of the image, indicating that it was unavailable for selection (Figure 16, top row, far left). Participants completed 10 trials in each block. The 12 decks were always presented in the same locations during the block.

In the RR and RE conditions, in addition to the instructions to try to find as many rewards (5 cents) as possible, participants were also asked to remember the decks associated with reward feedback or error feedback in the RR and RE conditions, respectively. Then, during the subsequent recall phase participants were given 6 s to prepare to report the target decks during the recall phase, as indicated by the word "Remember…" presented in white font on a black screen during this period (Figure 16, bottom row, far left). A white fixation cross was then presented in the center of the screen for 500 ms (Figure 16, bottom row, middle left) followed by the image of the 12 decks in the same locations as presented throughout the choice phase (Figure 16, bottom row, middle right). Participants were instructed to click on the target decks according to the trial outcomes in the preceding choice phase. The recall phase terminated when all targets (decks associated with reward feedback in the RR condition and decks associated with error feedback in the RE condition) were selected, or after 10 s, whichever came first. The number of correct responses, i.e. the number of times the target decks were selected, was then presented on the screen for 2 s (Figure 16, bottom row, far right). There were no

memory instructions given in the NR condition and the order of performing the RR, RE, and NR conditions was counterbalanced across participants. Following the recall phase in the RR and RE conditions and following the choice phase in the NR condition, end of the block was indicated by text on the screen, allowing subjects to continue to the next block at their own pace.

In each block, the 12 stimuli were randomly drawn from a set of 24 fractal images that were downloaded from The Fractal World Gallery at http://www.enchgallery.com/. I chose abstract images of fractals to minimize the types of memory strategies that participants might employ relative to stimuli with semantic content. Fractal images have been previously used for reinforcement learning (Gläscher, Daw, Dayan, & O'Doherty, 2010) and working memory (Ragland et al., 2002) task paradigms.

The choice phase of each block consisted of 10 trials with reward, error, or neutral outcomes. Unbeknownst to the participants, the outcomes were predetermined and independent of their behavior in order to control the number of trials in each condition. To be specific, the number of reward and error outcomes on each block varied randomly between 2 and 6 so that their sum was always 8, and the number of neutral outcomes was fixed at 2. The two neutral outcomes on each block were included to ensure that remembering the decks associated with either reward feedback or error feedback were not redundant tasks; without the neutral feedback, participants could have successfully performed the task by remembering only the decks associated with reward feedback or error feedback in both RR and RE conditions irrespective of the task instructions. The number of reward and error trials across blocks was selected at random from a uniform distribution between 2 and 6, which ensured a comparable number of trials for the reward

and error conditions across the 20 blocks. Therefore, across the 20 blocks of 10 trials, about 40% were associated with reward outcomes, 40% were associated with error outcomes, and 20% were associated with neutral outcomes, yielding about 80 reward trials and 80 error trials total. The monetary reward was calculated and paid to the participants at the end of each condition of the task (~13.20 CAD).

Upon the completion of the experiment all participants answered a brief, 5-question questionnaire that asked them to rate the difficulty of the conditions and explain what, if any, strategies they used to perform the task.

Data acquisition

The EEG was recorded from 53 electrode locations using BrainVision Recorder software (Brainproducts, Munich, Germany). Electrodes were arranged according to the standard 10-20 layout (Jasper, 1958) and were referenced online to the average voltage across the channels. Vertical and horizontal ocular movements were recorded by an electrode placed under the right eye (re-referenced offline to FP2), and two on the outer canthi of the right and left eyes (re-referenced offline to each other) respectively. Electrode impedances were kept under 20 kΩ. Data were sampled at 500 Hz and high pass filtered by the amplifiers at 0.017 Hz.

Data analysis

Data pre-processing was performed in BrainVision Analyzer 2. A band-pass filter (0.1-100 Hz) was applied to the EEG data and epochs of EEG activity were selected from 1 s before to 1 s after the onset of feedback stimuli. Data were subsequently re-referenced to the average mastoids. Ocular correction was performed using the Gratton, Coles, & Donchin (1983) algorithm as implemented in the Analyzer software. Feedback segments

were baseline-corrected by subtracting, for each channel, subject, and trial, the average voltage values during the 100 ms prior to the feedback stimulus, from the subsequent voltages in the epoch. EEG artifacts were identified and rejected according to the following criteria: Any abrupt change of voltage greater than 35 μV from one time sample to the next, any difference between the negative and positive peaks in a 200 ms interval that exceeded 150 μV, and any activity that was consistently smaller than 0.5 μV in a 100 ms interval were considered artifacts and the corresponding segment was rejected for all channels. Nearest neighborhood channel interpolation (Hjorth algorithm embedded in BrainVision analyzer) was used for channels that caused more than 15% trial rejection rate in the recording of each condition. After the interpolation, 7% of trials were rejected on average across all recordings for all subjects. Data were exported to MATLAB for the time-frequency analyses. Topographical scalp maps were plotted with EEGLAB (Delorme & Makeig, 2004).

To extract time-frequency information, the two-second epoch centered on the feedback presentation time was convoluted with a seven-cycle complex Morlet wavelet that was linearly scaled based on the frequency range of 1-100 Hz. The time-frequency power was extracted relative to a 100 ms baseline before the feedback on each trial as 10*log10 (trial power/average baseline power). Then the power values were averaged separately for each condition and subject. Beta power for every subject and condition was calculated as the average power in a window of interest spanning 20-30 Hz and 250-450 ms post-feedback (Chapter 2, 3, and 4). To reduce the spatial dimensionality of the data, PCA was applied to a covariance matrix that was computed over the average beta power values for each subject at each channel and condition (NR-reward, NR-error, RR-

reward, RR-error, RE-reward, RE-error), yielding multiple "virtual channels" representing localized clusters of EEG activity that covary spatially across the scalp (Spencer et al., 2001). The first 6 virtual channels that together accounted for 80% of the total variance were submitted to Varimax rotation. For each subject and each condition, beta power was obtained for each virtual channel by taking the inner product of the factor loadings with the beta power values recorded for each channel.

Because this study was driven by a specific hypothesis on the role of beta oscillations in reinforcement learning and working memory, I limited my analyses to that frequency band. All errors reported for the behavioral data constitute standard deviations. Statistics on variables with more than two levels were corrected using the Greenhouse-Geisser correction for sphericity.

Results

Behavioral analysis

Average reaction times across subjects during the choice phase were not

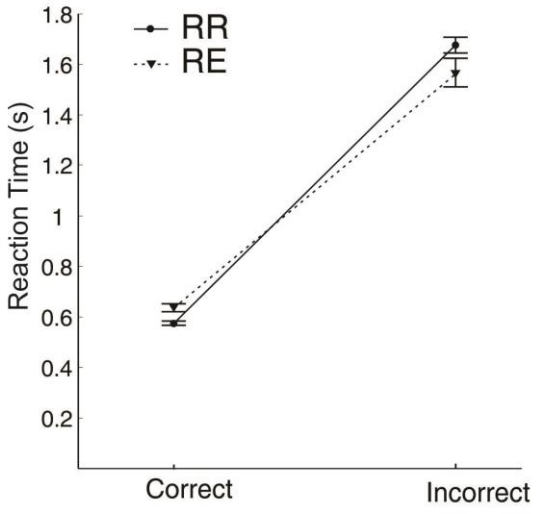

Figure 17: Reaction time at recall phase. Correct and incorrect recalls in the RR (remember rewards) and RE (remember errors) conditions are illustrated. Error bars show standard errors.

significantly different across conditions ($p > 0.05$) (491 ± 160 ms in the NR condition, 533 ± 145 ms in the RR condition, and 531 ± 161 ms in the RE condition). In the recall condition, a 2×2 ANOVA on reaction time with memory condition (RR, RE) and recall accuracy (correct, incorrect) revealed a significant effect of recall accuracy ($F(1,29) = 954.2$, $p < 0.001$) and a significant interaction between accuracy and memory condition ($F(1,29) = 19.7$, $p < 0.001$). Post-hoc paired t-tests revealed that in the RR condition subjects were significantly faster when correctly reporting the decks associated with reward feedback (571 ± 44 ms) compared to incorrectly reporting decks that were not associated with reward feedback (1674 ± 179 ms) ($t(29) = -40.5$, $p < 0.001$). Likewise, in the RE condition participants were significantly faster when correctly reporting the decks associated with error feedback (631 ± 88 ms) compared to incorrectly reporting the decks that were not associated with error feedback (1558 ± 314 ms) ($t(29) = -19.7$, $p < 0.001$). Further, correct recalls were significantly faster in the RR condition compared to the RE condition ($t(29) = -5.2$, $p < 0.001$) and incorrect recalls were significantly slower in the RR condition compared to the RE condition ($t(29) = 2.5$, $p = 0.019$) (Figure 17).

Recall accuracy was calculated as the percentage of correct recalls in each block averaged across the 20 blocks for the RR and RE conditions separately. Average recall accuracy was significantly higher in the RR condition (95.2 ± 4%) compared to the RE condition (90.7 ± 9%), ($t(29) = 3.3$, $p = 0.003$), indicating that participants were more accurate in recalling decks associated with reward feedback compared to decks associated with error feedback.

