THE HUMAN LATERAL ORBITOFRONTAL CORTEX AND REPRESENTATIONS OF MOTIVATIONAL CONTEXT FOR ACTION: BASIC FINDINGS AND RELEVANCE FOR PSYCHOPATHOLOGY

by

Stefan Ursu

M.D., University of Medical and Pharmaceutical Sciences, Iasi, Romania, 1995

Submitted to the Graduate Faculty of

Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2004

UNIVERSITY OF PITTSBURGH

FACULTY OF ARTS AND SCIENCES

This dissertation was presented

by

Stefan Ursu

It was defended on

July 28, 2004

and approved by

Hans C.R. Breiter, M.D.

Cameron S. Carter, M.D.
Committee Chairperson

Julie A. Fiez, Ph.D.

David A. Lewis, M.D.

Carl R. Olson, Ph.D.

Peter L. Strick, Ph.D.

THE HUMAN LATERAL ORBITOFRONTAL CORTEX AND REPRESENTATION OF MOTIVATIONAL CONTEXT FOR ACTION: BASIC FINDINGS AND RELEVANCE FOR PSYCHOPATHOLOGY

Stefan Ursu, Ph.D.

University of Pittsburgh, 2004

The human orbitofrontal cortex (OFC) is known to play a critical role in goal-directed behavior. However, it is still unknown whether the OFC contribution to guiding behavior is through top-down control of inappropriate responses or through providing a motivational context by representing potential outcomes. This dilemma stems from the fact that research to date has not been able to clearly distinguish the effects of changes in motivational states from associated adjustments in cognitive control. In order to answer this question, two functional magnetic resonance imaging (fMRI) studies, simultaneously and independently manipulated demands for inhibitory control and monetary incentives for correct performance. Across experiments, demands for control only engaged the OFC when they also increased the likelihood of a negative outcome, in the form of increased error rates. In contrast to these effects of control, expected outcomes modulated the OFC activity irrespective of whether the demands for control were high or low. Moreover, the lateral areas of the OFC were maximally engaged during expectation of negative outcomes. Thus, we provided convergent evidence that the OFC is specifically involved in establishing the motivational context of behavior through representation of possible outcomes. Furthermore, the nature of outcome representations in these two experiments have potential implications for existing theories of decision making, by providing evidence that the OFC

representations of potential outcomes are influenced the whole range of possible alternative outcomes.

In a third fMRI experiment, the framework laid out by the first two studies was applied in the analysis of data investigating the neural substrates of obsessive-compulsive disorder (OCD). Compared to control subjects, the OFC of patients showed increased activity in response to stimuli associated with relatively increased potential for negative outcomes, despite the fact that those stimuli were not directly relevant to obsessive-compulsive symptomatology. These findings suggest that the OFC hyperactivity in OCD may reflect an underlying neural dysfunction, and are consistent with the phenomenology of this disease, in which excessive concerns for potential negative outcomes of actions are a prominent feature of symptomatology.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

Publications

This dissertation is comprised of a general introductory chapter, followed by two chapters (second chapter represents a separate research project, and the third chapter describes two complementary projects), and a final discussion chapter. The three research projects are listed below and have been submitted for publication or are being prepared for submission.

Chapters 2 and 3: Stefan Ursu, Kristi A. Clark, V. Andrew Stenger and Cameron S. Carter. Dissociating motivational context from cognitive control in the human orbitofrontal cortex *(submitted for publication).*

Chapter 4: Stefan Ursu and Cameron S. Carter. Increased orbital-frontal activation to predictors of potential negative outcomes in obsessive-compulsive patients: implications for altered reward-related processes *(in preparation).*

Acknowledgements

I am grateful to many people who supported me in the completion of this body of work. Firstly, I thank my advisor, Cameron Carter, who has introduced me to neuroimaging research and was always available to guide me towards the final goal. His mentorship has also led me to fully appreciate the power of theoretically-driven cognitive neuroscientific methods in understanding the biological substrates of human behavior and its disturbance in psychiatric disorders. I would also like to thank my doctoral committee members, Hans Breiter, Julie Fiez, Davis Lewis, Carl Olson and Peter Strick. They were a constant source of deep insights and valuable input over the years, which had a major impact in shaping my thinking as a neuroscientist.

I would also like to acknowledge the support from members of the Clinical Cognitive Neuroscience Laboratory at the Western Psychiatric Institute and Clinic. First and foremost, Angus MacDonald, III, John Kerns and Ray Cho have been great role models of young researchers starting tremendously successful careers while making time to be great friends and mentors. Greg Siegle has repeatedly and patiently helped me think through statistical issues in neuroimaging, and generously provided software that sped up some of the data analyses included in my dissertation. Kate Fissell's efforts have gone beyond inventive and rigorous development of functional MRI data analysis software; without her tireless technical support, the work included in my dissertation would have taken much longer.

I would also like to thank my friends and colleagues Vivek Khatri and Vincent van Veen. Vivek was always there with his unfaltering friendship, even in the most difficult moments, while Vincent taught me a great deal by repeatedly lending me the sounding board of his

sophisticated cognitive psychologist mind and by introducing me to the complex world of fine Belgian beers.

Finally, it was only with the support of a few very special people that this work was possible. I am deeply grateful to my best friend, collaborator and partner Kristi Clark for her trust in the value of my work and for supporting me, numerous times, with her unique combination of brilliance and genuine enthusiasm for science. Finally, the unwavering love and faith of my parents Dorina and Dodu and my sister Stefania were always with me, regardless whether from a few miles away or, as was the case most of the time, from across the Atlantic and most of Europe.

CHAPTER 1

Introduction and background

1.1.The prefrontal cortex and executive functions

The adaptive control of behavior is the result of multiple cognitive processes, including representation and maintenance of goals, maintenance and manipulation of task rules associated with any particular goal, selection of appropriate responses from an array of possibilities, and execution of the selected response. In general, the term “executive control” can be used to refer to any or all of these types of processes. By the same token, motivational and emotional factors interact with cognition in implementation of executive control functions, presumably by modulating the strength or quality of the representations of goals and their subjective utility as well as by influencing the efficacy of the active processes leading to task execution. All the aspects of behavioral regulation listed above have been at some point linked to the prefrontal cortex, and these questions continue to be the focus of extensive research efforts.

While it is widely accepted that the PFC fulfills the functional anatomical requirements for a critical role in executive control of behavior, the precise mechanisms through which it could implement this control have remained poorly defined. Significant efforts have been made to study possible principles of functional organization within the PFC. Of direct interest to the topic of this dissertation are hypotheses proposing possible gradients of functional specialization in the dorsal-ventral axis, that is differences in specialized functions of the dorsolateral vs. ventral areas of the PFC. A long standing view suggests involvement of the dorsal and lateral

PFC in mnemonic and attentional processes, while the orbital areas of the PFC are critical for inhibitory function (Fuster, 1997). A more recent line of research proposed that inhibitory functions are a general property of the PFC, but that the type of information that is being inhibited differs between the dorsolateral PFC (DLPFC) and the orbitofrontal cortex (OFC) (Roberts, Robbins, & L., 1998; Roberts & Wallis, 2000). These theories imply the existence of active top-down inhibitory processes in the service of selection of the most appropriate response (or representation) according to the task demands. More recent theoretical work using computational models proposed a parsimonious alternative, which integrates many diverse findings in the literature related to executive control while not resorting to active inhibitory processes as support of the appropriate behavioral output. This model of prefrontal function posits that the prefrontal cortex is the specifically involved in control of behavior by actively maintaining patterns of activity that represent task goals and the intermediate steps necessary to achieve them (Miller & Cohen, 2001). The "executive" aspect of the activity in these prefrontal neural networks is thought to be exerted through a biasing effect on the activity of other subsystems involved in task execution at various levels, e.g. sensory processing, mnemonic operations (recognition, retrieval, etc), response selection. Thus, the PFC implements executive control of cognitive processes by biasing the "flow of neural activity" through pathways which ensure appropriate mapping of inputs onto outputs, with the ultimate goal of reaching optimal internal states. In other words, the PFC is thought to ensure that patterns of activation in networks directly involved in task execution are appropriate relative to the specific context provided by current goals and task rules. Expanding on models of visual attention based on biased competition (Desimone & Duncan, 1995), this framework generated computational models (O'Reilly, Noelle, Braver, & Cohen, 2002) which accounted for observed patterns of

deficits induced by lesions of the lateral vs. orbitofrontal cortex (Dias, Robbins, & Roberts, 1996a, 1996b, 1997). This model accounted for performance and deficits in a dynamic categorization task based only on active maintenance of featural and dimensional information regarding stimuli involved, combined with gating effects of rewarding feedback. However, this type of biased competition model did not require any active top-down inhibitory mechanism for the control of inappropriate responses or representations, as other models did (Dias et al., 1996a; Fuster, 1997; Roberts et al., 1998; Roberts & Wallis, 2000).

The work presented in this thesis is part of an effort to further our understanding of the role of the PFC in behavioral control and its functional organization. The overall model of PFC function adopted here is on the lines of that proposed by Miller and Cohen (2001). However, in terms of its functional organization, an alternative heuristic is proposed here, namely that of distinguishing the task context into two types: motivational context and task-set context. Motivational context can essentially be equated with the representation of incentives that motivate any given behavior, which in most cases coincides (though not always, as will be detailed later) with the overall utility[1] of the final outcome. While this term has been previously used in studies of PFC function, here it is extended to include not only the motivational context provided by positive reinforcers or rewards (Watanabe, Hikosaka, Sakagami, & Shirakawa, 2002), but also that of punishments and negative reinforcers. On the other hand, task-set context includes, under this heuristic, the task-relevant characteristic of the stimuli involved in a given task and the rules which associate them with particular responses.

[1] The use of the term utility throughout this thesis is basically interchangeable with "value", and is meant to capture the impact that a certain outcome will have on behavior, similar to their use in decision-making literature. The rather subtle differences that exist between these terms in economic studies are not relevant for the purpose of the work that will be presented here. In fact, the precise definition of such terms is still being updated, as are theories of decision-making based on rewards and punishments (Mellers, 2000). However, one important aspect to be noted at this time is that existing decision-making theories refer to utility and/or value as a function that determines choice of behavior, and that it is generally thought to be referenced to the status-quo. That is, the utility of a choice is determined, basically, by how advantageous the outcome would be *relative to the present situation*.

1.2.The orbitofrontal cortex (OFC) and executive control

Based on the inputs from various subdivisions of the mediodorsal thalamic nucleus to the prefrontal cortex, three main subdivisions of the primate PFC can be identified (Fuster, 1997) the dorsolateral PFC (DLPFC), the frontal eye fields and the orbitofrontal cortex (OFC). The aim of the work included in this thesis is to investigate the role of the latter in the control of behavior. Despite data suggesting considerable functional heterogeneity of the ventral prefrontal cortex, which therefore limits the usefulness of defining strict anatomical boundaries, the primate OFC generally includes the most ventral aspects of Walker areas 10 and 11 anteriorly, orbital area 12 (or 47 in humans) laterally, area 13 caudally and area 14 medially.

1.2.1. Anatomy of the OFC

While specific cytoarchitecture and connectivity can be described in connection to each of its subdivisions, several principles seem to characterize the anatomy of the OFC. One is the presence of two gradients of differentiation of the cortical layers, and the other is the bias in connectivity towards various other brain areas. In terms of level of differentiation, the most striking characteristic of the OFC subregions is that the presence of the granular layer IV varies from practically none (in the caudal periallocortex and proisocortex situated most caudally and medially) to a an incipient layer IV in central areas (areas 13) to a clearly present granular layer in more rostral areas 11 and 10 (Barbas & Pandya, 1989; Carmichael & Price, 1994). The same is true in the medio-lateral axis, where medial area 14 has only an incipient granular layer, which becomes progressively more populated in more lateral areas 11 and 12.

One interesting feature of the OFC cytoarchitecture which will prove particularly relevant for the arguments made in this work pertains to the characteristics of cortical area 12, which in humans corresponds roughly to BA 47 (Ongur & Price, 2000). This area includes a part of the prefrontal cortex which extends from the lateral aspects of the orbital face of the frontal lobe to its inferior lateral part. However, the cytoarchitecture of its orbital division is less differentiated compared to the lateral portion, the latter being more similar to the adjacent lateral prefrontal area 46.

The connectivity of the OFC also has quite unique characteristics within the prefrontal cortex. While the connectivity of the OFC has been explored in several species, we will focus in principal on the data generated by studies in non-human primates. As mentioned earlier, the OFC receives projections from multiple thalamic nuclei, with the most significant source of input from the magnocellular division of the medio-dorsal nucleus (Ray & Price, 1993). The OFC also receives significant amygdalar input originating mainly from the baso-lateral nucleus and ending directly or via the magnocellular division of the medio-dorsal thalamic nucleus (MDmc) in the posterior agranular parts of areas 13, 12 and gyrus rectus. The more anterior dysgranular portions (13 proper and caudal 11) receive amygdala input mostly via the MDmc. Importantly, the OFC is the main source of prefrontal output to the amygdala (i.e. the basolateral and accessory nuclei). Other important limbic and paralimbic OFC connections include the lateral hypothalamus and the entorhinal and cingulate cortices (Zald & Kim, 1996). While the precise functional significance of each of these types of connections remains to be established, several details are worth noting. Firstly, the OFC projection to the lateral hypothalamus, a structure critically involved in regulation of visceral responses to stimuli, is mostly unidirectional from the PFC to

the lateral hypothalamus. Secondly, the OFC is the main prefrontal source of input to the entorhinal cortex, which plays a central role in various mnemonic processes.