Time-frequency analysis

Figure 18A illustrates the scalp distributions of average beta power, 250-450 ms post-feedback, for the NR (left column), RR (middle column), and RE (right column)

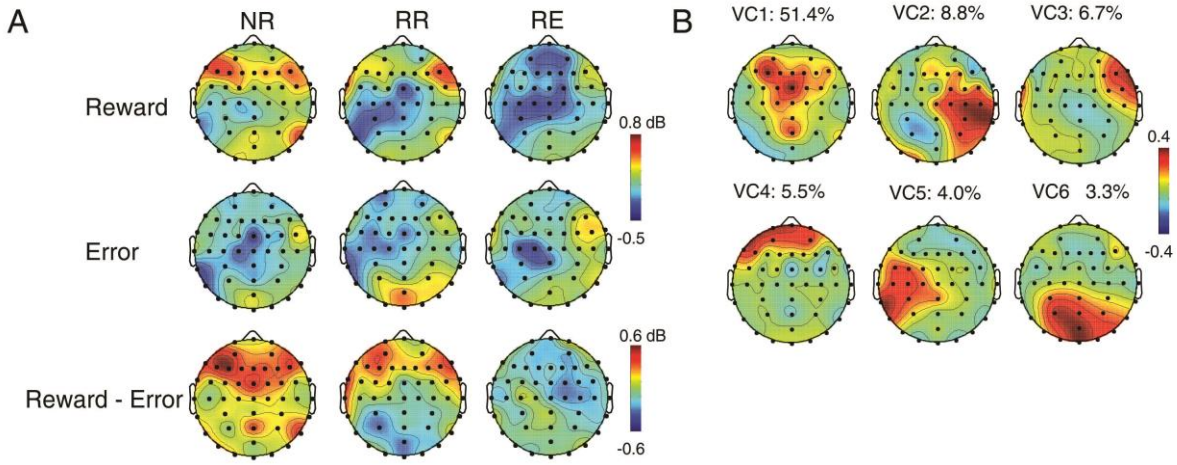

Figure 18: Scalp distribution of beta power. A) beta power averaged in 250-450 ms post-feedback associated with reward feedback (top), error feedback (middle) and the reward- error contrast (bottom), separately for the no-recall (NR, left), remember rewards (RR, middle), and remember errors (RE, right) conditions. B) Spatial PCA results illustrating six virtual channels explaining over 80% of the total variance in beta power.

conditions associated with reward feedback (top row), error feedback (middle row), and the contrast between these two feedback types (bottom row). Visual inspection indicates an increase in beta power over frontal areas of the scalp for rewards relative to errors in the NR and RR conditions but not in the RE condition (Figure 18A). Figure 19 provides time-frequency maps that illustrate the time course of beta at channel F3, where the beta contrast in reward vs. error was maximal in the NR and RR conditions; the dotted boxes indicate the time-frequency windows used for the scalp distributions in Figure 18A and for the statistical analyses.

Individual spatial factors were identified by submitting average beta power for every subject and condition to PCA using MATLAB. We retained the first 6 components that together explained over 80% of the total variance, comprising 6 'virtual channels' (Figure 18B); additional components were excluded from rotation because each added

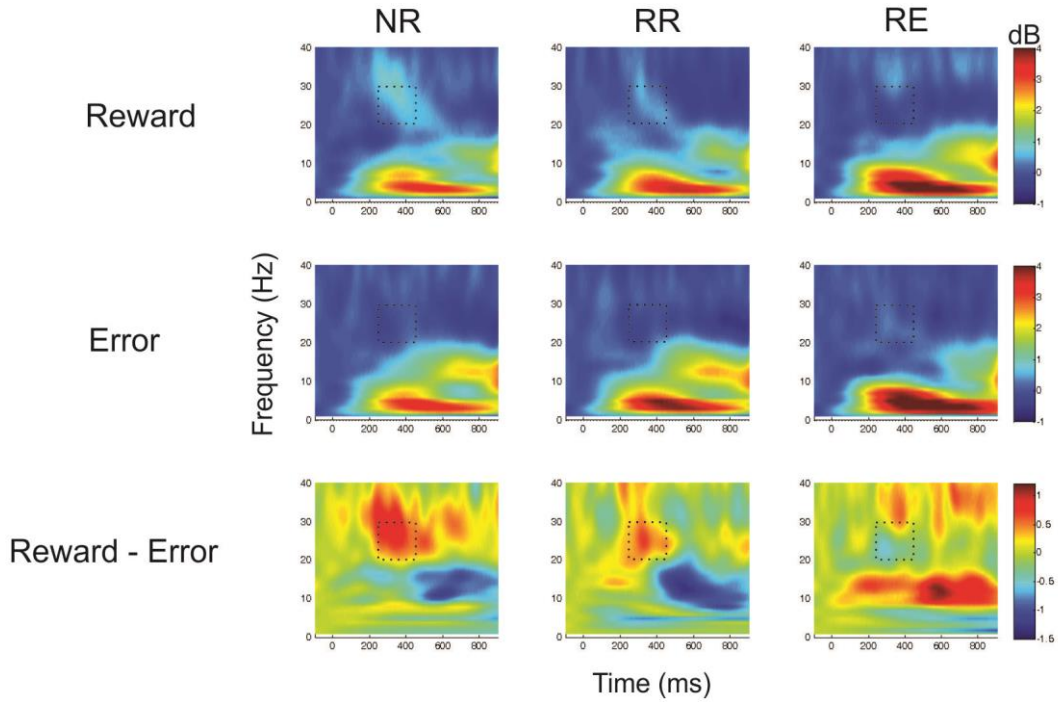

Figure 19: Time-frequency maps of EEG power recorded at channel F3 associated with reward (top row) and error (middle row) feedback, and the reward-error contrast (bottom row), for the no recall (NR, left column), remember rewards (RR, middle column), and remember errors (RE, right column) conditions.

less than 1% additional variance to the total (Figure 18B). Because reward-related beta oscillations concentrate over frontal areas of the scalp (Cunillera et al., 2012; HajiHosseini et al., 2012; Marco-Pallarés et al., 2008), I identified three virtual channels of interest with loadings that clustered over frontal areas: VC1 (frontal-central), VC3 (right frontal), and VC4 (frontal pole) (Figure 18B). A 3×3×2 ANOVA on beta power with virtual channel (VC1, VC3, VC4), memory condition (RR, RE, NR), and valence (reward, error) as factors revealed a significant main effect of valence ($F(1,29) = 4.25$, p

= 0.048) and an interaction between memory condition and valence ($F(2,58) = 3.48$, $p = 0.049$), with no main effects of virtual channel or memory. Due to the absence of a main effect or interaction with virtual channel, beta power was collapsed across VC1, VC3, and VC4 for the remaining analyses (Figure 20). Post-hoc paired t-tests revealed that beta power to reward feedback was larger in the NR condition than in the RE condition ($t(29) = 2.8$, $p = 0.01$), and marginally larger in the RR condition compared to RE condition ($t(29) = 1.9$, $p = 0.06$). Beta power to error feedback did not show any significant effect of memory condition ($p > 0.05$). Within the NR condition, beta power to reward feedback was significantly larger than beta power to error feedback ($t(29) = 2.6$, $p = 0.014$). Although beta power was larger to error feedback in the RE condition compared to the RR condition, this difference was not statistically significant ($p > .05$) (Figure 20).

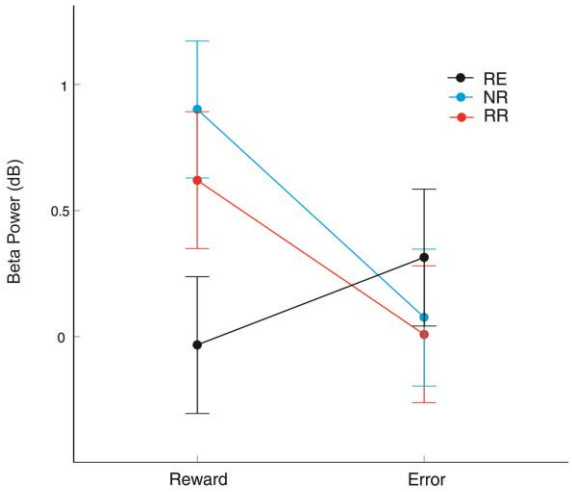

Figure 20: Beta power averaged across the three frontal virtual channels (VC1, VC3, and VC4) elicited by the Reward and Error feedback, in the three memory conditions: no recall (NR), remember rewards (RR), and remember errors (RE). Error bars show the 95% confidence intervals calculated by the pooled estimate of within-subject errors (Masson &

Beta power and behavior

When collapsed across the frontal virtual channels, post- reward-feedback beta power was negatively correlated with the reaction time for correctly recalled rewarded

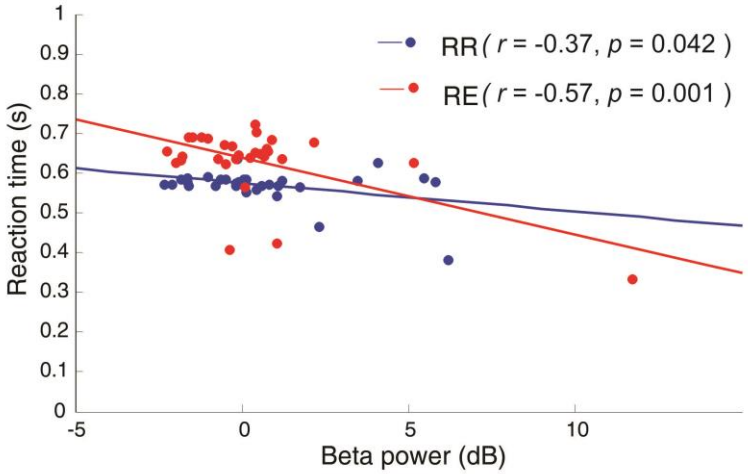

Figure 21: Relationship between beta power and speed of recall. Reaction time for correct recalls vs. beta power to reward feedback in the remember reward (RR) condition (blue) and beta power to error feedback in the remember error (RE) condition (red). Subjects were faster to recall target items correctly during the recall phase when the feedback to the target items elicited greater beta power in the previous choice phase.

trials in the RR condition (r = -0.37, p = 0.04). Further, post- error-feedback beta power was negatively correlated with the reaction time for correctly recalled error trials in the RE condition (r = -0.57, p = 0.001) (Figure 21). Although the latter correlation did not survive statistical significance when the data of an outlier (beta power > 10 dB) was removed, taken together these results suggest that across memory conditions the subjects who produced more beta power to feedback stimuli indicating the target decks (whether rewarded or not) were faster to recall correctly the target items during the recall phase.

Post-feedback beta power was not correlated with accuracy in either of the memory conditions (p > 0.05), perhaps because of a ceiling effect in accuracy levels (95.2 ± 4% in the RR and 90.7 ± 9% in the RE condition).

Discussion

Previous studies have observed that reward feedback stimuli elicit an increase in beta power over the frontal areas of the scalp around 250-450 ms post-feedback (Cohen et al., 2007; Cunillera et al., 2012; Marco-Pallares et al., 2008). Existing proposals associate this oscillatory activity with learning from rewards (Cohen et al., 2011) through an interplay between neurocognitive processes related to motivation, attention and memory (Marco-Pallarés et al., 2015). Previously, I source localized reward-related beta power to the DLPFC and proposed that it is related to maintenance and transfer of recently reinforced stimulus-response rules to brain areas that are responsible for task execution (Chapter 3). Further, I showed in Chapter 4 that frontal beta power is sensitive to the number of actions in action sequences as predicted by the oscillatory model for working memory (Lisman & Idiart, 1995; Lisman & Jensen, 2013).

Here, I showed that ‘reward-related beta power’ is not specific to reward feedback per se but rather depends on the informative value of the feedback stimuli for task performance. In a control condition, subjects were asked to perform a relatively typical reinforcement learning card choice task (NR), and in two other conditions, the task demands were modified by asking subjects to remember choices that were followed by either reward (RR) or error (RE) feedback for later recall. Consistent with previous studies, in the NR condition reward feedback compared to error feedback elicited more beta power distributed over frontal areas of the scalp (Figure 18A). Further, although this pattern was relatively unchanged by the instruction to remember rewards (RR), it was abolished by the instruction to remember errors (RE) such that reward feedback failed to elicit beta power over frontal cortex (Figure 19).

Instead, beta power was larger to error feedback compared to reward feedback in the RE condition, suggesting that beta can be elicited by error feedback stimuli when these stimuli convey important information for task performance. Although this comparison was not statistically significant, beta power was also inversely correlated with speed of correctly recalling both the stimuli associated with error feedback in the RE condition and the stimuli associated with reward feedback in the RR condition. This correlation between the speed of recall and beta power suggests that greater beta power to feedback reflects stronger encoding of the choices associated with rewards in the RR condition and with errors in the RE condition, implicating a broader function for beta beyond reward processing in particular.

The encoding process underlying later recall could be explained by at least three possible mechanisms. First, the items could be encoded via a simple reinforcement

learning process that strengthens individual action values. However, the striatal mechanisms underlying such associations are believed to require multiple pairings between actions and outcomes (Niv, 2009), but in this experiment each deck was selected only once on a given block of trials. Second, the information could be gated into working memory and then actively maintained as persistent activity in the PFC, or via cortico-striatal loops between the PFC and BG, until the end of the block (e.g., Frank, Loughry, & O'Reilly, 2001; Todd, Niv, & Cohen, 2008). Third, the deck choices could be rapidly encoded and subsequently retrieved from the hippocampus through PFC- hippocampus communication (Brincat & Miller, 2015; Preston & Eichenbaum, 2013). The last two possibilities are not mutually exclusive. Therefore, the findings from this and previous experiments (Chapters 3 and 4) suggest intercommunication between the PFC-BG and PFC-hippocampus.

This interpretation is in line with recordings from the principle sulci of monkeys, a homolog of human DLPFC (Goldman-Rakic, 1984), that have shown that beta oscillations reflect the activity of neural ensembles specific to the task at hand (Buschman et al., 2012), and that DLPFC neurons synchronize with striatal and hippocampal neurons when learning novel stimulus categories and associations (Antzoulatos & Miller, 2014; Brincat & Miller, 2015). In humans, joint analysis of EEG and functional magnetic resonance imaging in a gambling task also shows an association between beta power and the BOLD response in the striatum, hippocampus, and DLPFC that also supports the role of beta power in memory processes (Mas-Herrero et al., 2015). Further, frontal beta power increases during the maintenance period of a working memory task when performance is rewarded compared to when it is not (Kawasaki &

Yamaguchi, 2013). There is also evidence that frontal beta power increases with memory load (Howard et al., 2003), and that the theta-to-beta (low-gamma) frequency ratio is associated with working memory load (Axmacher et al., 2010) and working memory capacity (Kamiński, Brzezicka, & Wróbel, 2011).

Note that the homogeneity in reaction times during the choice phase suggests that subjects performed the reinforcement learning task in the same way across the three memory conditions. Therefore the memory instructions for the recall phase in the two recall conditions did not appear to affect choice performance. Further, post-feedback beta power did not correlate with recall accuracy, perhaps because of a ceiling effect in accuracy. Note that in order to maintain relatively equal difficulty across the reinforcement learning (choice phase) and working memory (recall phase) tasks, the deck locations remained constant on the screen throughout both choice and recall phases in each block, possibly explaining the relatively high recall accuracy in both conditions. Nevertheless, post-feedback beta power was inversely correlated with the speed of correctly recalled items in both the RR and the RE conditions, suggesting that post-feedback beta power contributes to faster item recall irrespective of feedback valence.

On the basis of these observations, I propose that frontal beta power reflects a general mechanism mediated by the DLPFC that 1) signals the need for active maintenance of recent actions underlying desired task performance, and 2) facilitates the communication of this information to other brain regions for future task execution. For this purpose, beta oscillations in the DLPFC communicate with the hippocampus regarding the maintenance of the information and with the BG regarding the execution. In conventional reinforcement learning paradigms, such as in the NR condition here, beta

power is elicited by reward feedback but not by error feedback because the former convey relatively more information for optimizing task performance. However, this process can be interrupted by explicit memory instructions that increase the utility of error feedback relative to reward feedback.

Chapter 6: General Discussion

Human frontal beta oscillations are poorly understood in terms of their functional significance and neural correlates. In particular, increased power of frontal beta oscillations has been observed following the delivery of reward or positive feedback stimuli in a variety of reinforcement learning tasks including gambling, time-estimation, probabilistic learning, and task switching paradigms (e.g., Cohen, Elger, & Ranganath, 2007; Marco-Pallares et al., 2008), yet a unified account on the significance of this phenomenon has yet to be developed. The two currently existing theories hinge on the well-known role of neural oscillations for communication across different brain areas (Cohen, Wilmes, & van de Vijver, 2011; Marco-Pallarés, Münte, & Rodríguez-Fornells, 2015). However, few of these ideas have been empirically tested.

The work of this dissertation was driven by a new hypothesis on the source and functional significance of frontal beta oscillations: I hypothesized that frontal beta oscillations reflect a general mechanism mediated by the DLPFC and ACC for the activation of neural ensembles related to actions immediately preceding feedback stimuli and maintenance and transfer of this information to other brain areas that are responsible for task execution. When feedback stimuli indicate that the preceding actions hold task-related information that are useful for achieving the desired performance, such as actions that lead to rewards in reinforcement learning paradigms, beta oscillations in the DLPFC reflect the communication of this information to areas responsible for task execution. Further, I proposed that beta oscillations in the ACC are related to the maintenance of information related to the formation and execution of hierarchical action sequences called

options. On this view, beta oscillations in the DLPFC and ACC govern the communication of task-related information to areas such as the BG for the purpose of enhancing task performance. This process is especially necessary when new rules are learned and diminishes as rule execution becomes overlearned. Once actions are executed automatically without the need from top-down control by the PFC, this communication would take place through alternative cortico-striatal loops established during learning (Antzoulatos & Miller, 2014).

This hypothesis was based on a few theories on PFC and ACC function in relation to goal-directed behavior and on the functional role of beta oscillations in communicating multi-item messages across the brain. The role of beta oscillations in the ACC in maintaining action sequences was interpreted based on the proposed role for this region in the HRL-ACC theory in the selection and maintenance of hierarchical action sequences called options. The role of beta oscillations in the DLPFC was interpreted based on the widely recognized function of the PFC in active representation of task-related information and monitoring and execution of task sets that bias task performance in other brain areas.

In what follows, I will summarize the findings of this dissertation as described in Chapters 2 to 5, explain how my proposal compares and contrasts with the other two accounts for the role of frontal beta oscillations, highlight unanswered questions, and discuss possible future directions.

Summary of the current findings

Because the burst of frontal EEG beta was originally observed in reinforcement learning paradigms, and because RPE signals are broadcast to frontal cortex by the midbrain dopamine system in reinforcement learning paradigms, I first examined whether these oscillations code for a dopamine-driven RPE signal. In Experiment One (Chapter 2) I tested the sensitivity of frontal beta power to an RPE signal, specifically, whether frontal beta power would reflect a specific interaction between outcome valence and probability. This study showed that unlike previous findings in a gambling task (HajiHosseini et al., 2012), in a time estimation task beta was not sensitive to outcome probability, which disconfirms its proposed relation to an RPE. Given that dopamine RPE signals appear to be the dominant reinforcement learning signal in the brain, this finding is in line with my hypothesis that beta power is not specifically related to reinforcement learning.

In Experiment Two, described in Chapter 3, a reinforcement learning paradigm with cue-response associations showed that frontal beta oscillations that are mainly sensitive to outcome valence were localized to the right DLPFC. This finding supported my hypothesis that beta oscillations originate in the DLPFC to send a top-down control signal to facilitate the communication of task-related information in order to enhance task performance. Further, the study replicated the insensitivity of this neural signal to the RPE in a new task (indicating that the absence of this effect is unlikely to be due to insufficient statistical power).

In Experiment Three, described in Chapter 4, I showed that frontal beta oscillations are sensitive to the length of the action sequence immediately preceding the feedback stimuli, and found coupling between beta power and theta phase following the

reward feedback in the task with a longer sequence. As in Experiment Two, I found greater beta power in the reward condition compared to the error condition in the DLPFC and further, greater beta power to feedback following a longer sequence of actions compared to feedback following a single action in the ACC. These findings, in accordance with the oscillatory model for communication of multi-item messages (Lisman & Jensen, 2013) and the HRL-ACC theory (Holroyd & Yeung, 2012), support my hypothesis on the association of frontal beta oscillations with a frontal mechanism governed by the DLPFC and ACC in maintaining and communicating task-related information to other brain areas.

In Experiment Four, described in Chapter 5, I showed that frontal beta oscillations that are elicited to reward-feedback stimuli in standard reinforcement learning tasks are not in fact specific to reward-related information; rather the reward-related beta oscillations constitute a special case of a more general signal that are sensitive to reward or error feedback depending on task demands. These frontal beta oscillations reflect top-down communication of the task-related information to areas that are responsible for task execution in order to reach desired task performance. In other words, the information communicated by frontal beta oscillations is not necessarily reward-related. This finding shows that a general theory about the functional significance of frontal beta oscillations that accounts for all of the empirical data must be achieved independent of reinforcement learning paradigms, leading to a more common account that can explain the processes governed by the PFC and ACC.

To my knowledge, these studies constitute the first hypothesis-driven experiments to address the source and functional significance of frontal beta oscillations. Therefore

the current findings make a valuable contribution to the literature on the function of PFC and ACC as understood according to the oscillatory behavior of neural networks.

Existing accounts

Existing theories on the role of frontal beta oscillations in feedback-guided learning require empirical support. Moreover, current models and theories of cognitive control in relation to DLPFC, ACC, and BG function are silent about the possible role of oscillations in goal-directed behavior. Therefore, the findings of this dissertation elucidate these previous accounts.