The OFC also receives multiple inputs from all sensory modalities (Barbas, 1992; Jones & Powell, 1970; Macko et al., 1982; Morecraft, Geula, & Mesulam, 1992; Rolls, Yaxley, & Sienkiewicz, 1990; Tanabe, Iino, & Takagi, 1975), with overall consistency in terms of level of differentiation, i.e. the more differentiated sensory association cortices project to more differentiated areas of the OFC, whereas inputs from more primitive sensory cortices connect to more posterior agranular or dysgranular parts of the OFC (Morecraft et al., 1992; Zald & Kim, 1996).

Finally, the projections from the OFC to the basal ganglia in primates show evidence of topographical distribution in the medio-lateral axis, with medial orbitofrontal areas projecting relatively stronger to the ventral striatum, while more lateral areas connect to more dorsal aspects of the striatum (Ferry, Ongur, An, & Price, 2000).

Taken together, these aspects of cytoarchitecture and connectivity place the OFC in a critical position for the integration of relevant sensory features of stimuli with information regarding their emotional and motivational salience.

1.2.2. Hypotheses regarding the OFC's role in regulation of behavior

The hypothesis that the OFC is directly relevant to behavioral control has a long history in cognitive neuroscience, starting with the case of Phineas Gage, a railroad worker who lived in the 1800's and suffered extensive damage to the ventral prefrontal cortex, leading to drastic changes in his behavior. More controlled studies of humans and animals with brain lesions have since confirmed that damage to the OFC or its animal homologues produce significant

behavioral impairments. Combined with data from electrophysiological studies in non-human primates and, more recently, with data from human neuroimaging studies, they lead to several models regarding OFC role in control of behavior, which I will detail below.

Inhibition of inappropriate responses.

The idea that the OFC is critical for the suppression of "inappropriate behaviors" has played an important role in a frequently used model of prefrontal function (Fuster, 1997), and was fueled mainly by the pattern of behavioral changes undergone by humans suffering from ventral prefrontal lesions. Analyses of behavioral performance on neuropsychological tests of human subjects with cortical lesions including various extents of the OFC have produced data supporting this view (Milner, 1964; Mishkin, 1964), even though in some studies the pattern of deficit could have been interpreted according to alternative views (Berlin, Rolls, & Kischka, 2004). In most cases, their deficits manifest as perseverative responding in conditions when previously rewarded stimuli cease to be associated with rewards, or simply as an inability to withhold prepotent but inappropriate responses. More recently, some imaging studies generated results consistent with this theory. For instance, in a Go/No-go task, response inhibition was associated with activation in the orbital areas of the PFC (Casey et al., 1997; Garavan, Ross, & Stein, 1999). A specific form of inhibitory processes, namely those involved in behavioral changes induced by reversals of reward contingencies, has received special attention, and will be discussed separately in the following section.

Reversal learning

Reversal learning refers, in its more recent conceptualization (Clark, Cools, & Robbins, 2004), to the capacity of inhibiting associations between rewards and stimuli or responses when reward contingencies change. Due to the causal role of the change in reward contingencies for

this type of process, it has also been referred to, in recent studies, as “affective shift” (Dias et al., 1996a, 1997). Evidence for involvement of the OFC in behavioral changes in response to reversals of stimulus-reward contingencies abounds, from studies of both animal and human subjects. As such, it has been repeatedly shown that lesions of the monkey OFC result in perseverative responding to previously rewarded stimuli (Iversen & Mishkin, 1970; Jones & Powell, 1970; Meunier, Bachevalier, & Mishkin, 1997) Similarly, humans with lesions of the ventral PFC have also been shown to exhibit similar deficits (Berlin et al., 2004; Daum, Channon, Polkey, & Gray, 1991; Daum, Schugens, Channon, Polkey, & Gray, 1991; Fellows & Farah, 2003; Hornak et al., 2004; Rolls, Hornak, Wade, & McGrath, 1994). The validity of this type of results has been supported by electrophysiological recordings in monkeys, which have identified neurons in the OFC which reverse their responding to reward visual cues when those cues switched from predicting availability to predicting absence of a reward (Rolls, Critchley, Mason, & Wakeman, 1996). In recent years, data from marmosets with lesions of the OFC or DLPFC have been used to propose a specific function of the OFC in reversal learning. Specifically, Dias and colleagues (1996a; Dias et al., 1996b, 1997; Roberts & Wallis, 2000) have used a task in which, after acquisition of a discrimination between stimuli with features along two dimensions (lines and shapes), the reward contingencies changed. The monkeys were required to detect when:

1) a novel feature of the same dimension was rewarded (i.e. if the last feature was a type of line, it stopped being rewarded and a new line was associated with reward; this was labeled “intra-dimensional shift”, IDS);

2) the reward switched to a stimulus feature of the same dimension but which had not been associated with reward until then (labeled “intra-dimensional reversal”, IDR);

3) the reward switched to a novel feature from the other dimension (i.e. in the same example, the reward shifted from a type of line to a novel stimulus shape; this was termed “extra-dimensional shift”, EDS).

Using this manipulation, Dias and colleagues (1997) observed a double dissociation in monkeys with lesions of the OFC or lateral PFC: the lateral PFC lesions only affected the first extra-dimensional shift, while the OFC lesions only impaired the first intra-dimensional reversal. Based on this type of result, Roberts & Wallis (2000) and, more recently, Clark et al. (2004), proposed that the OFC has a critical role in implementing a special inhibitory function, namely that of inhibiting associations between stimuli and affective value when these associations are no longer relevant. Elliott et al. (2000) have also argued for a similar mechanism, in the context of a proposed functional segregation. These authors have argued that in general OFC participates in control under conditions of uncertainty, whenever the likely reward value of stimuli and responses can help guide behavior. They also proposed that the medial OFC participates in ascribing a reward value to stimuli, whereas the lateral OFC is active whenever suppression of previously rewarded responses is necessary. While this hypothesis has face validity, it still is unclear what its overall value would be, for several reasons. Firstly, the fact that the effects of the lesions are inferred from performance measures (i.e. number of errors), cannot rule out purely motivational effects (i.e. impairments in the motivational impact of not receiving a reward in response to a stimulus which predicted it). This argument also applies to recent neuroimaging results which proposed similar active mechanisms for the OFC function (Arana et al., 2003; O'Doherty, Critchley, Deichmann, & Dolan, 2003). Secondly, it is unclear that an active control process (i.e. inhibition) is necessary in the first place. More parsimonious computational models based on biased competition between representations of featural (in the OFC) vs. dimensional (in

the lateral PFC) information were able to account for this pattern of deficit without the requirement of a top-down inhibitory process (O'Reilly et al., 2002).

Motivational processes (representation of rewards and punishments)

Another line of research, represented mainly by the work of Rolls et al., emphasizes the OFC involvement in behavioral changes induced by the expectation or experience of various rewards and punishments (Kringelbach & Rolls, 2004; Rolls, 2000). This over-arching theory, supported mainly by evidence of modulation of OFC neuronal activity in response to stimuli predicting various types of rewards or in response to the delivery of rewards, has taken various forms over the years. Until very recently, Rolls (1998) has argued that the OFC is critical in representing and rapidly reversing stimulus-reward associations, in a manner much similar to that proposed by Robbins and Roberts (see previous section). Such reward-induced changes in OFC activity have been documented in electrophysiological experiments on awake behaving monkeys (Hikosaka & Watanabe, 2000; Roesch & Olson, 2004; Rolls et al., 1996; Rolls et al., 1990; Schultz, Tremblay, & Hollerman, 2000; Tremblay & Schultz, 1999, 2000a), and in human neuroimaging studies (Breiter, Aharon, Kahneman, Dale, & Shizgal, 2001; Elliott, Newman, Longe, & Deakin, 2003; O'Doherty et al., 2003; O'Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001; Thut et al., 1997). The proposed involvement of the OFC in reward-based behavioral control was also supported by patterns of deficits in humans with ventral prefrontal lesions, who show perseveration on high-reward, high-risk choices despite overall losses incurred as a result of those choices (Bechara, Damasio, Damasio, & Anderson, 1994; Bechara, Damasio, Tranel, & Damasio, 1997; Bechara, Tranel, & Damasio, 2000; Hornak et al., 2004; Rolls et al., 1994). Several results from studies investigating the role of the OFC in sensory processes are also consistent with a role in representations of rewards and punishments: in

humans, anticipation and experience of pleasant and unpleasant stimuli like tastes or foods, odors, touch and faces, tend to modulate activity in the OFC (De Araujo & Rolls, 2004; O'Doherty et al., 2000; O'Doherty, Deichmann, Critchley, & Dolan, 2002; Rolls et al., 2003; Small, Zatorre, Dagher, Evans, & Jones Gotman, 2001).

1.2.3. Distinguishing outcome representation from active cognitive control

A careful examination of the hypothesis outlined above regarding the role of the OFC in behavioral control reveals several important aspects. Firstly, and the most relevant for the topic of this thesis, is the fact that none of the theories attempts to clearly distinguish the contribution of active top-down control from that of motivational processes linked to representation of rewards and punishments (i.e. representations of outcomes of actions). Any form of inhibitory function, as well as "reversal" of representations or associations, are by definition active control processes which actively participate in modification of behavior, and are therefore part of the "task-set" or "task-specific context" established by the representations of stimuli, rules and stimulus-response mappings which ensure correct performance of any given task (Miller & Cohen, 2001). In contrast, representation of outcomes and their utility, while ultimately having an effect on behavior, do not (at least in theory) carry this requirement by definition. In other words, active inhibitory control or reversals participate in establishing "how" to perform, whereas motivational context informs "why", without necessarily specifying "how". Obviously, the two types of representations may ultimately be subserved by the same area of the brain or by separate areas, but that is precisely the empirical question which has not been directly addressed in humans. Also, while this may be one of the most fundamental heuristics that can be used when studying the impact of motivational factors on behavior, it is also very difficult to test. As has

been pointed out recently with regards to modulation of attentional processes by rewards (Maunsell, 2004), the use of positive rewards to change the subjects' motivational state confounds reward representation with active control processes which necessarily are engaged when subjects modify their behavior in order to obtain the reward. Perhaps because of this tight relationship between motivation and control mechanisms, very few data exist that speak directly to this issue, especially in humans[2].

Much of the work included in this thesis (i.e. Chapters 2 and 3) was aimed at distinguishing two factors in the regulation of human behavior. Chapter 2 of this thesis will present data from an fMRI experiment in human subjects which tested the hypothesis that the human OFC is involved in representing the potential outcome of future behavior, and not in active implementation of control processes which lead to changes in behavior. Understanding the functions of the human OFC in behavioral control is also critical for the study of major psychiatric illnesses such as mood and anxiety disorders. In the next section, I will review previous studies which identified OFC hyperactivity as a characteristic of several anxiety disorders, such as obsessive-compulsive disorder (OCD), simple phobia and post-traumatic stress disorder (Rauch, Savage, Alpert, Fischman, & Jenike, 1997). In this thesis, Chapter 4 describes an analysis of fMRI data from OCD patients and healthy subjects which tested a prediction regarding the relevance of OFC hyperactivity to OCD as well as, in this context, a prediction generated by the hypothesis that the human OFC is involved in representation of potential negative outcomes.

[2] In the time since the work presented in this thesis has been started, only one study in awake behaving monkeys generated results consistent with the hypothesis that the primate OFC is involved in representing the reward value of reinforcers and not in control mechanisms (Roesch & Olson, 2004).

1.3.Orbitofrontal cortex and psychopathology

Aside from its fundamental scientific value, characterizing the involvement of the PFC in executive functions has direct implications for human psychopathology. Extensive evidence points toward a critical role for various PFC areas in the pathogenesis of major psychiatric disorders such as schizophrenia, major depressive disorder and bipolar disorder and anxiety disorders such as obsessive-compulsive disorder (Phillips, Drevets, Rauch, & Lane, 2003; Saxena, Brody, Schwartz, & Baxter, 1998).

1.3.1. Orbitofrontal cortex dysfunction in OCD

Obsessive-compulsive disorder (OCD) is a common psychiatric disorder, with a lifetime prevalence of 1-2.5% (Robins et al., 1984). It is a heterogeneous condition, characterized by the presence of either obsessions (recurrent and persistent thoughts, impulses or sensations) or compulsions (intentional repetitive behaviors like hand washing and checking, or mental acts such as counting, that the person feels driven to perform in response to obsessions or according to rigid rules).

Despite the complexity of the clinical picture, it is currently well established that OCD has a biological basis. Reports of comorbidity with neurological disorders like postencephalitic parkinsonism, Sydenham's chorea, Tourette's syndrome, as well as effects of lesions affecting the basal ganglia and/or frontal structures (Berthier, Kulisevsky, Gironell, & Heras, 1996) suggested that components of the basal ganglia (such as caudate nucleus, globus pallidus, putamen), of the prefrontal cortex (such as the orbitofrontal, anterior cingulate or dorsolateral prefrontal cortices) and the thalamus, are critical for the manifestations of OCD. With the development of neuroimaging methods, a growing body of data strengthened the link between

the presence of OCD symptoms and these structures (Saxena et al., 1998). Studies using Positron Emission Tomography (PET) (Baxter et al., 1988; Sawle, Hymas, Lees, & Frackowiak, 1991; Swedo et al., 1989) or Single Photon Emission Computed Tomography (SPECT) (Adams, Warneke, McEwan, & Fraser, 1993; Machlin et al., 1991) have linked the presence of OCD symptoms with increases in resting metabolism of the OFC, CdN, ACC and thalamus. Other studies used the same imaging techniques (Benkelfat et al., 1990; Rauch et al., 1994) or functional MRI (Breiter & Rauch, 1996) to compare indices of neural activity at rest with that during provocation of OC symptoms; this approach pointed towards abnormalities in the same brain structures. Finally, several studies explored the OCD pathophysiology by assessing resting rCBF or rCMglu (regional cerebral glucose metabolism) in OCD patients before and after treatment. Again, the most reliable findings were correlations of improvement in symptoms with decreases of activity in the caudate nucleus, OFC or anterior cingulate (Baxter et al., 1992; Benkelfat et al., 1990; Swedo et al., 1992; Swedo et al., 1989). The importance of the interactions within the circuit including the OFC, the ACC and the caudate nucleus for OCD pathophysiology is consistent with neuroanatomical findings in non-human primates, which point out that areas of the OFC and ACC are rather unique sources of prefrontal input to the caudate nucleus (i.e. input specific to the striosome compartment (Eblen & Graybiel, 1995)). Despite the overall consistency of these results, few models of OCD pathogenesis exist to date (Pitman, 1987; Rauch, 2000), with even fewer tests of predictions generated by such models (Gehring, Himle, & Nisenson, 2000; Ownby, 1998; Ursu, Stenger, Shear, Jones, & Carter, 2003). Therefore, the precise interactions between these structures and precisely how each of them is linked to the manifestation of OCD symptoms remain unclear. There are several possible accounts for the observed OFC hyperactivity, none of which can be ruled out based on existing

data. For instance, it is possible that this hyperactivity represents an epiphenomenon brought about by the symptomatic state. Another possibility is that it reflects a permanent underlying brain dysfunction which results in increased vulnerability for pathological anxiety; furthermore, this type of dysfunction may or may not be specific to OCD as opposed to other anxiety disorders. Chapter 4 describes results consistent with the latter, as well as with the general framework that the human OFC participates in executive functions through representation of potential outcomes.

1.4.Goals

Chapters 2 and 3 describe basic fMRI experiments aimed at distinguishing whether the role of the human OFC in behavioral control is through top-down control of inappropriate behaviors or through representation of possible outcomes which provide the motivational context for behavior. In the experiment described in Chapter 2, subjects were scanned while performing a cognitive task with varied demands for cognitive control. In parallel, the incentives for correct performance were varied on a trial by trial basis, independently from demands for control. Thus it was possible to directly test the hypothesis that the OFC is involved in representing the potential outcome of actions (i.e. the incentive manipulation), and not in implementing control. Chapter 3 describes a complementary experiment (using a modified version of the task described in Chapter 2) to test the hypothesis that the OFC is engaged by cognitive control only when changes in control coincide with changes in the subjects' experience with rewards and punishments. The incentives manipulation used in this experiment also enabled a more detailed characterization of the nature of outcome representation in the human OFC, namely its specificity for relatively negative outcomes. Chapter 4 presents a test of predictions generated by this hypothesis regarding human OFC function, with relevance to psychopathology. Functional

MRI data, obtained from OCD patients while performing a cognitive task, tested the hypothesis that OFC hyperactivity can be elicited with disease-irrelevant stimuli in conditions when these stimuli predict a high likelihood of a negative outcome. This will help establish the relevance of OFC dysfunction for the pathophysiology of OCD.

1.5.Note on isolating anticipatory from feedback activity

The results presented in this thesis were all based on a two-phase design, in which each trial consisted of the sequential presentation of a sample and a target stimulus. Thus, they included a preparatory phase of task execution, followed by the execution of a response which was, in two of the experiments, followed immediately by performance feedback which included information about monetary gains or losses used as incentives. In the analyses of both task and incentive effects in these two experiments, we focused on the anticipatory sample-target interval, when the effects of demands for control and incentives (the two factors manipulated) could be clearly distinguished. Additional confirmatory analyses of effects of demands for control on response-related activity were also performed.

With regards to feedback activity, however, it was not possible to fully control the effects of expectation on feedback-related activity in this design. A significant body of literature regarding decision making in humans, reviewed by Mellers (2000), suggests that the utility of a given outcome is modified by the feelings of disappointment or elation experienced given the best and worst possible outcomes. In this context, the effects of varying levels of demands for control on the differences between positive vs. negative feedback represents a complex issue which could not be addressed beyond the speculative level by these experiments. Considering this, the feedback related activity was not a focus of the work presented in this thesis.

CHAPTER 2

Dissociating top-down control from outcome representation in the OFC

2.1.Introduction and rationale

Mechanisms of active implementation of executive functions are tightly related to motivational factors which modulate them. When organisms are motivated by available rewards or the threat of punishments (relative to a baseline state), control mechanisms are also more active, in support of approach or avoidance behavior, respectively. This argument applies to consummatory behavior, when the individual changes its behavior subsequent to experiencing a reward or a punishment, as well as to incentive motivation, that is when behavior changes in response to stimuli which predict the availability of rewards or punishments.

For instance, when studying control processes through lesion experiments, failure of control mechanisms is inferred from changes in frequency of a certain overt behavior (Aron, Robbins, & Poldrack, 2004; Dias et al., 1997; Roberts & Wallis, 2000), leaving open the possibility that the behavioral changes are the effect of abnormal processing of the rewards (or punishments) which are used to motivate behavior. In particular, perseverative commission errors have long been considered as a hallmark of failure of control processes, namely of the capacity to inhibit a response elicited by presentation of a stimulus which had been previously associated with reward (Dias et al., 1996b; Fuster, 1997; Milner, 1964; Mishkin, 1964; Roberts & Wallis, 2000). However, responding to a stimulus even after it is no longer rewarded could also be caused by decreased impact of the relative punishment of not receiving a reward when it

had been expected (i.e. a decreased impact of “negative punishments”), or by an exaggerated sensitivity to the previously established association between those stimuli and rewards. In imaging studies, neural substrates of inhibitory function are explored by comparing an experimental condition which requires inhibition of a prepotent response with one in which the correct response is more automatic (Casey et al., 1997; Elliott & Dolan, 1999). However, in those instances the increased demands for inhibitory control coincide with changes in performance accuracy (i.e. increased errors) which, in turn, change the experience and expectation of rewards and/or punishments. Similarly, in experiments based on response reversals (Clarke, Dalley, Crofts, Robbins, & Roberts, 2004; Cools, Clark, Owen, & Robbins, 2002; O'Doherty et al., 2003; O'Doherty et al., 2001; Rogers et al., 1999), the correlates of control mechanisms cannot be completely separated from the impact of negative feedback. This is because in such tasks the stimulus-reward contingencies have to be inferred from the nature of the feedback, and therefore the subjects’ change in behavior coincides, presumably, with significant impact of the negative feedback in the previous trial. For example, in a recent study, O’Doherty et al. (2003) used a probabilistic reward reversal task to explore whether the OFC was involved in behavioral control or in representing valence of outcome. Subjects had to detect which stimulus, of two choices, was associated in 80% of trials with winnings and in 20% with losses (the other stimulus had reversed probabilities of winning and losing). After acquiring this contingency, the proportion of wins/losses reversed without warning, such that the old “good” choice now started yielding losses in 80% of the trials. Using this design, O’Doherty and colleagues found more activation in the lateral OFC in response to stimuli which, after the reversal in contingency, led to losses and reversal of response in the subsequent trial, relative to stimuli which led to losses but which did not lead to a reversal in response. This difference in

activity was interpreted as evidence that the OFC was involved in switching from one pattern of behavior to the alternative one, and not in representation of the outcome valence. However, because of the probabilistic nature of the rewards (i.e. 20% of trials resulted in losses even when the "good" stimulus was chosen), after the switch in contingency the first few losses to the "good" stimulus (which in the meantime has become "bad' without the subjects' knowledge) can be regarded as part of the small percentage of expected loss trials. Therefore, it is possible that trials which are followed by the actual switch in behavior are also the trials in which subjects' stimulus-reward association switches from "good" to "bad". Consequently, these trials still confound control processes (resulting in response switch) and outcome representation ("good" vs. "bad" stimulus).

Existing studies exploring the neural substrates of reward representation are largely affected by the same problem. In general, these studies infer effects of rewards or punishments from subsequent changes in behavior accuracy or type of response or from changes in reaction time (Hikosaka & Watanabe, 2000; Ramnani & Miall, 2003; Taylor et al., 2004; Tremblay & Schultz, 1999; 2000b,O'Doherty, 2001 #1062). Furthermore, studies which implicated OFC in reward evaluation based on data from sensory discrimination type paradigms (De Araujo & Rolls, 2004; O'Doherty et al., 2000; Rolls et al., 1996; Rolls et al., 2003) cannot explicitly control for the role of cognitive processes associated with evaluation of the potential for reward/punishment of each type of stimulus. Overall, a careful analysis of this literature reveals that, with one notable exception (Roesch & Olson, 2004), increases in neural activity associated with rewards or punishments can still be attributed to active processes which result in behavioral changes.

In order to clearly establish whether a brain structure has a role in representation of reward or in active control of behavior, one needs to create experimental conditions in which the two factors change independently of each other and make different predictions with regards to indices of neural activity. In the present study, 19 human subjects were scanned while performing a cognitive task with two levels of demands for cognitive control. In parallel with demands for control and independent of them, monetary incentives available for correct performance were either present or absent. Therefore, we were able to test whether the OFC activity would parallel the load and/or efficacy of control processes (suggestive of a role in active cognitive control), versus tracking the presence or absence of monetary incentives (consistent with a role in establishing motivational context through outcome representation).

Another issue that was addressed with the present design was the fact that previous studies which interpreted OFC activations as possible reward- or punishment-related representations frequently confounded evaluation/anticipatory processes with execution and feedback-related ones. This confound was created either by the actual task design, which required evaluation of stimuli followed immediately by responding (Elliott & Dolan, 1999; Elliott et al., 2003; O'Doherty et al., 2003; O'Doherty et al., 2001), or by the combination of design and analysis (Ramnani & Miall, 2003). In one of the few studies which detected OFC activations while clearly separating anticipatory and feedback phases of reward processing (Breiter et al., 2001), a region of interest analysis showed that some OFC foci responded most strongly to anticipation of monetary rewards, some to anticipation of losses, while others reacted to the subsequent experience of reward or loss. However, the rewards and losses were delivered in the context of a “wheel of fortune” game, in which outcomes were not contingent on performance, and therefore it was not possible to exert much experimental control on the

associated cognitive processes (such as conscious assessment of outcome probabilities and their expected impact on the overall gains). The advantage of the design used in the current experiment (similar to Experiment 1) was that concomitant performance of a cognitive task and the fact that outcomes were determined by performance (and not probabilistically) limited the impact of such processes. Furthermore, conditioning outcome by performance also provided behavioral measures of both control and motivational processes (i.e. reaction times and accuracy) that were not dependent on self report, which is usually susceptible to influences of cognitive and personality biases.

2.2.Methods

2.2.1. Subjects

Twenty healthy adult subjects participated in the experiment. One subject, whose data showed high scan-to-scan movement (mean scan-to-scan movement higher than ½ voxel) was not included in further analyses. The statistical analyses were therefore performed on data from 19 subjects (9 female) with a mean age of 23.5 years (range 18-35), who had been all screened for history of major neurological or psychiatric disorders. All subjects gave informed consent to participate in an experimental protocol approved by the Institutional Board Review at the University of Pittsburgh.

2.2.2. Behavioral task

Stimulus presentation was done using specialized software (Eprime, PST Inc, Pittsburgh, Pennsylvania) which ran on a IBM-compatible computer under the Windows 98 OS. Stimuli were projected on a translucent screen placed in the scanner approximately 50cm from the

center of the bore, and viewed via an MR-safe prismatic mirror installed on the scanner head coil, directly above the subjects' eyes. Responses during the scanning session were collected using specialized fiberoptic response devices attached to each of the subjects' hands.

Manipulation of demands for cognitive control

The task that subjects had performed while in the MRI scanner varied the demands for control (i.e. the spatial bias to respond on the same side as an identified target) independently from factors influencing the motivational context (i.e. performance feedback and incentives for correct performance). The basic paradigm required subjects to remember a sample stimulus over an interval, identify it in a target stimulus and respond to its spatial location according to one of two rules (i.e. tasks). Depending on which of the two rules was applied, one of the two tasks carried increased demands for cognitive control in order to overcome a prepotent response tendency. Each trial (see Figure 1) consisted of a sample stimulus (a 1 second presentation of a single star-shaped stimulus), a delay and a target (a 1 second side-by-side presentation of the sample plus a different stimulus chosen randomly from three other similar stimuli); immediately following the target, performance feedback was briefly presented for 0.5 seconds, followed by the inter-trial interval.