Learning from feedback

One of these theories holds that beta oscillations have a specific role in learning from rewards by synchronizing the activity of prefrontal cortical structures including the OFC, DLPFC, and motor cortex according to task demands (Cohen et al., 2011), whereas theta oscillations are specifically involved in learning from errors by synchronizing medial prefrontal and posterior cingulate cortices. Although the evidence provided in this dissertation is in line with the idea that the source of beta oscillations might depend on task demands, I found that beta is not a signature for learning from rewards per se as it does not necessarily depend on the reward value of the feedback. Therefore, the idea that learning from errors is driven by theta oscillations whereas learning from rewards is driven by beta oscillations is not correct as these oscillations can be modified by a more general mechanism for communication across brain areas rather than pure reinforcement learning (see also Holroyd, HajiHosseini, & Baker, 2012).

Interaction between cognitive functions

A recent proposal by Marco-Pallares and colleagues posits that beta oscillations that are elicited to unexpected rewards send a fast motivational signal to other brain structures to communicate information about changes in the environment and the need to adapt to new contingencies. On this view, beta oscillations might have a key role in mediating 'cross-talk' between attention and memory processes. They also proposed the source of beta oscillations to be the ventromedial PFC or the dorsal ACC. But it is not clear how these possible sources govern this cross-talk across brain areas and how that is related to cognitive behavior. My current findings showed that beta oscillations are produced in the DLPFC and ACC depending on task demands and are sensitive to task-related information irrespective of feedback valence. These findings provide evidence on the association of frontal beta with communication of task-related information that need to be communicated and maintained for behavioral adjustment, which indicates the involvement of a working memory process. Therefore, my proposal is in line with a proposed role in working memory but it clearly differentiates between the sources of beta power depending on the task demands and explains how this might be related to the functional role of the DLPFC and ACC in goal-directed behavior.

Frontal beta, PFC, ACC, learning, and memory

Table 1 summarizes the studies where frontal beta oscillations were elicited to reward feedback stimuli including the four experiments described in this dissertation. This summary shows a consistent effect of valence that is distributed over frontal-central and frontal-lateral areas of the scalp, and which is consistently interpreted as reflecting enhanced synchronization of neural activity at this frequency. This table also indicates

Table 1: Electroencephalogram studies of frontal beta oscillations. A list of studies, tasks, parameters, sensitivity of frontal beta, relation with behavior, source and/or distribution, and the interpretation of authors about beta function is presented.

Study	Task	Parameters	Sensitivity	Behavior	Source/distribution	Interpretation
Cohen et al. (2007)	probabilistic learning	probability, valence	probability	-	frontal-central	enhanced connectivity following gains
Marco-Pallares et al., (2008)	gambling	magnitude, valence	magnitude, valence	-	frontal-central, left-lateralized	synchronization in brain reward network
van de Vijver et al. (2011)*	time-estimation	valence, next trial performance	valence, next trial performance	correlated with next trial performance	medial-frontal	continuation of the current state after positive feedback
HajiHosseini et al. (2012)	gambling	magnitude, probability, valence	probability, valence	-	frontal-central, right-lateralized	prediction error only for gains, fronto-striatal interactions in reward network
Cunillera, et al., (2012)	Wisconsin Card Sort Test	valence, rule change	valence, rule change	-	frontal, right-lateralized	processing of novel task-relevant information after positive feedback for behavioral adjustment
Mas-Herrero et al. (2015)	gambling	probability, valence	valence	-	frontal-central, right-lateralized	coordination of reward network
Experiment One	time-estimation	probability, valence	valence	-	frontal –lateral, frontal-central	not a reward prediction error signal, frontal mechanism for boosting reward-related information
Experiment Two	T-maze	probability, valence	valence	-	rDLPFC/frontal-lateral	boosting reward-related information in working memory
Experiment Three	R-maze and H-maze	valence, memory load	valence, memory load	predicting correct performance in the H-maze	lDLPFC, ACC/frontal-central, frontal-lateral	encoding and maintenance of reward-related information
Experiment Four	card choice	valence, memory condition	valence× memory condition	correlated with the speed of correct recall	frontal-central, frontal-lateral	encoding of task-related information, working memory

* time window for beta analysis was 300-800 ms post-feedback

that the experiments presented in this thesis have provided evidence on the neural source of beta oscillations and on specific predictions regarding task demands on working memory.

There is evidence from LFP recordings in non-human primates on the relationship between beta oscillations in the DLPFC in the creation of task-specialized neuronal ensembles. Studies have also shown synchrony in the beta frequency band between neurons in the PFC and the striatum during category learning, which is argued to contribute to the creation of task-related functional-anatomical loops between these areas (Antzoulatos & Miller, 2014). It has also been shown that when monkeys learn novel stimulus associations, neural synchrony between the PFC and hippocampus in lower beta and theta band oscillations are related to correct and incorrect responses, respectively (Brincat & Miller, 2015).

Therefore, it is possible that EEG frontal beta oscillations observed in the studies presented in Table 1 relate to learning by encoding and maintenance of the most recent task-related information through communication between the PFC, striatum, and hippocampus. When task demands are low or when behavior becomes automatic, there is less need for a top-down signal for communicating task-related information and beta power diminishes.

Other considerations

It is important to note that the association of beta with the 'working memory' demands of the task in my proposal does not mean that specific stimulus-related information is stored in the PFC or ACC (e.g., a sequence of stimuli representing the alleys). Rather, the task-related information attributed to these brain areas refers to the

abstract rules that are created when subjects encounter stimulus-response mappings useful for desired task performance (e.g., right-left-right). These abstract rules, which are created and maintained jointly by the PFC and ACC, are applied in a top-down manner over other brain areas that are responsible for task execution such as the BG, and for long-term storage and retrieval such as the hippocampus. In line with this idea, single-unit recordings from monkeys have recently shown that the firing of PFC neurons is not related to encoding the specific stimulus features, such as color, in working memory but instead is more related to resource allocation via long-range communication with other brain areas (Lara & Wallis, 2014). This idea converges with other theoretical frameworks proposed for working memory and PFC function that propose a role for the PFC in biasing the execution of stimulus-response mappings by other brain structures (Miller & Cohen, 2001), and in linking recent stimulus-related information with forthcoming actions (Fuster & Bressler, 2012).

Empirical support for this proposal is provided by my experiments presented in this thesis that associate frontal beta power with the working memory demands of the task and evidence for the involvement of DLPFC and ACC in this process. Other aspects of my proposal -- such as whether frontal beta is related to the formation of cortico-striatal and hippocampal loops that encode task-related abstract rules, and how ACC and DLPFC communicate for the purpose of maintaining action sequences -- are inferences based on the previous literature that remain to be tested in future experiments.

Further, beta oscillations are also proposed to be related to dopamine levels. To be specific, it has been proposed that salient events in the environment induce phasic bursts of dopamine activity in the midbrain and BG leading to a rise in dopamine level, thereby

providing an index of the likelihood of the need for new voluntary action (Jenkinson & Brown, 2011). Loss of dopamine, as in Parkinson's disease, disrupts this function and contributes to excessive beta oscillations in the BG and cortex (Little & Brown, 2014). Also, dopamine medication can attenuate beta oscillations recorded in the BG in Parkinson's disease, but the cellular mechanisms of how dopamine level contributes to excessive beta oscillations is still largely unknown (Hammond et al., 2007). Consequently, it is not possible to make any interpretations about a direct relationship between dopamine activity and frontal beta oscillations regarding their function in working memory and goal-directed behavior according to the experiments presented in this thesis.

The findings presented in this dissertation can open new research avenues on the possibility of using these oscillations to guide therapeutic treatments for neuropsychiatric disorders associated with abnormal beta oscillations and cognitive and motor behavior. This information can be potentially used in non-invasive brain stimulation methods as treatments for disorders such as Parkinson's disease. I will explain one of these ideas in the next section.

Future directions

The association of oscillatory bands with distinct cognitive roles that I reviewed in Chapter 1 might raise the question of whether other oscillations can govern opposite effects on learning and memory compared to what I propose for frontal beta oscillations. Although the experiments reported in this thesis were designed based on a hypothesis on frontal beta oscillations, they also showed increased alpha power following error

(negative) feedback in Experiment Three (Figure 10) and following reward feedback in the remember error condition in Experiment Four (Figure 19). One of the avenues to pursue in the future might be to test whether oscillations in the alpha frequency band inhibit the activation of the information related to behavior that is *not* useful for desired task performance such as erroneous responses. This idea is based on the suggested role for alpha oscillations in inhibiting task-irrelevant rules and regions (Buschman et al., 2012; Jensen & Mazaheri, 2010). But whether there is evidence for such function for EEG alpha oscillations in reinforcement learning is an open question.

As mentioned above, this information can be useful for therapeutic purposes using non-invasive brain stimulation techniques. Below I present ideas for future research using the findings of this thesis on frontal beta to develop a novel, non-invasive treatment for Parkinson's disease.

Parkinson's disease is caused by degeneration of dopamine neurons in a collection of subcortical nuclei called the substantia nigra pars compacta and ventral tegmental area, which produce and distribute dopamine widely across the brain, including the BG and PFC. Direct projection of dopamine to the BG underlies movement regulation by these nuclei. Influential hypotheses of BG function propose a role in action selection and suppression of competing actions (Mink, 1996), adjustment of movement scale (Schmidt et al., 2008), and motor learning (Doyon et al., 2009) that could explain motor symptoms of Parkinson's disease such as rigidity, slowness of movement, and tremor due to impaired dopamine projection to this area. Also, some of the non-motor features of Parkinson's disease such as cognitive dysfunction, depression, anxiety, and apathy are likely related to dopamine depletion in the striatum, part of the BG network,

and the PFC. Important dopamine mediated cognitive functions such as reinforcement learning (Schultz et al., 1997), and working memory (Seamans & Yang, 2004) are impaired in Parkinson's disease (Beato et al., 2008; Frank, Samanta, Moustafa, & Sherman, 2007; Frank et al., 2004; Lee et al., 2010; Lewis et al., 2005). In this thesis, I showed that working memory aspects of a task can affect an oscillatory signature of reward, suggesting an interaction between different cognitive processes in learning from outcomes. This also suggests a relation between the working memory impairments and dysfunctional reinforcement learning in Parkinson's disease.

However, motor symptoms that are the main focus in Parkinson's disease treatments are also a result of impaired movement regulation using motivational/incentive factors by the BG. To be specific, it is proposed that movement performance or 'vigor' is adjusted by context-specific cost and reward functions (Turner & Desmurget, 2010). This proposal is supported by studies showing that patients with BG lesions are able to control grip forces in response to explicit instructions but do not increase grip forces despite understanding that higher forces will obtain greater monetary rewards (Schmidt et al., 2008). This has led to the suggestion that slowness of movement in Parkinson's disease is due to the lack of implicit motivation to expend more energy caused by the impaired link between motivation and control of movement speed and size (Mazzoni, Hristova, & Krakauer, 2007). Therefore alleviation of motor symptoms could be achieved by manipulation of the signals that carry such information related to motivation and rewards inside the brain.