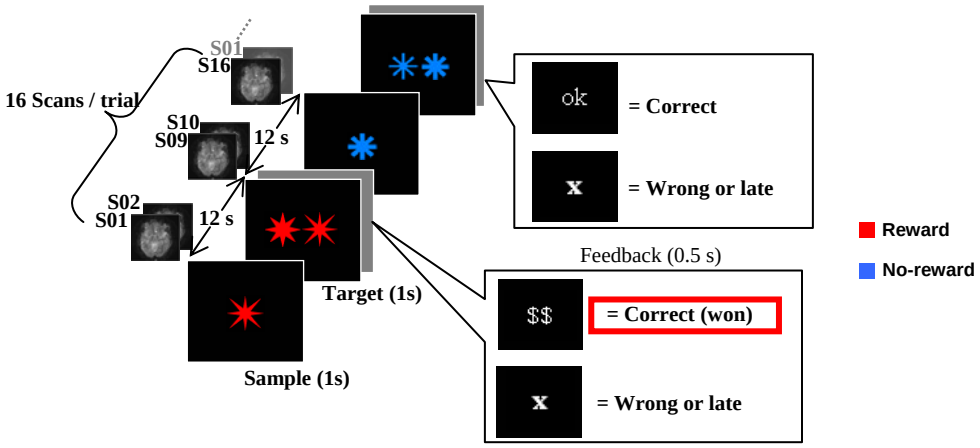

Figure 1. Task design in Experiment 1.

Subjects were required to remember the sample over the first 12 seconds of a trial. Upon presentation of the target, they were required to identify the sample stimulus in the target and respond according to one of two rules: with the index finger of the hand on the same side as the stimulus identical to the sample (Match task blocks) or the index finger of the hand on the side of the stimulus which did not match the sample (Non-match task blocks). Subjects were informed that responses were only recorded while the target was on the screen (1 second) and that late responses would count as errors. The color of the stimuli informed whether or not a monetary reward could be obtained upon correct and fast responding. A 0.5 second feedback screen followed immediately after target offset.

At the beginning of each block of trials, subjects were given the response rule to be applied for each target: to respond on the same side with the stimulus which matched the sample in shape (Match task), or to respond on the side of the unfamiliar stimulus (Non-match task). Therefore, since Non-match required identification of the sample stimulus in the target (as in the Match trials) but suppression of the automatic tendency to respond on the same side, the demands for control were increased in the Non-match task relative to Match (Ridderinkhof, 2002).

Manipulation of incentives

In parallel, and irrelevant to the Match/Non-match rules, the color of the stimuli changed randomly from trial to trial, and informed the subjects about the monetary incentive for a correct

response in the current trial: one color signaled that a $0.50 reward would be won upon correct performance, while the other color signaled that no reward would be obtained. (see Figure 1). The mapping of the colors onto incentive levels was counterbalanced across subjects. Subjects were also informed that responses were only counted while the target was on the screen, and that late responses would count as errors, resulting in not obtaining the reward in the trials in which the monetary reward was available. At the end of the 1 second of target presentation, the target was replaced for 0.5 seconds with performance feedback. The word "ok" meant that the response had been correct and fast (in a No-reward trial), "$$" meant that the response was correct and fast and $0.5 was won (in a Reward trial), and "x" meant that the response was incorrect or too late (resulting in a missing the monetary reward in Reward trials). Subjects were instructed to be as accurate and as fast as possible at all times in order to maximize their winnings, but they were not given a running count of their winnings as the task progressed.

2.2.3. fMRI data collection and statistical analysis

Analyses of behavioral effects were conducted using the random effects ANOVA model implemented in the Unixstat package installed on a SGI platform running IRIX 6.5. The reaction time data were analyzed after excluding trials in which subjects made wrong responses and trials with reaction times shorter than 200ms, to avoid confounds from erroneous responses or guesses. Since subjects made extremely few guessing responses (RT under 200ms), the majority of exclusions were due to errors and averaged approximately 10% of trials across subjects. For analyses of response accuracy, trials with RTs under 200ms were also excluded.

Scanning was performed using a 3T GE Signa scanner with a standard head coil. Functional scans were acquired in the same location as an immediately preceding set of

structural images: thirty-four 3.2 mm thick slices aligned to the AC-PC plane, acquired using a "reverse" spiral scanning protocol (TR = 1500ms, TE = 25ms, flip = 70°, FOV = 200mm), which improved the signal-to-noise-ratio in areas of the brain susceptible to magnetic field inhomogeneity artifacts (Glover & Law, 2001). The basic principles of this technique and examples of its effects on fMRI signal are presented in Appendix A.

After detection and correction of outliers in k-space data, images were reconstructed, movement corrected (post-correction parameters revealed mean movement of less than 1mm in all subjects) (Woods, Cherry, & Mazziotta, 1992) and linearly detrended on a voxel-by-voxel basis to eliminate linear signal drifts within each run. Data were then co-registered to a common reference structural scan using a 2^{nd} order (60 parameters) non-linear warping algorithm (Woods et al., 1992). Finally, data were smoothed with a 6mm FWHM three-dimensional Gaussian filter. The reference brain used for co-registration was also warped to match the Montreal Neurological Institute reference brain, and the warping parameters obtained were down-sampled to the matrix of the statistical map of the functional data. This transformation matrix was applied, after the statistical analysis, to all statistical maps in order to facilitate mapping of activations onto the MNI brain and the Talairach space.

Statistical analysis was carried out on a voxel-by-voxel basis using random effects (subject as a random factor) repeated measures ANOVA model, with MR intensity in each scan of the epoch of interest as a repeated measure (MacDonald, Cohen, Stenger, & Carter, 2000; Ursu et al., 2003). The statistical F maps obtained were used for delimiting 2D (i.e. in-slice) clusters of active voxels based on voxel-wise alpha values and cluster thresholds determined through Monte Carlo simulations of fMRI data (Forman et al., 1995). Correcting for the 32 slices containing frontal cortex and for the resultant smoothing of the statistical maps (kernel width of

approximately one voxel dimension in all three axes), this ensured an image-wise protection against type 1 error of 0.05. Subsequently, the 2D voxel clusters were merged into functional regions of interest (ROI) by allowing clusters that were contiguous in adjacent slices to be collapsed into one ROI. Then time-series of signal change (normalized to the signal in the first scan of the trial) were computed, for the average signal of all voxels in the ROI as well as for the signal of the peak voxel (the voxel of maximum F). Activation in any ROI was interpreted if it withstood all of the following three criteria: a) at least one of the conditions had positive signal changes in the average ROI time series, b) the patterns of signal change in the time series of the peak voxel matched the average time series, and the Condition x Scan ANOVA of both maximum voxel and ROI-averaged signal changes were significant ($p < 0.05$); c) a significant planned independent samples t test ($p < 0.05$, 2 tailed) in the peak voxel time series. For this, the signal changes relative to scan 1 were computed for all middle scans (scans 2 through 7 for analyses of sample-target interval, or 9 through 15 for the inter-trial interval), in the timeseries of the peak voxel; then averages of these signal changes at each scan were computed, for each condition of interest. Then the timecourse of these signal changes was plotted, and the scan in which at least one condition had a positive signal change and with the highest difference between the means of each condition was identified. The planned t test was conducted, for the entire group of subjects, on these mean values of the signal change for each condition of interest.

2.3.Results

2.3.1. Behavioral results

As can be seen in Figure 2, which summarizes the significant behavioral effects observed in this experiment, both demands for control and the presence of incentives had the intended effects on

performance. In the Task x Incentive ANOVA of mean reaction times, the effect of demands for control were reflected in a significant main effect of task [$F(1, 18) = 23.34$, $p < 0.001$] which confirmed that subjects took longer to respond to the Non-match than to Match targets. The effects of incentives were also clearly reflected in the latency of responding: since rewards could only be obtained if the responses fell under the 1 second deadline, subjects responded faster to Reward trials than to No-reward trials (main effect of Incentive, $F(1, 18) = 11.50$, $p = 0.003$). There was no interaction between incentives and task [$F(1,18) = 0.01$, $p = 0.92$]. Subjects committed errors in approximately 10% or trials, and error rates were not differentially affected by either task [$F(1, 18) = 0.97$, $p = 0.34$] or incentive level [$F(1, 18) = 0.20$, $p = 0.66$].

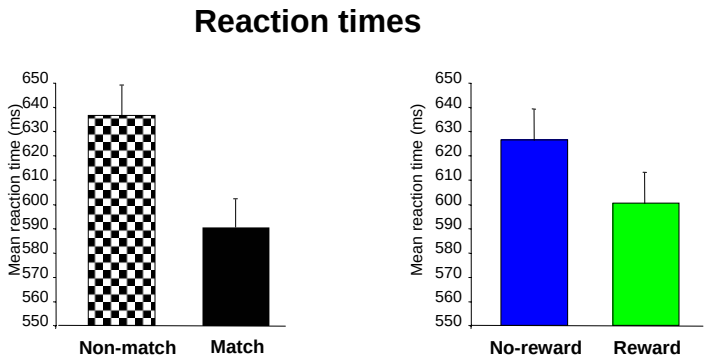

Figure 2. Behavioral effects of task and incentives in Experiment 1.

Subjects responded significantly faster to targets of Match than Non-match trials (left graph). Regardless of task performed, they were also faster in responding during Reward trials compared to No-reward (right graph).

2.3.2. Effects of demands for control (Task x Scan interaction)

As expected, the interaction of Task by Scan identified several prefrontal regions with differential signal dynamics in the Match relative to Non-match tasks (see Table 1). Examination of the signal changes relative to the first scan of each trial revealed higher activation during preparation for Non-match trials in the motor and premotor cortex, pre-SMA (BA8) and left lateral PFC (see Figure 3 and Table 1). The peak voxel of one of the main left lateral PFC

regions of interest was localized in Brodmann's area (BA) 10 (-36, 53, -4), and extended posteriorly and dorsally in the neighboring area BA46. Most importantly, this analysis did not isolate any regions of significant Match/Non-match differences in the OFC (see Figure 3). As can be seen in Table 1, the most ventral a cluster with significant Task x Scan interaction was identified in the left inferior convexity, which the Talairach transformation localized in BA10/47 (middle frontal gyrus, peak voxel at Talairach coordinates -47, 41, -7).

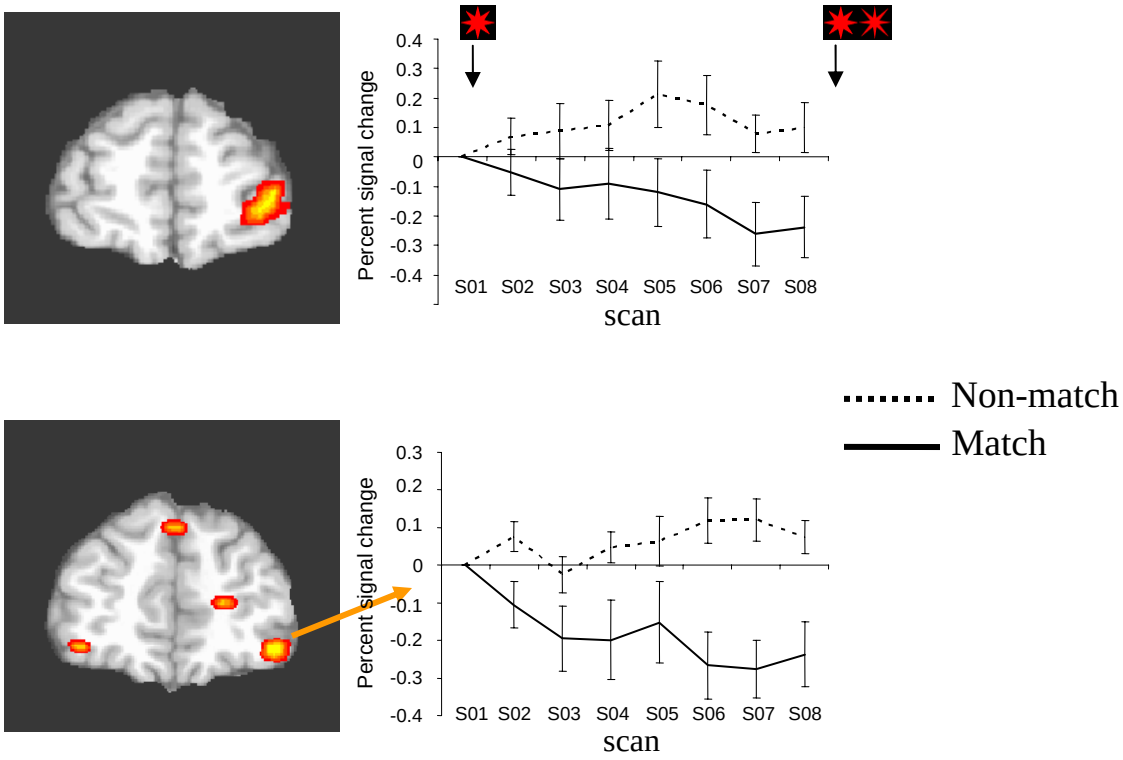

Figure 3. Effects of task on preparatory activity in the prefrontal cortex (Experiment 1)

The left lateral PFC (BA 10/46, Talairach coordinates of peak voxel -36, 53, -4) showed more activity during the sample-target interval of Non-match trials (dotted line). The left inferior convexity (BA47, -47, 41, -7) showed a similar pattern of activation, while the homologous area on the right showed deactivations to both conditions.

Table 1. Task by Scan interaction during the preparation interval of Experiment 1

Region	Laterality	BA	Direction of peak activity	# voxels	Tal x	Tal y	Tal z
Sup. medial frontal g.	B	8	Non-match	5	-5	39	40
Middle frontal g.	L	10/46	Non-match	26	-36	53	-4
Inf. frontal g.	L	10/47	Non-match	12	-47	41	-7
Precentral g.	L	6	Match	48	-35	-20	65
Sup. Parietal	L	7	Match	7	-32	-47	70
Middle temporal	B	21	Non-match	32	62	-23	-11
Sup. medial frontal g.	R	8	Deactivation	4	8	35	48
Inf. frontal g.	R	47	Deactivation	4	36	39	-8

To ensure that the lack of OFC activity in response to manipulation of demands for control was not simply a false negative result, we performed several additional analyses. Firstly, we relaxed the threshold used for examining the statistical maps to a level which corrected only for the few slices containing OFC (voxel-wise $p = 0.01$, 2D clustering threshold of 3 contiguous voxels, as per Forman et al., 1995) and verified that no orbitofrontal areas showed increased activity during Non-match trials. Secondly, we sought to confirm that the behavioral task placed significant demands on inhibitory control processes at the time of responses to targets. For this, we performed an exploratory analysis of the BOLD responses time-locked to the scan when the response was produced. This analysis was a Task x Scan interaction similar to that performed on the sample-target interval (scans S1-S8), except this time was performed on the scans covering the post-response interval (scans S9-S16). Examination of the activation areas revealed two clusters of active voxels in the left BA 9, several other areas of activation in the prefrontal and parietal cortices, but no significant Task by Scan interactions in the orbitofrontal cortex (all active regions had peak voxels at Talairach coordinates more than 11mm dorsal to the AC-PC plane, see Table 2). In parallel to the lack of activations in the orbitofrontal cortex, this analysis

identified cortical areas previously involved in suppression of interference from competing response sets (Bunge, Dudukovic, Thomason, Vaidya, & Gabrieli, 2002; Konishi, Nakajima, Uchida, Sekihara, & Miyashita, 1998) and in response inhibition (Aron et al., 2004; Bunge et al., 2002; Waldvogel et al., 2000). Such areas included inferior lateral prefrontal, motor and premotor cortices. Since some of these areas may have been involved in task execution at the preparatory phase too, the post-hoc testing of the direction of activity was done on signal changes relative to the first scan of the trial (S01), to control for effects of any residual sustained task-related preparatory activity. This analysis showed higher activity after Non-match relative to Match targets in the majority of the regions identified by the response-locked exploratory ANOVA (see Table 2).

Table 2. Task by Scan interaction during the post-response interval of Experiment 1

Region	Laterality	BA	Direction of peak activity	# voxels	Tal x	Tal y	Tal z
Sup. frontal g.	R	6	Match=Non-match	13	13	21	57
Middle frontal g.	L	8	Match=Non-match	42	-37	16	46
Sup. frontal g.	L	6	Non-match	18	-21	-2	74
Medial sup. frontal	B	6	Non-match	22	5	-20	69
Medial frontal g.	L	8/9	Non-match	26	-9	43	35
Precentral g.	B	6	Non-match	11	65	-3	13
Middle frontal g.	R	10/9	Non-match	24	30	44	12
Inf. frontal g.	L	46/45	Non-match	4	-47	29	11
Sup. parietal	L	7	Non-match	189	-47	-61	50
Caudate	B	-	Non-match	37	-7	-1	12
Middle temporal	L	21	Match	5	-46	4	-25
Middle temporal	R	21	Match=Non-match	9	59	-54	4
Sup. temporal	R	42	Non-match	15	57	-17	12
Angular	L	-	deactivation	4	-39	-70	57

Finally, after identifying the OFC regions sensitive to changes in incentives (see section 2.3.4), we verified that the effect of task in that OFC region of interest was not masked by the presence of incentives in only half of the trials. Theoretically, this could have caused the No-reward trials to result in a reactive disengagement of control processes (due to lack of motivation to perform) to the point where it may have diluted a possible effect of task during the Reward trials. Therefore, we performed a confirmatory Task x Incentive x Scan ANOVA on the signal change of the OFC region identified by the Incentive x Scan analysis and found that the Task x Incentive x Scan interaction in those areas were highly non-significant (all $p > 0.4$).

2.3.3. Effects of incentives (Incentive x Scan interaction)

The Incentive by Scan interaction in the exploratory ANOVA of the fMRI data was significant in several regions of the prefrontal cortex, as well as in some subcortical structures, and temporal, parietal and occipital cortices (see Table 3).

Table 3. Incentive by Scan interaction during the preparation interval of Experiment 1

Region	Laterality	BA	Direction of peak activity	# voxels	Tal x	Tal y	Tal z
Middle frontal g.	L	6/8/9	Reward	19	-40	14	53
Precentral g.	R	6	Reward	34	59	6	35
Sup. medial frontal cortex	B	9	Reward	31	-9	54	30
Medial frontal g.	B	32/10	Reward=No-reward	10	-2	50	-1
Middle frontal g.	R	46/10	Reward	4	38	49	9
Inf. frontal g.	L	45	Reward	4	-56	19	5
Insula	L	13	Reward	4	-42	2	0
G. rectus	B	11	No-reward		-1	53	-18
Inf. frontal g.	R	44	No-reward	7	38	10	29
Lateral orbital cortex	B	11/47	No-reward	22	-24	36	-15
Postcentral	R	1/2	Reward	75	32	-34	71
Sup. Parietal	R	7	Reward	15	23	-58	65
Supramarginal	R	40	No-reward	4	58	-36	41
Postcentral	L	6	No-reward	3	-60	-3	30
Middle temporal	R	39	Reward	87	50	-67	25
Middle temporal	L	39	No-reward	43	-53	-62	24
Sup. Parietal	B	22	No-reward	43	67	-27	4
Amygdala	R		Reward	10	24	-9	-15
Thalamus	R		Reward	4	10	-30	5
Angular	L	40/7	No-reward	19	-41	-63	48
Precuneus	R	7	Reward	223	1	-56	43
Calcarine	B	18	Reward	44	-1	-91	23
Lingual	L	29	Reward	10	-14	-47	8
Fusiform	L	18	Reward	20	-26	-71	-2
Occipital	R	19	Reward	12	46	-74	-3
Sup. frontal g.	R	4	Decrease	11	13	34	55

As in the analysis of task effects, we performed planned t tests to determine the direction of signal changes in the OFC (left and right BA11/47 and medial OFC), as well as in other regions of interest in which it had been previously shown that neural activity changes in response to availability of rewards: medial PFC, lateral and dorsal PFC, amygdala/sublenticular extended amygdala. In the left lateral OFC (see Figure 4), peak activity during preparation for No-reward trials was higher relative to Reward trials in both clusters identified by the ANOVA [$t(18) = 4.19$

and t (18) = 4.06 respectively, p < 0.001], as was in the right lateral OFC [t(18) = 3.0, p = 0.008]. Examination of the signal changes in the medial OFC region of activation revealed a more complex pattern (see Figure 4): overall the activity in the Reward condition was stronger that in the more lateral parts of the OFC, and in scan 4 of the delay the means signal changes were even reversed (higher for Reward than Non-reward). While the peak difference (at scan 7) was still significantly higher in the No-reward condition [t(18) = 2.15, p = 0.04], we also tested reversed difference (at scan 4) and found it to be also significant [t(18) = 2.47, p = 0.02].

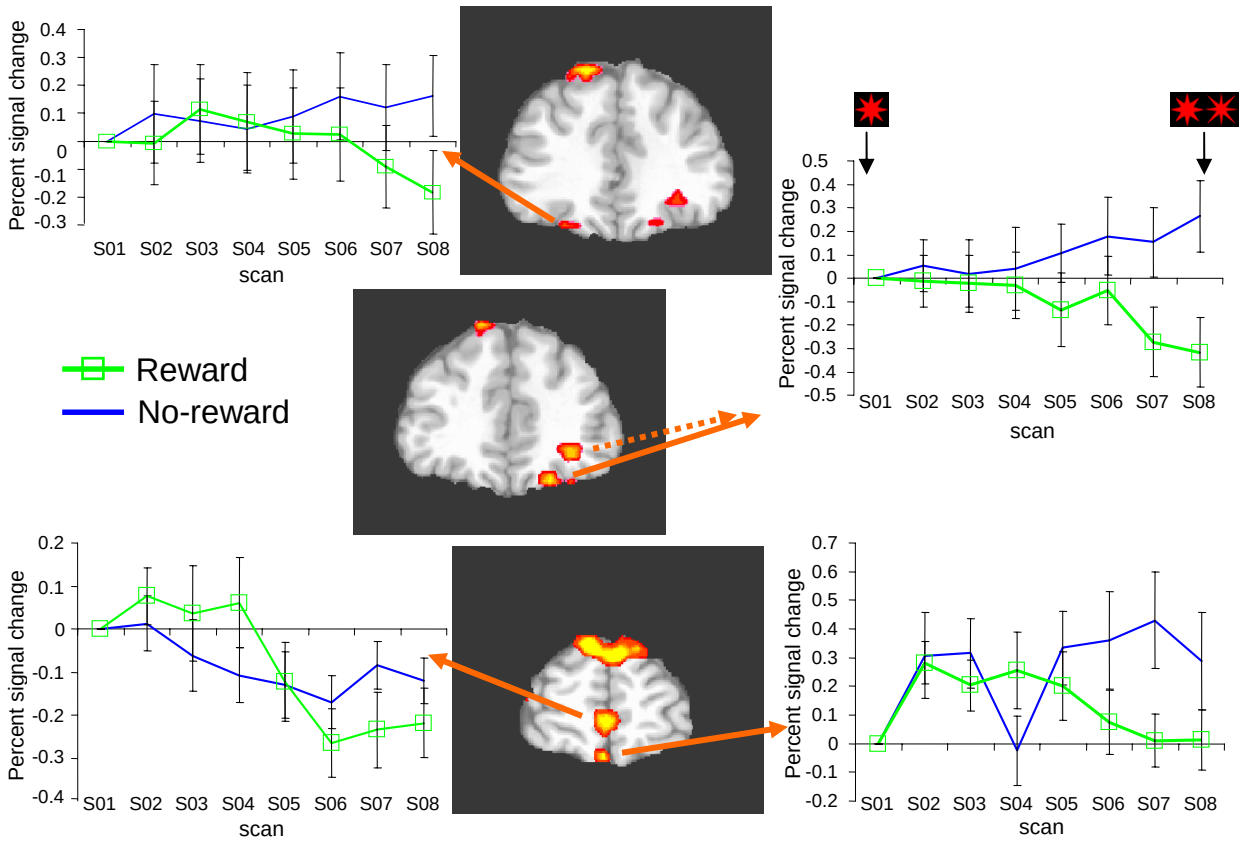

Figure 4. Orbital and medial PFC areas activated differentially by Reward vs. No-reward cues

Both left and right lateral OFC activity (BA 11/47) was significantly higher during No-reward trials relative to Reward trials. Both medial OFC and medial frontal cortex (BA 10/14 and BA 10/32, respectively, bottom graphs) showed evidence for increased Reward activity early during the delay. In the medial OFC region of interest an increase in activity to No-reward cues was noted late during the anticipatory interval (bottom right timeseries).

We also examined the effects of incentives in more dorsal aspects of the PFC (i.e. left lateral PFC, at the border of BA8, 9 and 6), as well as in the amygdala/sublenticular extended amygdala region of interest, and the results are summarized in Figure 5. In the lateral PFC, the activity peaked higher during Reward trials [$t(18) = 4.04$, $p = 0.001$], whereas the amygdala activation showed stronger activation to Reward early during the delay but more activity to No-reward trials in later scans (see Figure 5). As can be seen in Figure 4, a trend for similar pattern of signal changes was noted in the medial prefrontal cortex (BA 10/32, frequently associated with reward processing): the Incentive by Scan interaction in the signal changes was significant [$F(7, 126) = 5.00$, $p < 0.001$), and the means of signal change peaked higher for Reward early in the interval (scan 4), but the latter effect did not reach statistical significance in the post-hoc test ($p = 0.11$).