As mentioned above, one of the pathological signs of Parkinson's disease is exaggerated beta band oscillations (Moran et al., 2011) in the BG and cortex (Brittain &

Brown, 2014). Beta oscillations are suggested to be an indication of dopamine levels at the sites of cortical inputs to BG (Little & Brown, 2014). Further, DBS treatments are based on overcoming the excessive beta in the subthalamic nucleus and globus pallidus, parts of a broader BG network. The causal relationship between cortical and BG beta is not very clear but it is proposed that dopamine depletion may cause increased connectivity between the cortex and BG and therefore communicate the pathological beta rhythm to downstream areas (Little & Brown, 2014). Also, non-motor Parkinson's disease symptoms indicate impaired cognitive functions normally governed by the PFC. Despite this intriguing evidence, the role of the PFC in Parkinson's disease has not been studied well as most treatments target the BG.

Non-invasive brain stimulation techniques have been used as therapeutic treatments over the years (Utz, Dimova, Oppenländer, & Kerkhoff, 2010) although their underlying mechanisms of action are becoming evident only recently (Romero Lauro et al., 2014). In particular, noisy galvanic vestibular stimulation (nGVS) is reported to enhance cognitive abilities (Wilkinson, Nicholls, Pattenden, Kilduff, & Milberg, 2008) and improve motor responsiveness in Parkinson's disease (Yamamoto, Struzik, Soma, Ohashi, & Kwak, 2005). It is shown that nGVS modifies EEG oscillations in different frequency bands and notably enhances beta oscillations in frontal regions (Kim et al., 2013). It is also shown that nGVS can specifically improve the vigor of corrective sub-movements that are impaired in Parkinson's disease during manual tracking performance. Effective therapeutic stimulation can be designed based on the theoretical basis of dopamine dysfunction and its effects on oscillatory signatures. Therefore one can investigate how the oscillatory signatures of goal-directed behavior as described in this

dissertation can guide therapeutic brain stimulation for Parkinson's disease. This is specifically interesting because of dopamine deficiency accompanied by abnormal beta oscillations in this disorder.

I proposed in this dissertation that beta is related to a general mechanism that guides learning from rewards as a special case, which suggests that beta must affect the computation of movement vigor too. Therefore pathological beta in Parkinson's disease can be associated with impaired calculation of movement vigor. But is there is an association between beta and behavior in Parkinson's disease? Can brain stimulation by nGVS modify beta as an effective treatment for Parkinson's disease? I hypothesize that manipulation of pathological beta through brain stimulation might affect reward-related beta and as a result compensate for impaired vigor adjustments in Parkinson's disease. These ideas can be tested through empirical research paradigms that can potentially contribute to customized treatments for individual patients and elucidate the theoretical basis of functional role of these oscillations.

Concluding remarks

Brain oscillations that are non-invasively recorded from the scalp provide substantial information about neurocognitive brain processes. With the availability of advanced analysis techniques, the association of these oscillations with cognitive and motor behavior can be studied in various paradigms. Here I focused on human frontal beta oscillations recorded by EEG and showed that this oscillatory activity reflects important information about learning and memory processes governed by the DLPFC and ACC in goal-directed behavior. This information can contribute to a better understanding

of the neural mechanisms of goal-directed behavior that is grounded in theory, and can be used in clinical applications such as in brain stimulation techniques for therapeutic purposes. Further research on EEG oscillations driven by theories of cortical function can elucidate additional aspects of cognitive and motor behavior.

Bibliography

Antzoulatos, E. G., & Miller, E. K. (2014). Increases in functional connectivity between prefrontal cortex and striatum during category learning. *Neuron*, *83*(1), 216–225. doi:10.1016/j.neuron.2014.05.005

Asada, H., Fukuda, Y., Tsunoda, S., Yamaguchi, M., & Tonoike, M. (1999). Frontal midline theta rhythms reflect alternative activation of prefrontal cortex and anterior cingulate cortex in humans. *Neuroscience Letters*, *274*(1), 29–32. doi:10.1016/S0304-3940(99)00679-5

Axmacher, N., Henseler, M. M., Jensen, O., Weinreich, I., Elger, C. E., & Fell, J. (2010). Cross-frequency coupling supports multi-item working memory in the human hippocampus. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(7), 3228–3233. doi:10.1073/pnas.0911531107

Baker, T. E., & Holroyd, C. B. (2009). Which way do I go? neural activation in response to feedback and spatial processing in a virtual t-maze. *Cerebral Cortex*. doi:10.1093/cercor/bhn223

Baker, T. E., & Holroyd, C. B. (2013). The topographical N170: Electrophysiological evidence of a neural mechanism for human spatial navigation. *Biological Psychology*, *94*(1), 90–105. doi:10.1016/j.biopsycho.2013.05.004

Barbey, A. K., Koenigs, M., & Grafman, J. (2013). Dorsolateral prefrontal contributions to human working memory. *Cortex*, *49*, 1195–1205. doi:10.1016/j.cortex.2012.05.022

Başar, E., Başar-Eroglu, C., Karakaş, S., & Schürmann, M. (1999). Are cognitive processes manifested in event-related gamma, alpha, theta and delta oscillations in the EEG? *Neuroscience Letters*, *259*(3), 165–168. doi:10.1016/S0304-3940(98)00934-3

Başar, E., Başar-Eroglu, C., Karakaş, S., & Schürmann, M. (2000). Gamma, alpha, delta, and theta oscillations govern cognitive processes. *International Journal of Psychophysiology*, *39*(2-3), 241–248. doi:10.1016/S0167-8760(00)00145-8

Basar-Eroglu, C., Brand, A., Hildebrandt, H., Karolina Kedzior, K., Mathes, B., & Schmiedt, C. (2007). Working memory related gamma oscillations in schizophrenia patients. *International Journal of Psychophysiology*, *64*(1), 39–45. doi:10.1016/j.ijpsycho.2006.07.007

Beato, R., Levy, R., Pillon, B., Vidal, C., Du Montcel, S. T., Deweer, B., … Cardoso, F. (2008). Working memory in Parkinson's disease patients: Clinical features and response to levodopa. *Arquivos de Neuro-Psiquiatria*, *66*(2 A), 147–151. doi:10.1590/S0004-282X2008000200001

Benchenane, K., Tiesinga, P. H., & Battaglia, F. P. (2011). Oscillations in the prefrontal cortex: A gateway to memory and attention. *Current Opinion in Neurobiology*, *21*(3), 475–485. doi:10.1016/j.conb.2011.01.004

Berke, J. D., Hetrick, V., Breck, J., & Greene, R. W. (2008). Transient 23-30 Hz oscillations in mouse hippocampus during exploration of novel environments. *Hippocampus*, *18*(5), 519–529. doi:10.1002/hipo.20435

Bernat, E. M., Nelson, L. D., Holroyd, C. B., Gehring, W. J., & Patrick, C. J. (2008). Separating cognitive processes with principal components analysis of EEG time-frequency distributions. *Proceedings of SPIE*, *7074*, 70740S–70740S–10. doi:10.1117/12.801362

Botvinick, M. M., Niv, Y., & Barto, A. C. (2009). Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. *Cognition*, *113*(3), 262–280. doi:10.1016/j.cognition.2008.08.011

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. doi:10.1163/156856897X00357

Braver, T., & Cohen, J. (2000). On the control of control: The role of dopamine in regulating prefrontal function and working memory. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 713–736). Cambridge, MA: MIT Press.

Brincat, S. L., & Miller, E. K. (2015). Frequency-specific hippocampal-prefrontal interactions during associative learning. *Nature Neuroscience*, *18*(4), 576–581. doi:10.1038/nn.3954

Brittain, J. S., & Brown, P. (2014). Oscillations and the basal ganglia: Motor control and beyond. *NeuroImage*, *85*, 637–647. doi:10.1016/j.neuroimage.2013.05.084

Buffalo, E. A., Fries, P., Landman, R., Buschman, T. J., & Desimone, R. (2011). Laminar differences in gamma and alpha coherence in the ventral stream. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(27), 11262–11267. doi:10.1073/pnas.1011284108

Buschman, T. J., Denovellis, E. L., Diogo, C., Bullock, D., & Miller, E. K. (2012). Synchronous oscillatory neural ensembles for rules in the prefrontal cortex. *Neuron*, *76*(4), 838–46. doi:10.1016/j.neuron.2012.09.029

Buzsáki, G. (2006). *Rhythms of the brain*. New York: Oxford Univ Press.

Buzsáki, G., & Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science*, *304*(5679), 1926–1929. doi:10.1126/science.1099745

Canolty, R. T., Edwards, E., Dalal, S. S., Soltani, M., Nagarajan, S. S., Kirsch, H. E., … Knight, R. T. (2006). High gamma power is phase-locked to theta oscillations in human neocortex. *Science*, *313*(5793), 1626–1628. doi:10.1126/science.1128115

Caplin, A., & Dean, M. (2008). Axiomatic methods, dopamine and reward prediction error. *Current Opinion in Neurobiology*, *18*(2), 197–202. doi:10.1016/j.conb.2008.07.007

Cavanagh, J. F. (2015). Cortical delta activity reflects reward prediction error and related behavioral adjustments, but at different times. *NeuroImage*, *110*, 205–216. doi:10.1016/j.neuroimage.2015.02.007

Cavanagh, J. F., Eisenberg, I., Guitart-Masip, M., Huys, Q., & Frank, M. J. (2013). Frontal theta overrides pavlovian learning biases. *The Journal of Neuroscience*, *33*(19), 8541–8. doi:10.1523/JNEUROSCI.5754-12.2013

Cavanagh, J. F., Figueroa, C. M., Cohen, M. X., & Frank, M. J. (2012). Frontal theta reflects uncertainty and unexpectedness during exploration and exploitation. *Cerebral Cortex*, *22*(11), 2575–86. doi:10.1093/cercor/bhr332

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, *18*(8), 414–421. doi:10.1016/j.tics.2014.04.012

Cavanagh, J. F., Frank, M. J., Klein, T. J., & Allen, J. J. B. (2010). Frontal theta links prediction errors to behavioral adaptation in reinforcement learning. *NeuroImage*, *49*(4), 3198–209. doi:10.1016/j.neuroimage.2009.11.080

Cavanagh, J. F., Zambrano-Vazquez, L., & Allen, J. J. B. (2012). Theta lingua franca: A common mid-frontal substrate for action monitoring processes. *Psychophysiology*, *49*(2), 220–238. doi:10.1111/j.1469-8986.2011.01293.x