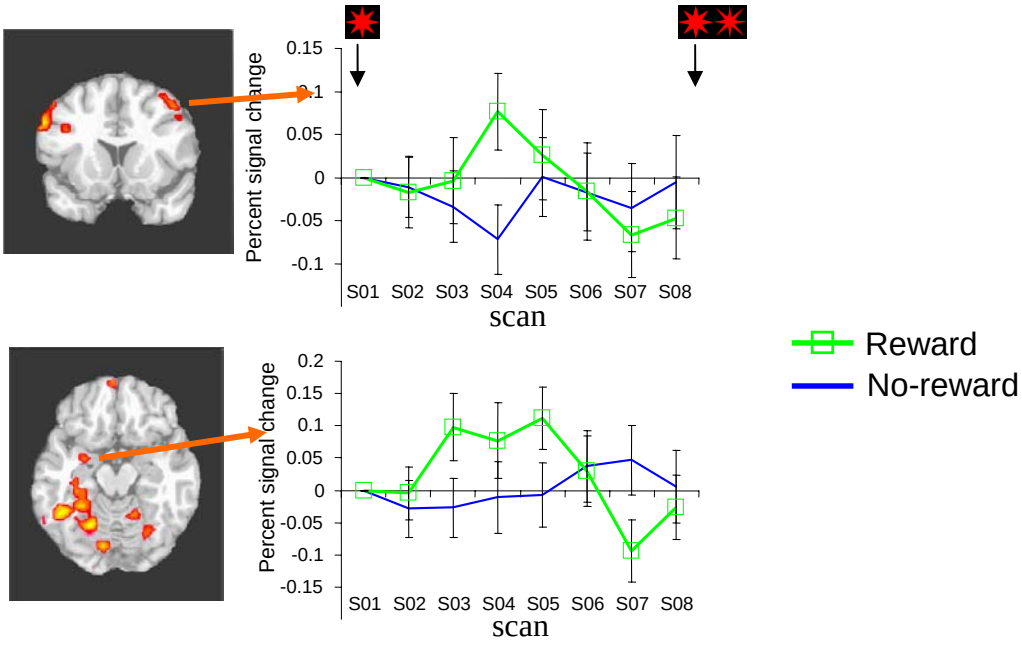

Figure 5. Other brain structures with increased activity in response to Reward cues

The effects of incentives were also significant in areas such as the premotor and dorsolateral PFC (top graph) and amygdala/sublenticular extended amygdala (bottom graph). Activity in both areas peaked higher during preparation for Reward trials than No-reward trials.

Since the lateral prefrontal areas showing more activity during preparation for reward trials were more posterior relative to those identified by the Task by Scan interaction, we verified the effect of incentives in the left lateral region of interest from the latter analysis. As in the lateral PFC region exemplified in Figure 5, the signal changes at all middle scans had higher means during preparation for Reward trials, but the effect was not statistically significant [$F(7, 126) = 0.56$, $p > 0.7$, see Figure 7).

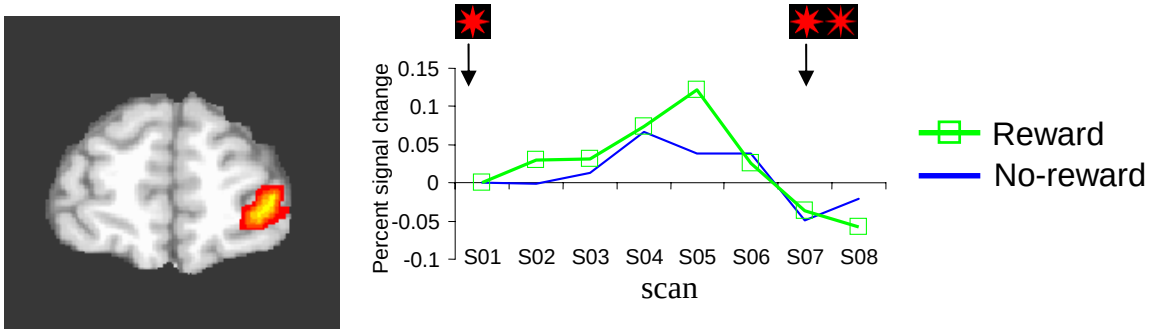

Figure 6. Effect of incentives in the left BA10/46

In the left lateral prefrontal region identified by the Task by Scan interaction, the signal changes was nominally higher after Reward than after No-reward cues, but the effect was not statistically significant.

2.4.Discussion

2.4.1. The human OFC represents potential outcomes but does not implement control

In this experiment, we manipulated demands for control and incentive for correct performance independently of each other, and thus were able to test the prediction that the activity in the

human OFC will parallel representations of potential outcomes rather than the engagement of control processes.

In this experiment, activity in the lateral OFC was modulated by whether subjects expected to win a reward for good performance or not (Reward vs. No-reward), and not by the degree to which the task they had to prepare to suppress a prepotent response tendency (Non-match vs. Match, respectively). Therefore, this study provided evidence in support of the hypothesis that the OFC is involved in representing the motivational context of actions (Hikosaka & Watanabe, 2000; Roesch & Olson, 2004; Tremblay & Schultz, 1999), and inconsistent with active implementation of control processes (Clark et al., 2004; Clarke et al., 2004; Fuster, 1997; Roberts & Wallis, 2000). This conclusion is strengthened by the finding that, in the analysis of incentive effects, *the highest activity in the OFC was seen in the condition associated with little engagement of control processes (No-reward).* These findings are summarized in Figure 7.

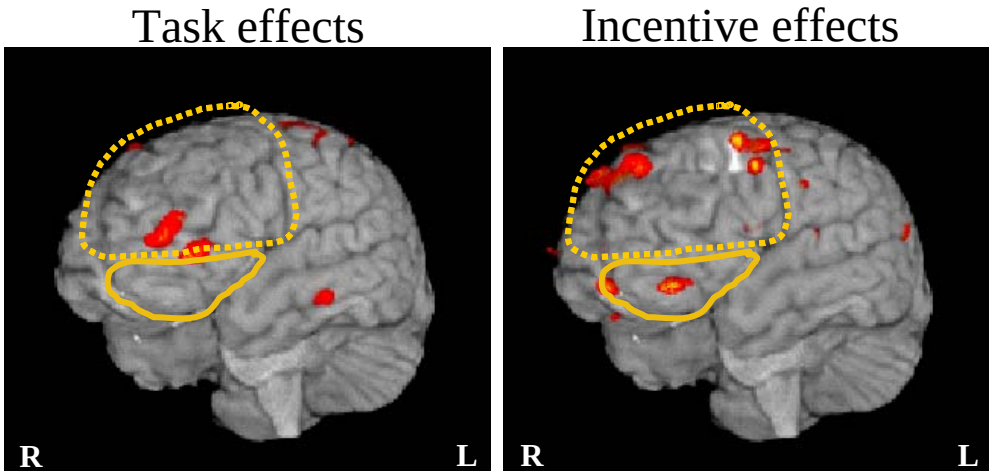

Figure 7. Summary of the main results of Experiment 1.

The boundaries of the lateral (dotted line) and the orbital (continuous line) PFC are outlined, to highlight the fact that the lateral PFC activity during preparation to respond was modulated both by task (left map) as well as by the manipulation of incentives (right map), consistent with a role in implementation of control processes; in contrast, the OFC only responded to manipulation of incentives, suggestive of a specific role in representation of potential outcomes.

Very recently, (Roesch & Olson, 2004) have used a similar approach in awake behaving monkeys and provided evidence supporting the hypothesis that the neurons in the OFC represent the hedonic value of rewards rather than the motivational impact that rewards have on executive functions. In OFC neurons, stimuli which included cues for immediate large reward elicited higher activity than stimuli which included predictors of either a short or a long time-out in reward delivery. Since monkeys performed better in immediate reward and long time-out trials relative to short time-out trials, the specificity of OFC neuronal activity to reward trials cannot be accounted for by an increase in control processes, and is therefore consistent with a representation of the reward value of the outcome. While some differences still exist between these findings and other non-human primate studies of motivational processing in the OFC (Hikosaka & Watanabe, 2000; Tremblay & Schultz, 1999, 2000b), these results are the first, in the non-human primate literature, to provide support for opposite predictions made by reward representation vs. active control hypotheses.

The patterns of activity in the lateral and orbital PFC observed in our study could account for previous findings used as arguments in favor of the proposed OFC involvement in inhibitory control. For instance, Roberts has argued (Clarke et al., 2004; Roberts & Wallis, 2000) that the OFC is involved in top-down inhibitory control of associations between stimuli or responses and affective valence, when such associations are no longer adaptive for behavioral regulation. This hypothesis was thought to be supported by findings that OFC lesions have effects on reversals of responses when reward contingencies change whereas lateral PFC lesions impair switches between higher level, abstract task rules (Clark et al., 2004; Dias et al., 1996a, 1996b; Roberts & Wallis, 2000). However, on closer examination, these deficits are only apparent on the first of a series of reversals, which seems to suggest that performing those tasks is a result of a complex

interaction between several mechanisms and subject to rapid compensation from intact neural circuits. Furthermore, computational modeling work has pointed out that such deficits could be accounted only by assuming a role for the OFC in representing featural information and for the DLPFC in representing more abstract, dimensional aspects of stimuli; this eliminated the need for a active control mechanisms to be implemented by the OFC (O'Reilly et al., 2002). Finally, a role for OFC in constructing representations of possible punishing outcomes, in interaction with a DLPFC involvement in active attentional bias can account for the deficits induced by OFC or DLPFC lesions on those tasks. As such, an impaired OFC would lead to rapid degradation of the information which associates a particular stimulus with a bad outcome, causing perseverative errors in trials immediately following a reversal in contingencies. However, a DLPFC lesion would not necessarily predict such an impairment in the presence of intact OFC: lack of expected reward in the first trial following the contingency switch would lead to automatic reward-seeking behaviors, which in subsequent trials would lead to choosing the alternative stimulus. Conversely, switching the association of rewards from one abstract dimension to another (e.g. from shapes to lines), has been tested by presenting the animal with new choices of shape-line combination, and measuring how fast the animal started responding to the alternate dimension. Again, one might indeed expect this behavior to be impaired by lesions in an area closely associated with representation of task context or with working memory (Cohen et al., 1997; Funahashi, Chafee, & Goldman-Rakic, 1993; Miller & Cohen, 2001), but not by OFC lesions, since the latter were known not to affect, in isolation, acquisition of novel object discriminations based on associations with rewards.

Our results of OFC engagement during an interval which eventually led to indices of poor performance speak against a general involvement of the OFC in control processes. However, it

could still be argued that the OFC is important for some specific control process necessary when switching from a task with low demands for control to one with high demands, which in this design were minimized by blocking Match and Non-match trials (this was done to avoid confounds from general processes associated with task switching, thought to depend on ventral areas of the lateral PFC (Konishi, Nakajima, Uchida, Kameyama et al., 1998; Sohn, Ursu, Anderson, Stenger, & Carter, 2000)). This possible interpretation will be addressed in the next chapter, by maintaining the current differences in demands for control between Match and Non-match tasks, and showing that other factors (i.e. differential outcome representation induced by changes in response accuracy between the two tasks) can account for OFC activations. Nevertheless, in order to verify that inhibitory control processes control processes were significantly engaged by performing the Non-match task, a second analysis of the post-response interval sought to parallel a previously published detailed examination of Go/No-go responses with the dominant hand (Waldvogel et al., 2000). In that study, authors used fMRI to scan human subjects while they performed a Go/No-go task with the dominant hand, and compared the differences in fMRI signal between Go and No-go trials in the pre-supplementary motor areas (pre-SMA) and in the controlateral motor cortex. While the activity in the pre-SMA did not differ between Go and No-go trials, it was significantly stronger for Go trials than for No-go in the motor cortex, consistent with effective suppression of the response with the right hand in the No-go condition. In our data, comparing executed responses of the right hand (i.e. right-sided sample, Match rule) with suppressed ones (right sided sample, Non-match rule) replicated the results of Waldvogel et al. (see Figure 8). This suggested that the demands on inhibitory control imposed by the Non-match task were comparable with those of a typical No-go task (details about statistical analysis are included in Appendix B).

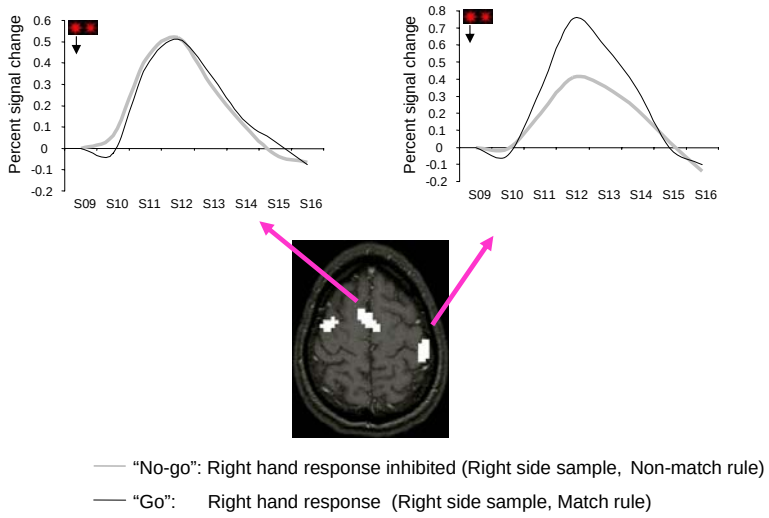

Figure 8. Neural correlates of inhibitory control during execution of Non-match responses

When the recognized stimulus was on the right side of the target, the pre-supplementary motor area shows equal activation regardless of whether the automatic response with the right hand was actually produced (i.e. Match trials, labeled "Go" in the figure), or withheld (i.e. Non-match trials, labeled "No-go"). In contrast, the motor cortex controlateral to the response hand (left motor cortex, shown on the right in the image) shows decreased activation when the automatic right hand response is inappropriate and is successfully suppressed (Non-match rule or "No-go", grey thick line) relative to when it is produced (Match rule, "Go", thin black line). This replicates the differences between classical No-go and Go trials, respectively, seen by (Waldvogel et al., 2000).

2.4.2. The OFC activity to cues signaling lack of reward

Aside from providing the critical piece of evidence for clearly refuting involvement of the OFC in active control of behavior, the increased activity in the lateral OFC during preparation for No-reward trials has important implications for understanding the role of the OFC in representation of outcomes and, consequently, in establishing motivational context. A few recent human imaging studies have interpreted patterns of activations in the lateral OFC as evidence for specific processing of punishments (Breiter et al., 2001; O'Doherty et al., 2001; Small et al., 2001). However, multiple similar studies report OFC activations in response to increasing rewards levels (Critchley, Mathias, & Dolan, 2001; Elliott et al., 2003; Elliott, Newman, Longe, & William Deakin, 2004; Elliott, Sahakian, Michael, Paykel, & Dolan, 1998; O'Doherty et al.,

2003; O'Doherty et al., 2002). Furthermore, some studies of non-human primates report identifying populations of OFC neurons which respond specifically to non-preferred foods (Hikosaka & Watanabe, 2000), or OFC neurons which increased their firing rates as learning of associations between stimuli and *absence* of rewards progressed (Tremblay & Schultz, 2000b). In contrast, other studies report that the majority of OFC neurons increase their activity to rewards or reward predicting stimuli (Roesch & Olson, 2004; Rosenkilde, Bauer, & Fuster, 1981; Tremblay & Schultz, 1999, 2000a; Wallis & Miller, 2003). These contrasting findings are reminiscent of another subtle but important difference between two types of hypotheses regarding OFC function. Some authors emphasize its role in reward processing while pointing out its strong connectivity within the dopaminergic system (Elliott, Dolan et al., 2000; Elliott et al., 2003; O'Doherty et al., 2002; Rolls, 2000; Rolls, 2004; Rolls et al., 1996; Volkow et al., 2001). On the other hand, evidence accumulates that activity of important dopaminergic targets such as the striatum is modulated by unpredictability or saliency of both rewards and punishments (Delgado, Nystrom, Fissell, Noll, & Fiez, 2000; McClure, Berns, & Montague, 2003; Pagnoni, Zink, Montague, & Berns, 2002; Zink, Pagnoni, Martin, Dhamala, & Berns, 2003; Zink, Pagnoni, Martin-Skurski, Chappelow, & Berns, 2004). Overall, these results point out that more controlled, hypothesis-driven experiments are needed in order to precisely define the involvement of the human OFC in motivational processes.

It has been shown, in the human decision-making literature, that decisions are influenced by expected utilities, and utilities are modulated by the disappointment and elation that are expected given the alternative outcomes (Mellers, 2000). In this context, the expected absence of reward renders the No-reward trials of this task a relatively punishing situation (by comparison to the almost certain gain experienced in Reward trials). Consequently, the location of the OFC

activity noted during the preparation interval of No-reward trials is consistent with the view that lateral OFC regions are specifically involved in processing punishments. Consequently, the contrast between these findings and the aforementioned literature emphasizing the OFC's role in processing rewards became part of the motivation for the subsequent study described in this thesis. The need for further exploration of the regional sub-specialization within the human OFC became obvious after examination of the pattern of activation in ventro-medial prefrontal regions identified by the exploratory analysis. In a medial OFC region of interest, post-hoc tests of the direction of activation showed significantly increased activity in Reward trials at scan 4. Furthermore, a trend for increased Reward activity was also noted in the medial prefrontal cortex (BA10/32). These findings would be consistent with the hypothesis that, in contrast to the lateral OFC regions, the medial OFC is more involved in processing rewards than punishments (O'Doherty et al., 2001; Small et al., 2001). However, a close examination of the timeseries in these medial region revealed a quite complex pattern of signal change, in which the early signal changes in response to reward cues was higher than to those signaling lack of reward, but these differences tended to reverse later in the interval. One explanation for this pattern of activity is that the magnitude of rewards used in this study was simply not sufficient to elicit robust activation of these structures. Another possibility is that the medial OFC responds to rewards under conditions of uncertainty, which was minimized in this design. However, the latter explanation, though it has been previously formulated in the context of gambling-type paradigms (Rogers et al., 1999), does not clearly distinguish between processing of rewards and punishments since, as we mentioned before, uncertain rewards can be viewed as potential punishments. Therefore, more research appears to be necessary to establish the degree of

specificity of the medial OFC activation for rewards; this warrants caution in formulating hypotheses regarding the precise role of the medial OFC in representation of outcomes.

CHAPTER 3

Representations of negative outcomes in the OFC

3.1.Introduction and rationale

In the previous chapter we presented results which were consistent with a role of the human lateral OFC in representation of potential outcomes of actions and not in active implementation of top-down control. In support of this hypothesis, neither increasing the demands for control (i.e. Non-match trials relative to Match) nor the engagement of control in response to rewards (Reward vs. No-rewards contrast) resulted in increased activity on the orbital surface of the prefrontal cortex. This paralleled well the subjects' response accuracy, which was not increased in Non-match relative to Match trials. However, it could still be argued that the differences in task design relative to other studies which reported OFC activations (Casey et al., 1997; Menon, Adleman, White, Glover, & Reiss, 2001; O'Doherty et al., 2003) were responsible for the lack of OFC activity. Therefore, we sought a more direct test of the hypothesis that OFC activations in cognitive control tasks reflect in fact the impact of errors, which are more frequently committed in conditions which, coincidentally, require engagement of control processes. Previous studies have generated results in which activity in the OFC (or areas directly connected to it) was noted when subjects made more errors (Casey et al., 1997; Cools, Clark, & Robbins, 2004), and when subjects evaluated stimuli which repeatedly switched their motivational significance between "reward" and "punishment" (Cools et al., 2004; O'Doherty et al., 2003; O'Doherty et al., 2001). However, these studies did not include the control condition, involving significant demands on

control processes but absence of errors. This experimental condition was achieved in our first experiment presented in Chapter 2. Consequently, one aim of the present study was to use the same cognitive task as in the previous experiment, while manipulating its parameters such as to result in a higher rate of commission errors. Thus, we were able to test the hypothesis that increases in errors are indeed associated with increased activity in the OFC.

Another important aim of this study was to clarify the nature of outcome representations supported by the OFC. As was mentioned before, it is still unclear whether the human OFC responds to rewards, punishments, or both. The experiment described in Chapter 2 identified robust activation in response to stimuli which predicted likelihood of a relatively punishing outcome (i.e. lack of reward). However, the first experiment lacked a condition with possible consequences that were punishing in themselves. Consequently, we added a third incentive level in which subjects would be motivated to perform in order to avoid a punishing outcome (i.e. monetary penalty). This addressed several issues, which will be outlined below.

Firstly, while the No-reward condition of Experiment 1 was critical for distinguishing control processes from representation of outcomes, it also made the interpretation of the lateral OFC activation more difficult to generalize. The presence of Penalty trials allowed us to directly test the prediction that the lateral OFC activity would be highest in preparation for those trials.

Secondly, in a study of brain activations to monetary gains and losses in a “wheel of fortune” paradigm, Breiter et al. (2001) found preliminary evidence that structures closely connected to the OFC might be part of the neural substrates of so called “counterfactual effects” of alternative outcomes on monetary gains and losses (Larsen, McGraw, Mellers, & Cacioppo, 2004; Mellers, Schwartz, Ho, & Ritov, 1997). Counterfactual effects refer to the fact that a certain outcome can be perceived as more or less favorable depending on what the alternative

outcome could have been. For instance, most people report feeling better upon a 50$ outcome of a gamble when the odds of winning are 10% than when the odds are 90%. These effects have been shown to influence in the same way anticipated feelings (Mellers, Schwartz, & Ritov, 1999). Therefore, in this study we sought to investigate whether the OFC might be part of the neural substrates of counterfactual effects on anticipated feelings, by examining the differences in OFC anticipatory activity to Reward and No-reward trials in the current design relative to the activity observed in Experiment 1.

In order to clarify these issues, the present experiment used a modification of the task described in Chapter 2. Subjects were scanned while responding according to the same "match" and "non-match" rules, and the incentives for correct performance were again varied from trial to trial. Instead of a fixed response window, as in Experiment 1, this time we used an individually titrated response deadline, which shortened when subjects made frequent correct and fast responses to incentive trials, and lengthened if the subjects' performance degraded. This had the effect of inducing differential error rates across experimental conditions, which allowed us to test the prediction that when demands for control will be accompanied by increased error rates, they will be associated with increased OFC activity.

In this design, a third stimulus color was introduced, which informed the subjects that incorrect or slow responses will result in a penalty. Introducing the Penalty trials was aimed at addressing two related questions. Firstly, it allowed us to directly test the prediction that the anticipatory OFC activity is specific to negative outcomes. Specifically, we expected to find that the greatest OFC activation during preparation for Penalty trials. Secondly, the changes in magnitudes and probabilities of all possible outcomes in this study (relative to the previous experiment), allowed us to examine whether the "counterfactual effects" of alternative outcomes

are present in the OFC during outcome anticipation. Furthermore, we could investigate whether the OFC activity might be consistent with predictions from either one of two main theories of decision-making: prospect theory (Kahneman & Tversky, 1988) and decision affect theory (Mellers et al., 1997).

3.2.Methods

3.2.1. Subjects

Seventeen adult right handed subjects (8 females, age range 18-38) participated in this study. Volunteers were screened for history of neurological or psychiatric disorders prior to enrollment in the study, and for substance dependence or abuse in the six months prior to the date of the interview. The experimental protocol was approved by the Institutional Review Board of the University of Pittsburgh.

3.2.2. Behavioral task

In this experiment, the behavioral task was very similar to that used in Experiment 1. The same type of stimuli were used (four star shaped variants), the stimulus presentation time was identical (1 second for sample and target, 500ms feedback, 12 second stimulus onset asynchrony). The two major differences from the previous task were: 1) the addition of a third stimulus color, coding for trials in which correct and fast responses were not rewarded, but in which errors or late responses were penalized with a loss from the total winnings, and 2) unbeknownst to the subjects, the response window was dynamically adapting to each subject's individual performance. Subjects were informed that the response deadline will be slightly shorter than the stimulus presentation time, and therefore they should be as fast as possible in order to maximize their winnings. During pre-test practice, the response window was set at 750ms. Within each

block of trials, as subjects became more proficient at performing the task, streaks of 5 correct and fast responses to incentive trials decreased the response window by 10%. After a deadline decrease, a correct but slow response or a streak of two errors or non-responses increased the deadline by 10%. This modification was implemented in order to avoid speed-accuracy trade-offs in the penalty condition, which would have complicated the interpretation of increased activation during Penalty trials relative to Reward trials.

Other modifications adopted in Experiment 2 were: the amount of the reward and penalty used in each trial was increased to $1, to preserve the theoretical total winnings available in the experiment compared to Experiment 1; responding was done using the index and middle finger of the right (dominant) hand, in order to eliminate interpretive problems related to preparation and execution of bimanual responses.

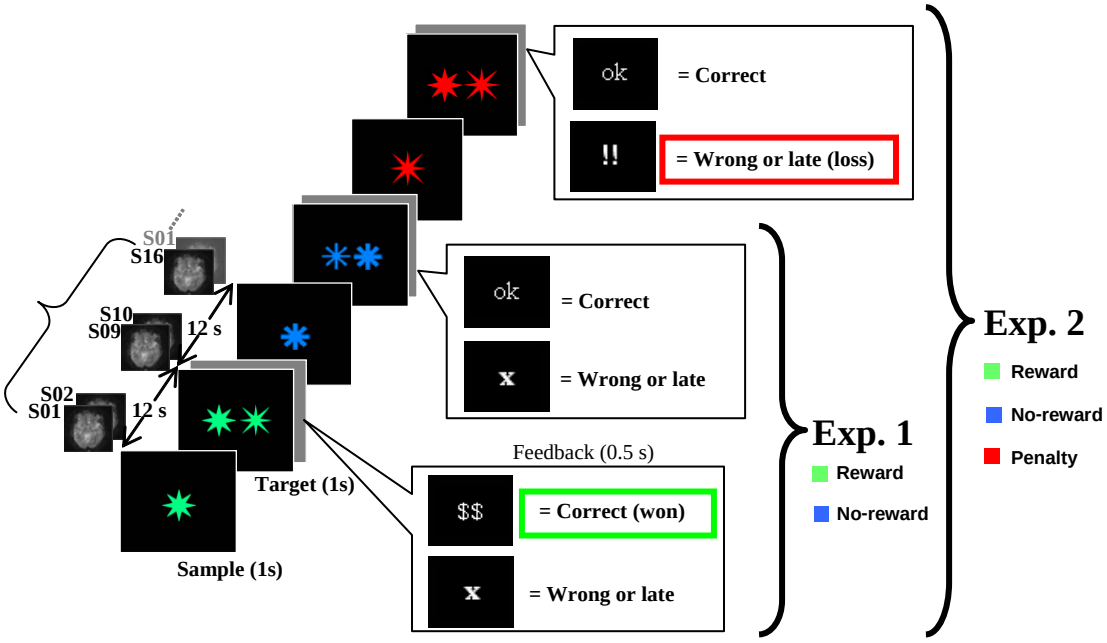

Figure 9. Task design in Experiment 2.