Chen, C. C., Brücke, C., Kempf, F., Kupsch, A., Lu, C. S., Lee, S. T., … Brown, P. (2006). Deep brain stimulation of the subthalamic nucleus: A two-edged sword. *Current Biology*, *16*(22). doi:10.1016/j.cub.2006.10.013

Chrobak, J. J., & Buzsáki, G. (1998). Gamma oscillations in the entorhinal cortex of the freely behaving rat. *The Journal of Neuroscience*, *18*(1), 388–398. doi:0270-6474/97/180388-11$05.00/0

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychological Review, 97*, 332–361. doi:10.1037/0033-295X.97.3.332

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

Cohen, M. X., & Cavanagh, J. F. (2011). Single-trial regression elucidates the role of prefrontal theta oscillations in response conflict. *Frontiers in Psychology*, *2*, 30. doi:10.3389/fpsyg.2011.00030

Cohen, M. X., Elger, C. E., & Ranganath, C. (2007). Reward expectation modulates feedback-related negativity and EEG spectra. *NeuroImage*, *35*(2), 968–78. doi:10.1016/j.neuroimage.2006.11.056

Cohen, M. X., Wilmes, K. A., & van de Vijver, I. (2011). Cortical electrophysiological network dynamics of feedback learning. *Trends in Cognitive Sciences*, *15*(12), 558–566. doi:10.1016/j.tics.2011.10.004

Collins, A. G. E., Cavanagh, J. F., & Frank, M. J. (2014). Human EEG uncovers latent generalizable rule structure during learning. *The Journal of Neuroscience*, *34*(13), 4677–85. doi:10.1523/JNEUROSCI.3900-13.2014

Collins, A. G. E., & Frank, M. J. (2012). How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis. *European Journal of Neuroscience*, *35*(7), 1024–35. doi:10.1111/j.1460-9568.2011.07980.x

Cunillera, T., Fuentemilla, L., Periañez, J., Marco-Pallarès, J., Krämer, U. M., Càmara, E., … Rodríguez-Fornells, A. (2012). Brain oscillatory activity associated with task switching and feedback processing. *Cognitive, Affective & Behavioral Neuroscience*, *12*(1), 16–33. doi:10.3758/s13415-011-0075-5

D'Ardenne, K., Eshel, N., Luka, J., Lenartowicz, A., Nystrom, L. E., & Cohen, J. D. (2012). Role of prefrontal cortex and the midbrain dopamine system in working memory updating. *Proceedings of the National Academy of Sciences*, *109*(49), 19900–19909. doi:10.1073/pnas.1116727109

Dang-Vu, T. T., Desseilles, M., Laureys, S., Degueldre, C., Perrin, F., Phillips, C., … Peigneux, P. (2005). Cerebral correlates of delta waves during non-REM sleep revisited. *NeuroImage, 28*(1), 14–21. doi:10.1016/j.neuroimage.2005.05.028

Davis, N. J., Tomlinson, S. P., & Morgan, H. M. (2012). The role of beta-frequency neural oscillations in motor control. *The Journal of Neuroscience*. doi:10.1523/JNEUROSCI.5106-11.2012

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *The Journal of Neuroscience Methods*, *134*, 9–21. doi:10.1016/j.jneumeth.2003.10.009

Doñamayor, N., Marco-Pallarés, J., Heldmann, M., Schoenfeld, M. A., & Münte, T. F. (2011). Temporal dynamics of reward processing revealed by magnetoencephalography. *Human Brain Mapping*, *32*, 2228–40. doi:10.1002/hbm.21184

Dosenbach, N. U. F., Visscher, K. M., Palmer, E. D., Miezin, F. M., Wenger, K. K., Kang, H. C., … Petersen, S. E. (2006). A core system for the implementation of task sets. *Neuron*, *50*(5), 799–812. doi:10.1016/j.neuron.2006.04.031

Doyon, J., Bellec, P., Amsel, R., Penhune, V., Monchi, O., Carrier, J., … Benali, H. (2009). Contributions of the basal ganglia and functionally related brain structures to motor learning. *Behavioural Brain Research*, *199*(1), 61–75. doi:10.1016/j.bbr.2008.11.012

Durstewitz, D., & Seamans, J. K. (2008). The dual-state theory of prefrontal cortex dopamine function with relevance to catechol-o-methyltransferase genotypes and schizophrenia. *Biological Psychiatry*, *64*(9), 739–749. doi:10.1016/j.biopsych.2008.05.015

Engel, A. K., & Fries, P. (2010). Beta-band oscillations--signalling the status quo? *Current Opinion in Neurobiology*, *20*(2), 156–65. doi:10.1016/j.conb.2010.02.015

Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top-down processing. *Nature Reviews. Neuroscience*, *2*(10), 704–716. doi:10.1038/35094565

Feingold, J. (2011). *Beta oscillations in frontal cortex and striatum represent post-processing of successful behavior*. Harvard University--MIT Division of Health Sciences and Technology. Retrieved from http://hdl.handle.net/1721.1/68453

Feurra, M., Bianco, G., Santarnecchi, E., Del Testa, M., Rossi, A., & Rossi, S. (2011). Frequency-dependent tuning of the human motor system induced by transcranial oscillatory potentials. *The Journal of Neuroscience*, *31*(34), 12165–12170. doi:10.1523/JNEUROSCI.0978-11.2011

Frank, M. J. (2005). Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *Journal of Cognitive Neuroscience*, *17*, 51–72. doi:10.1162/0898929052880093

Frank, M. J., Loughry, B., & O'Reilly, R. C. (2001). Interactions between frontal cortex and basal ganglia in working memory: a computational model. *Cognitive, Affective & Behavioral Neuroscience*, *1*(2), 137–60. doi:10.3758/CABN.1.2.137

Frank, M. J., Samanta, J., Moustafa, A. a, & Sherman, S. J. (2007). Hold your horses: impulsivity, deep brain stimulation, and medication in parkinsonism. *Science*, *318*(5854), 1309–1312. doi:10.1126/science.1146157

Frank, M. J., Seeberger, L. C., & O'reilly, R. C. (2004). By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science*, *306*(5703), 1940–3. doi:10.1126/science.1102941

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, *9*(10), 474–480. doi:10.1016/j.tics.2005.08.011

Fries, P., Nikolić, D., & Singer, W. (2007). The gamma cycle. *Trends in Neurosciences*, *30*(7), 309–316. doi:10.1016/j.tins.2007.05.005

Funahashi, S. (2006). Prefrontal cortex and working memory processes. *Neuroscience*, *139*(1), 251–261. doi:10.1016/j.neuroscience.2005.07.003

Fuster, J. M. (2008). *The Prefrontal Cortex* (4th editio., Vol. 1). Academic Press.

Fuster, J. M., & Alexander, G. E. (1971). Neuron activity related to short-term memory. *Science*, *173*(997), 652–654. doi:10.1126/science.173.3997.652

Fuster, J. M., & Bressler, S. L. (2012). Cognit activation: A mechanism enabling temporal integration in working memory. *Trends in Cognitive Sciences*, *16*(4), 207–217. doi:10.1016/j.tics.2012.03.005

Gais, S., & Born, J. (2004). Declarative memory consolidation: Mechanisms acting during human sleep. *Learning & Memory*, *11*(6), 679–685. doi:10.1101/lm.80504

Genzel, L., Kroes, M. C. W., Dresler, M., & Battaglia, F. P. (2014). Light sleep versus slow wave sleep in memory consolidation: A question of global versus local processes? *Trends in Neurosciences*, *37*(1), 10–19. doi:10.1016/j.tins.2013.10.002

Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, *66*(4), 585–595. doi:10.1016/j.neuron.2010.04.016

Gold, J. M., Waltz, J. A., Prentice, K. J., Morris, S. E., & Heerey, E. A. (2008). Reward processing in schizophrenia: A deficit in the representation of value. *Schizophrenia Bulletin*, *34*(5), 835–847. doi:10.1093/schbul/sbn068

Goldman-Rakic, P. S. (1984). Modular organization of prefrontal cortex. *Trends in Neurosciences*, *7*(11), 419–424. doi:10.1016/S0166-2236(84)80146-0

Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*, 468–484. doi:10.1016/0013-4694(83)90135-9

Grossberg, S. (2009). Beta oscillations and hippocampal place cell learning during exploration of novel environments. *Hippocampus*, *19*(9), 881–885. doi:10.1002/hipo.20602

HajiHosseini, A., & Holroyd, C. B. (2013). Frontal midline theta and N200 amplitude reflect complementary information about expectancy and outcome evaluation. *Psychophysiology*, *50*(6), 550–562. doi:10.1111/psyp.12040

HajiHosseini, A., & Holroyd, C. B. (2015a). Reward feedback stimuli elicit high-beta EEG oscillations in human dorsolateral prefrontal cortex. *Scientific Reports*. doi:10.1038/srep13021

HajiHosseini, A., & Holroyd, C. B. (2015b). Sensitivity of frontal beta oscillations to reward valence but not probability. *Neuroscience Letters*, *602*, 99–103. doi:10.1016/j.neulet.2015.06.054

HajiHosseini, A., Rodríguez-Fornells, A., & Marco-Pallarés, J. (2012). The role of beta-gamma oscillations in unexpected rewards processing. *NeuroImage*, *60*(3), 1678–85. doi:10.1016/j.neuroimage.2012.01.125

Hälbig, T. D., Tse, W., Frisina, P. G., Baker, B. R., Hollander, E., Shapiro, H., … Olanow, C. W. (2009). Subthalamic deep brain stimulation and impulse control in Parkinson's disease. *European Journal of Neurology*, *16*(4), 493–497. doi:10.1111/j.1468-1331.2008.02509.x

Hammond, C., Bergman, H., & Brown, P. (2007). Pathological synchronization in Parkinson's disease: networks, models and treatments. *Trends in Neurosciences*, *30*(7), 357–364. doi:10.1016/j.tins.2007.05.004

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709. doi:10.1037//0033-295X.109.4.679

Holroyd, C. B., HajiHosseini, A., & Baker, T. E. (2012). ERPs and EEG oscillations, best friends forever: Comment on Cohen et al. *Trends in Cognitive Sciences*, *16*(4), 192. doi:10.1016/j.tics.2012.02.008

Holroyd, C. B., & Krigolson, O. E. (2007). Reward prediction error signals associated with a modified time estimation task. *Psychophysiology*, *44*, 913–917. doi:10.1111/j.1469-8986.2007.00561.x

Holroyd, C. B., Krigolson, O. E., Baker, R., Lee, S., & Gibson, J. (2009). When is an error not a prediction error? An electrophysiological investigation. *Cognitive, Affective & Behavioral Neuroscience*, *9*, 59–70. doi:10.3758/CABN.9.1.59

Holroyd, C. B., Krigolson, O. E., & Lee, S. (2011). Reward positivity elicited by predictive cues. *Neuroreport*, *22*, 249–252. doi:10.1097/WNR.0b013e328345441d

Holroyd, C. B., & McClure, S. M. (2015). Hierarchical control over effortful behavior by rodent medial frontal cortex. *Psychological Review*, *122*(1), 1–107. doi:10.1037/a0038339

Holroyd, C. B., Nieuwenhuis, S., Yeung, N., & Cohen, J. D. (2003). Errors in reward prediction are reflected in the event-related brain potential. *Neuroreport*, *14*(18), 2481–4. doi:10.1097/01.wnr.0000099601.41403.a5

Holroyd, C. B., Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-related positivity: sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology*, *45*(5), 688–97. doi:10.1111/j.1469-8986.2008.00668.x

Holroyd, C. B., & Yeung, N. (2011). An integrative theory of anterior cingulate cortex function: Option selection in hierarchical reinforcement learning. In R. B. Mars, J. Sallet, M. Rushworth, & N. Yeung (Eds.), *Neural Basis of Motivational and Cognitive Control* (pp. 333–349). Cambridge, MA: MIT Press.

Holroyd, C. B., & Yeung, N. (2012). Motivation of extended behaviors by anterior cingulate cortex. *Trends in Cognitive Sciences*, *16*(2), 122–128. doi:10.1016/j.tics.2011.12.008

Horschig, J. M., Zumer, J. M., & Bahramisharif, A. (2014). Hypothesis-driven methods to augment human cognition by optimizing cortical oscillations. *Frontiers in Systems Neuroscience*, *8*(June), 1–21. doi:10.3389/fnsys.2014.00119

Hoshi, E., Shima, K., & Tanji, J. (2000). Neuronal activity in the primate prefrontal cortex in the process of motor selection based on two behavioral rules. *Journal of Neurophysiology*, *83*(4), 2355–2373.

Howard, M. W., Rizzuto, D. S., Caplan, J. B., Madsen, J. R., Lisman, J., Aschenbrenner-Scheibe, R., … Kahana, M. J. (2003). Gamma oscillations correlate with working memory load in humans. *Cerebral Cortex*, *13*(12), 1369–1374. doi:10.1093/cercor/bhg084

Hsieh, L. T., & Ranganath, C. (2014). Frontal midline theta oscillations during working memory maintenance and episodic encoding and retrieval. *NeuroImage, 85*, 721–729. doi:10.1016/j.neuroimage.2013.08.003

Ishii, R., Shinosaki, K., Ukai, S., Inouye, T., Ishihara, T., Yoshimine, T., … Takeda, M. (1999). Medial prefrontal cortex generates frontal midline theta rhythm. *NeuroReport, 10*(4), 675–679. doi:10.1097/00001756-199903170-00003

Jasper, H. H. (1958). The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, *10*, 371–375. doi:10.1016/0013-4694(58)90053-1

Jenkinson, N., & Brown, P. (2011). New insights into the relationship between dopamine, beta oscillations and motor function. *Trends in Neurosciences*, *34*(12), 611–618. doi:10.1016/j.tins.2011.09.003

Jensen, O., & Colgin, L. L. (2007). Cross-frequency coupling between neuronal oscillations. *Trends in Cognitive Sciences*, *11*(7), 267–269. doi:10.1016/j.tics.2007.05.003

Jensen, O., Kaiser, J., & Lachaux, J. P. (2007). Human gamma-frequency oscillations associated with attention and memory. *Trends in Neurosciences*, *30*(7), 317–324. doi:10.1016/j.tins.2007.05.001

Jensen, O., & Lisman, J. E. (1998). An oscillatory short-term memory buffer model can account for data on the Sternberg task. *The Journal of Neuroscience*, *18*(24), 10688–10699.

Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Frontiers in Human Neuroscience*, *4*, 186. doi:10.3389/fnhum.2010.00186

Jensen, O., & 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. doi:10.1046/j.1460-9568.2002.01975.x

Kamiński, J., Brzezicka, A., & Wróbel, A. (2011). Short-term memory capacity (7±2) predicted by theta to gamma cycle length ratio. *Neurobiology of Learning and Memory*, *95*(1), 19–23. doi:10.1016/j.nlm.2010.10.001

Kawasaki, M., & Yamaguchi, Y. (2013). Frontal theta and beta synchronizations for monetary reward increase visual working memory capacity. *Social Cognitive and Affective Neuroscience*, *8*(5), 523–530. doi:10.1093/scan/nss027

Kilner, J. M., Mattout, J., Henson, R., & Friston, K. J. (2005). Hemodynamic correlates of EEG: A heuristic. *NeuroImage*, *28*(1), 280–286. doi:10.1016/j.neuroimage.2005.06.008

Kim, D. J., Yogendrakumar, V., Chiang, J., Ty, E., Wang, Z. J., & McKeown, M. J. (2013). Noisy galvanic vestibular stimulation modulates the amplitude of EEG synchrony patterns. *PLoS ONE*, *8*(7). doi:10.1371/journal.pone.0069055

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews*, *29*(2-3), 169–195. doi:10.1016/S0165-0173(98)00056-3

Knoch, D., Gianotti, L. R. R., Pascual-Leone, A., Treyer, V., Regard, M., Hohmann, M., & Brugger, P. (2006). Disruption of right prefrontal cortex by low-frequency repetitive transcranial magnetic stimulation induces risk-taking behavior. *The Journal of Neuroscience*, *26*(24), 6469–6472. doi:10.1523/JNEUROSCI.0804-06.2006

Knyazev, G. G., Slobodskoj-Plusnin, J. Y., Bocharov, A. V., & Pylkova, L. V. (2011). The default mode network and EEG alpha oscillations: An independent component analysis. *Brain Research*, *1402*, 67–79. doi:10.1016/j.brainres.2011.05.052

Kopell, N. (2013). Interview with Nancy Kopell. *Trends in Neurosciences*, *36*(6), 313–4. doi:10.1016/j.tins.2013.03.005

Kopell, N., Ermentrout, G. B., Whittington, M. A., & Traub, R. D. (2000). Gamma rhythms and beta rhythms have different synchronization properties. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(4), 1867–1872. doi:10.1073/pnas.97.4.1867

Kringelbach, M. L. (2005). The human orbitofrontal cortex: linking reward to hedonic experience. *Nature Reviews Neuroscience*, *6*, 691–702. doi:10.1038/nrn1748

Lara, A. H., & Wallis, J. D. (2014). Executive control processes underlying multi-item working memory. *Nature Neuroscience*, *17*(6), 876–83. doi:10.1038/nn.3702

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

Lenartowicz, A., & McIntosh, A. R. (2005). The Role of Anterior Cingulate Cortex in Working Memory is Shaped by Functional Connectivity. *Journal of Cognitive Neuroscience*, *17*(7), 1026–1042. doi:10.1162/0898929054475127

Lewis, S. J. G., Slabosz, A., Robbins, T. W., Barker, R. A., & Owen, A. M. (2005). Dopaminergic basis for deficits in working memory but not attentional set-shifting in Parkinson's disease. *Neuropsychologia*, *43*(6), 823–832. doi:10.1016/j.neuropsychologia.2004.10.001

Lisman, J. E., & Idiart, M. A. (1995). Storage of 7 +/- 2 short-term memories in oscillatory subcycles. *Science*, *267*(5203), 1512–1515. doi:10.1126/science.7878473

Lisman, J. E., & Jensen, O. (2013). The Theta-Gamma Neural Code. *Neuron*, *77*(6), 1002–1016. doi:10.1016/j.neuron.2013.03.007

Little, S., & Brown, P. (2014). The functional role of beta oscillations in Parkinson's disease. *Parkinsonism and Related Disorders*, *20*(SUPPL.1). doi:10.1016/S1353-8020(13)70013-0

Lohrenz, T., McCabe, K., Camerer, C. F., & Montague, P. R. (2007). Neural signature of fictive learning signals in a sequential investment task. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(22), 9493–8. doi:10.1073/pnas.0608842104

Luft, C. D. B. (2014). Learning from feedback: the neural mechanisms of feedback processing facilitating better performance. *Behavioural Brain Research*, *261*, 356–68. doi:10.1016/j.bbr.2013.12.043

Luigjes, J., Mantione, M., van den Brink, W., Schuurman, P. R., van den Munckhof, P., & Denys, D. (2011). Deep brain stimulation increases impulsivity in two patients with obsessive–compulsive disorder. *International Clinical Psychopharmacology*, *26*(6), 338–340. doi:10.1097/YIC.0b013e32834af505

Marco-Pallarés, J., Cucurell, D., Cunillera, T., García, R., Andrés-Pueyo, A., Münte, T. F., & Rodríguez-Fornells, A. (2008). Human oscillatory activity associated to reward processing in a gambling task. *Neuropsychologia*, *46*(1), 241–8. doi:10.1016/j.neuropsychologia.2007.07.016

Marco-Pallarés, J., Cucurell, D., Cunillera, T., Krämer, U. M., Càmara, E., Nager, W., … Rodriguez-Fornells, A. (2009). Genetic variability in the dopamine system (dopamine receptor D4, catechol-O-methyltransferase) modulates neurophysiological responses to gains and losses. *Biological Psychiatry*, *66*(2), 154–61. doi:10.1016/j.biopsych.2009.01.006

Marco-Pallarés, J., Münte, T. F., & Rodríguez-Fornells, A. (2015). The role of high-frequency oscillatory activity in reward processing and learning. *Neuroscience & Biobehavioral Reviews*, *49*, 1–7. doi:10.1016/j.neubiorev.2014.11.014

Mas-Herrero, E., & Marco-Pallarés, J. (2014). Frontal theta oscillatory activity is a common mechanism for the computation of unexpected outcomes and learning rate. *Journal of Cognitive Neuroscience*, *26*, 447–58. doi:10.1162/jocn_a_00516

Mas-Herrero, E., Ripollés, P., HajiHosseini, A., Rodríguez-Fornells, A., & Marco-Pallarés, J. (2015). Beta oscillations and reward processing: Coupling oscillatory activity and hemodynamic responses. *NeuroImage*, *119*, 13–19. doi:10.1016/j.neuroimage.2015.05.095

Masson, M. E. J., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology*, *57*(3), 203–220. doi:10.1037/h0087426

Mazzoni, P., Hristova, A., & Krakauer, J. W. (2007). Why don't we move faster? Parkinson's disease, movement vigor, and implicit motivation. *The Journal of Neuroscience*, *27*(27), 7105–7116. doi:10.1523/JNEUROSCI.0264-07.2007

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202. doi:10.1146/annurev.neuro.24.1.167

Miltner, W. H., Braun, C. H., & Coles, M. G. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a "generic" neural system for error detection. *Journal of Cognitive Neuroscience*, *9*(6), 788–98. doi:10.1162/jocn.1997.9.6.788

Mink, J. W. (1996). The basal ganglia: Focused selection and inhibition of competing motor programs. *Progress in Neurobiology*, *50*(4), 381–425. doi:10.1016/S0301-0082(96)00042-1

Mitchell, D. J., McNaughton, N., Flanagan, D., & Kirk, I. J. (2008). Frontal-midline theta from the perspective of hippocampal "theta." *Progress in Neurobiology*, *86*(3), 156–185. doi:10.1016/j.pneurobio.2008.09.005

Moran, R. J., Mallet, N., Litvak, V., Dolan, R. J., Magill, P. J., Friston, K. J., & Brown, P. (2011). Alterations in brain connectivity underlying beta oscillations in parkinsonism. *PLoS Computational Biology*, *7*(8). doi:10.1371/journal.pcbi.1002124

Nieuwenhuis, S., De Geus, E. J., & Aston-Jones, G. (2011). The anatomical and functional relationship between the P3 and autonomic components of the orienting response. *Psychophysiology*, *48*(2), 162–175. doi:10.1111/j.1469-8986.2010.01057.x

Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, *53*(3), 139–154. doi:10.1016/j.jmp.2008.12.005

Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: Opportunity costs and the control of response vigor. *Psychopharmacology*, *191*(3), 507–520. doi:10.1007/s00213-006-0502-4

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, *304*(5669), 452–454. doi:10.1126/science.1094285

O'Keefe, J., & Recce, M. L. (1993). Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus*, *3*(3), 317–330. doi:10.1002/hipo.450030307

Pandya, D. N., & Barnes, C. L. (1987). Architecture and connections of the frontal lobe. In E. Perecman (Ed.), *The frontal lobes revisited* (pp. 41–72). New York, NY: IRBN Press.

Pandya, D. N., & Yeterian, E. H. (1990). Prefrontal cortex in relation to other cortical areas in rhesus monkey: architecture and connections. In H. B. . Uylings, C. G. Van Eden, J. P. . De Bruin, M. A. Corner, & M. G. P. Feenstra (Eds.), *Progress in brain research* (Vol. 85, pp. 63–94). Amsterdam, The Netherlands: Elsevier B.V.

Pascual-marqui, R. D. (1999). Review of methods for solving the EEG inverse problem. *International Journal of Bioelectromagnetism*, *1*(1), 75–86. doi:10.1186/1743-0003-5-25

Pascual-Marqui, R. D. (2002). Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. *Methods and Findings in Experimental and Clinical Pharmacology*, *24 Suppl D*, 5–12.

Petit, L., Courtney, S. M., Ungerleider, L. G., & Haxby, J. V. (1998). Sustained activity in the medial wall during working memory delays. *The Journal of Neuroscience*, *18*(22), 9429–9437.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., … Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology*, *37*(2), 127–152. doi:10.1111/1469-8986.3720127

Pinel, J. P. J., & Edwards, M. (2008). *A colorful introduction to the anatomy of human brain: A brain and psychology coloring book* (Second.). Pearson Education, Inc.

Pogosyan, A., Gaynor, L. D., Eusebio, A., & Brown, P. (2009). Boosting cortical activity at beta-band frequencies slows movement in humans. *Current Biology*, *19*(19), 1637–1641. doi:10.1016/j.cub.2009.07.074

Preston, A. R., & Eichenbaum, H. (2013). Interplay of hippocampus and prefrontal cortex in memory. *Current Biology, 23*(17). doi:10.1016/j.cub.2013.05.041

Proudfit, G. H. (2015). The reward positivity: From basic research on reward to a biomarker for depression. *Psychophysiology, 52*(4), 449–459. doi:10.1111/psyp.12370

Ragland, J. D., Turetsky, B. I., Gur, R. C., Gunning-Dixon, F., Turner, T., Schroeder, L., … Gur, R. E. (2002). Working memory for complex figures: an fMRI comparison of letter and fractal n-back tasks. *Neuropsychology, 16*(3), 370–379. doi:10.1037/0894-4105.16.3.370

Reinhart, R. M. G., & Woodman, G. F. (2014). Causal Control of Medial-Frontal Cortex Governs Electrophysiological and Behavioral Indices of Performance Monitoring and Learning. *The Journal of Neuroscience*, *34*(12), 4214–4227. doi:10.1523/JNEUROSCI.5421-13.2014

Romero Lauro, L. J., Rosanova, M., Mattavelli, G., Convento, S., Pisoni, A., Opitz, A., … Vallar, G. (2014). TDCS increases cortical excitability: Direct evidence from TMS-EEG. *Cortex, 58*, 99–111. doi:10.1016/j.cortex.2014.05.003

Sakai, K. (2008). Task set and prefrontal cortex. *Annual Review of Neuroscience, 31*, 219–45. doi:10.1146/annurev.neuro.31.060407.125642

Sakai, K., & Passingham, R. E. (2003). Prefrontal interactions reflect future task operations. *Nature Neuroscience, 6*, 75–81. doi:10.1038/nn987

Sambrook, T. D., & Goslin, J. (2015). A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychological Bulletin, 141*(1), 213–235. doi:10.1037/bul0000006

Sauseng, P., & Klimesch, W. (2008). What does phase information of oscillatory brain activity tell us about cognitive processes? *Neuroscience and Biobehavioral Reviews, 32*(5), 1001–1013. doi:10.1016/j.neubiorev.2008.03.014

Sauseng, P., Klimesch, W., Heise, K. F., Gruber, W. R., Holz, E., Karim, A. A., … Hummel, F. C. (2009). Brain oscillatory substrates of visual short-term memory capacity. *Current Biology, 19*(21), 1846–1852. doi:10.1016/j.cub.2009.08.062

Schmidt, L., D'Arc, B. F., Lafargue, G., Galanaud, D., Czernecki, V., Grabli, D., … Pessiglione, M. (2008). Disconnecting force from money: Effects of basal ganglia damage on incentive motivation. *Brain, 131*(5), 1303–1310. doi:10.1093/brain/awn045

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron, 36*(2), 241–263. doi:10.1016/S0896-6273(02)00967-4

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*, 1593–1599. doi:10.1126/science.275.5306.1593

Seamans, J. K., & Yang, C. R. (2004). The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Progress in Neurobiology*, *74*(1), 1–57. doi:10.1016/j.pneurobio.2004.05.006

Singer, W., Engel, a K., Kreiter, a K., Munk, M. H., Neuenschwander, S., & Roelfsema, P. R. (1997). Neuronal assemblies: necessity, signature and detectability. *Trends in Cognitive Sciences*, *1*(7), 252–61. doi:10.1016/S1364-6613(97)01079-6

Spencer, K. M., Dien, J., & Donchin, E. (2001). Spatiotemporal analysis of the late ERP responses to deviant stimuli. *Psychophysiology*, *38*(2), 343–358. doi:10.1017/S0048577201000324

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning, An Introduction*. Cambridge, MA: The MIT Press.

Sutton, R. S., Precup, D., & Singh, S. (1999). Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. *Artificial Intelligence*, *112*(1), 181–211. doi:10.1016/S0004-3702(99)00052-1

Tallon-Baudry, C., & Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends in Cognitive Sciences*, *3*(4), 151–162. doi:10.1016/S1364-6613(99)01299-1

Todd, M. T., Niv, Y., & Cohen, J. D. (2008). Learning to use working memory in partially observable environments through dopaminergic reinforcement. In D. Koller, D. Achuurmans, Y. Bengio, & L. Bootou (Eds.), *Advances in Neural Information Processing Systems 21 (NIPS 2008)* (Vol. 21, pp. 1689–1696). NIPS.

Turner, R. S., & Desmurget, M. (2010). Basal ganglia contributions to motor control: A vigorous tutor. *Current Opinion in Neurobiology*, *20*(6), 704–716. doi:10.1016/j.conb.2010.08.022

Uhlhaas, P. J., & Singer, W. (2006). Neural synchrony in brain disorders: Relevance for cognitive dysfunctions and pathophysiology. *Neuron*, *52*(1), 155–168. doi:10.1016/j.neuron.2006.09.020

Utz, K. S., Dimova, V., Oppenländer, K., & Kerkhoff, G. (2010). Electrified minds: Transcranial direct current stimulation (tDCS) and Galvanic Vestibular Stimulation (GVS) as methods of non-invasive brain stimulation in neuropsychology-A review of current data and future implications. *Neuropsychologia*, *48*(10), 2789–2810. doi:10.1016/j.neuropsychologia.2010.06.002

Van de Vijver, I., Ridderinkhof, K. R., & Cohen, M. X. (2011). Frontal Oscillatory Dynamics Predict Feedback Learning and Action Adjustment. *Journal of Cognitive Neuroscience*, *23*(12), 4106–4121. doi:10.1162/jocn_a_00110

Vanni, S., Tanskanen, T., Seppä, M., Uutela, K., & Hari, R. (2001). Coinciding early activation of the human primary visual cortex and anteromedial cuneus. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(5), 2776–2780. doi:10.1073/pnas.041600898

Von Stein, A., & Sarnthein, J. (2000). Different frequencies for different scales of cortical integration: From local gamma to long range alpha/theta synchronization. *International Journal of Psychophysiology*, *38*(3), 301–313. doi:10.1016/S0167-8760(00)00172-0

Wang, X. J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiological Reviews*, *90*, 1195–1268. doi:10.1152/physrev.00035.2008

Wilkinson, D., Nicholls, S., Pattenden, C., Kilduff, P., & Milberg, W. (2008). Galvanic vestibular stimulation speeds visual memory recall. *Experimental Brain Research*, *189*(2), 243–248. doi:10.1007/s00221-008-1463-0

Yamamoto, Y., Struzik, Z. R., Soma, R., Ohashi, K., & Kwak, S. (2005). Noisy vestibular stimulation improves autonomic and motor responsiveness in central neurodegenerative disorders. *Annals of Neurology*, *58*(2), 175–181. doi:10.1002/ana.20574