The two response rules (tasks) were the same as in Experiment 1: Match (identify the sample in the target stimulus, press the finger on the same side) and Non-match (identify the sample, press the finger on the opposite side). Similar to Experiment 1, two stimulus colors (green and blue in the example) meant that a reward ($1) could be won or not, respectively. A third color (red in this example) informed subjects that a correct and fast response will not change their winnings, but an error or a late response would decrease their winnings by $1. The color-incentive mapping was counterbalanced across subjects. Subjects were informed that late responses counted as errors, and that the response window was very short and therefore they should be as fast and as accurate as possible, in order to maximize their winnings. Unbeknownst to subjects, the response deadline shortened by 10% after streaks of 5 correct and fast responses to incentive trials, and lengthened by 10% when a deadline decrease was followed by a correct but slow response, or by streaks of two errors or no responses.

3.2.3. fMRI data collection and statistical analysis

The fMRI acquisition methods were identical to those described in Chapter 2, with the exception of acquisition plane (coronal, i.e. perpendicular to the AC-PC plane), slice thickness (5mm), and number of slices in each volume (28 per volume, no inter-slice gap). The statistical methods used in analyses were also similar: we used the same pre-processing steps as described in section 2.2.3., with the exception of the choice of a reference brain used as common space for data co-registration, which in this case was a version of the Montreal Neurological Institute reference

brain, also used as reference space by other fMRI analysis software packages. As in the experiment described in Chapter 2, the statistical model used in analyses was a random effects, repeated measures ANOVA of the MR signal during the scans covering the interval of interest. The criteria used for planned tests of the direction of signal changes were also identical to those described in section 2.2.3.

3.3.Results

3.3.1. Behavioral results

Task x Incentive ANOVAs of mean reaction times and error rates showed that, as in the previous experiment, performance was robustly influenced by the type of task performed as well as by presence or absence of incentives (Figure 10). The effects on the latency of correct responses were similar to those observed in Experiment 1: the task manipulation resulted in slower responding to Non-match trials compared to Match ones [$F(1, 16) = 17.72$, $p = 0.001$], while varying incentives lead to faster responding to both Reward and Penalty trials relative to No-reward ones [$F(2, 32) = 14.78$, $p < 0.001$]. In terms of accuracy, in this experiment both the task and the incentive manipulations impacted performance. The subjects' error rates were higher in Non-match trials relative to Match trials [$F(1, 16) = 13.63$, $p = 0.002$], and in the No-reward trials relative to either Reward or Penalty trials (see Figure 11). The differential effect of incentives on error rates was apparent both as overall accuracy [$F(2, 32) = 3.62$, $p = 0.04$] and as frequency of responses that were both correct and fast [$F(2, 32) = 10.80$, $p < 0.001$]. All measures of performance (reaction time and accuracy) showed equal improvement from positive and negative incentives, relative to neutral trials (t tests of Penalty vs. Reward trials yielded p

values of 0.3, 0.5 and 0.9 respectively). Similar to Study 1, there was no interaction between Incentives and Task in either latency or accuracy of responding (p values greater than 0.35).

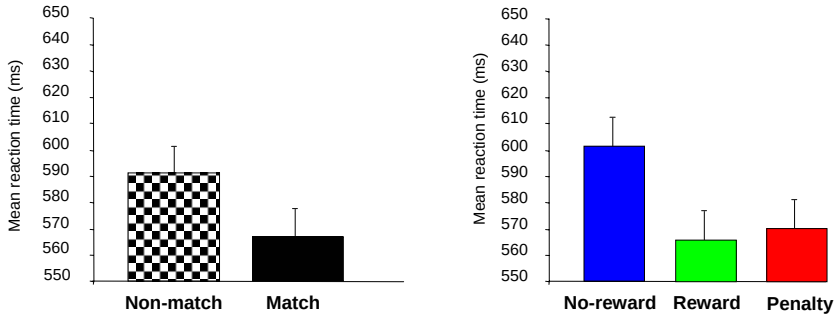

Figure 10. Reaction times of correct responses in Experiment 2

As in Experiment 1, subjects took longer to perform the Non-match task relative to Match (left graph), and responded faster in trials in which monetary incentives were available (regardless whether they were seeking to obtain a reward or to avoid a penalty) than when no incentives were used (right graph). The difference between RTs of Reward and Penalty trials were not significant ($p > 0.3$).

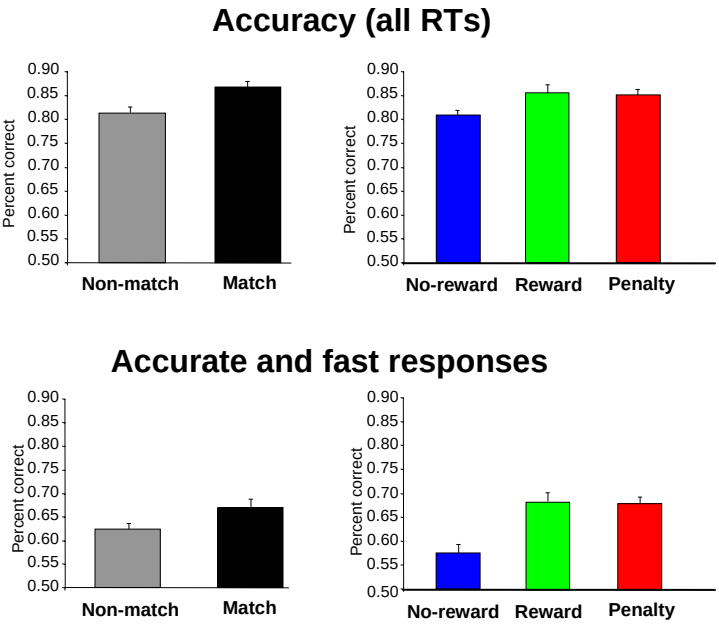

Figure 11. Accuracy effects of task and incentives in Experiment 2

The proportion of correct responses in Match vs. Non-match trials (left) and in Reward vs. Penalty vs. No-reward (right). The top graphs represent the accuracy of responses recorded in the 1 second of target presentation, regardless whether responses were faster or not than the dynamic response deadline. The bottom graphs represent the percentage of responses that were both correct and faster than the dynamic response deadline.

3.3.2. Effects of control (Task x Scan interaction) during preparation

As in Experiment 1, preparation for Non-match trials was associated with higher activity relative to Match trials in the left lateral PFC, replicating the findings of the first study. In this version of the task, this activation appeared to peak more posterior and more superior (including voxels in BA 45, 9 and 46, peak Talairach coordinates -52, 20, 16) relative to the one identified by Experiment 1 (see Table 4 and Figure 12).

Importantly, as predicted by the presence of differential error rates in the two tasks (Non-match vs. Match), this analysis also identified effects of task in the lateral OFC. The left lateral OFC activation (BA 11/47, peak voxel at Talairach coordinates -34, 33, -10) overlapped in part with areas showing effects of incentives in both experiments. Examination of the signal time series confirmed that, even though the task manipulation was very similar to that of the previous experiment, the Non-match task now showed evidence of increased OFC activity relative to Match, paralleling the error rates profile (Figure 11): in the left OFC the Task by Scan interaction was significant in the signal changes [$F(7, 112) = 3.19$, $p = 0.004$], and the planned t test at scan 7 showed a trend for higher activity in preparation for Non-match [$t(18) = 1.39$, $p = 0.18$]. In the right OFC ROI the timeseries pattern was more complex, with a significant Task by Scan ANOVA [$F(7,112) = 4.19$, $p < 0.001$], and a maximal difference between condition means in the opposite direction (significant t test of signal changes at scan 4, $t(18) = 2.17$, $p = 0.04$, Match > Non-match). In order to clarify these effects we also ran t tests in the ROI-averaged data, and found the effects to be in the same direction as in the peak voxel tests: for the left OFC, Non-match > Match [$t(18) = 2.28$, $p = 0.04$], while in the right OFC Match > Non-match, [$t(18) = 2.31$, $p = 0.035$].

Table 4. Task by Scan interaction during the preparation interval of Experiment 2

Region	Laterality	BA	Direction of peak activity	# voxels	Tal x	Tal y	Tal z
Sup. frontal g.	R	8	Match	8	19	26	38
Middle frontal g.	L	9/46	Match	14	-37	40	28
Middle frontal g.	R	6/8	Match	4	38	17	47
Precentral g.	L	4	Match	68	-55	-17	46
Inf. frontal g.	R	44	Match=Non-match	7	44	5	11
Lateral orbital cortex	R	11	Match=Non-match	8	25	24	-15
G. rectus	B	11	Match=Non-match	4	-6	58	-11
Caudal orbital cortex	R	11/13	Match=Non-match	4	22	14	-18
Medial sup. frontal cortex	B	6	Non-match	82	-3	12	50
Medial frontal cortex	L	8	Non-match	8	0	36	41
Inf. frontal g.	L	45/9/46	Non-match	68	-52	20	16
Lateral orbital cortex	L	11/47	Non-match	8	-34	33	-10
Sup. Temporal	L	42	Match	5	-47	-19	15
Sup. Temporal	L	21/22	Non-match	4	-55	9	-6
Thalamus	R		Match=Non-match	4	9	-29	12

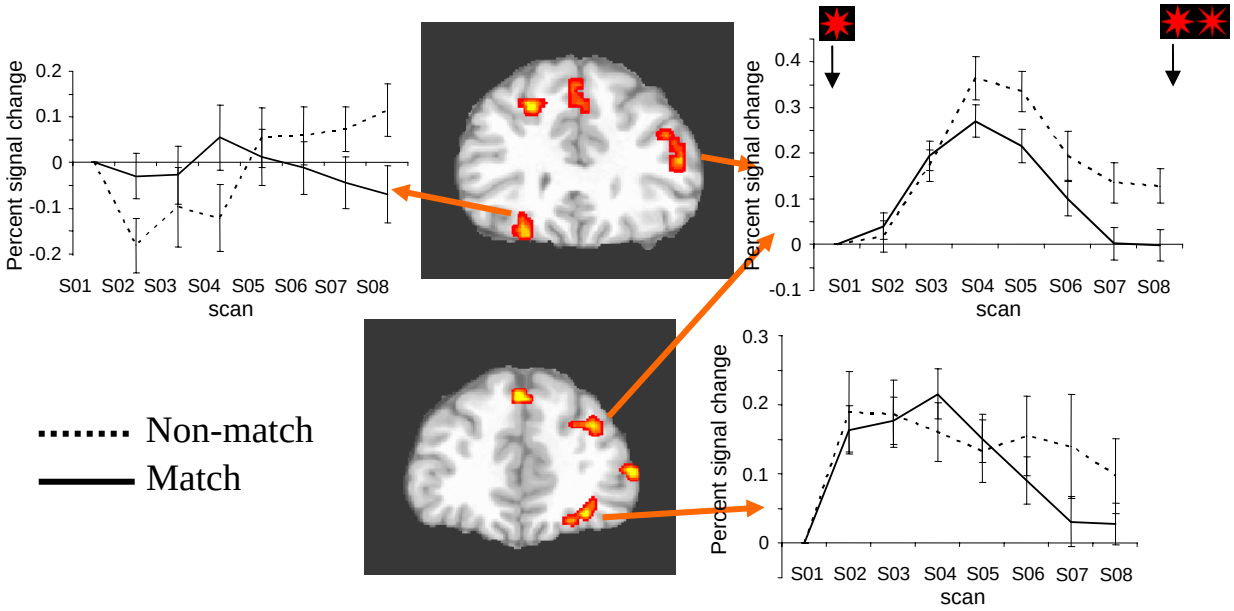

Figure 12. Effects of task on preparatory activity in the prefrontal cortex (Experiment 2)

Together with areas of the lateral prefrontal cortex, in this experiment the lateral OFC also shows evidence of higher activity during preparation for Non-match (dotted line). In the OFC, the higher peaks of activity during Non-match trials reached statistical significance when the t tests were conducted on signal changes of the entire region of interest. The left DLPFC (top right graph) showed the expected increase in activity during Non-match, while the right OFC showed an effect in the opposite direction.

Finally, to parallel the analyses performed in Chapter 2, we verified that effects of task in the OFC were not manifested as interactions between task and incentives. To this end, we performed an exploratory ANOVA of Task by Incentive by Scan and confirmed that there were no activation areas in the OFC. Furthermore, the Task x Incentive x Scan interaction was not significant [$F(14, 224) = 0.50$, $p > 0.93$] in the peak voxel time series of the left OFC region of interest identified by the Task by Scan interaction.

Similar to Experiment 1, we also performed an analysis of the Task x Scan interaction time locked to target presentation and response. As in the previous experiment, the most ventral focus of activation was laterally located in the left inferior frontal gyrus (BA 47, Talairach coordinates: -52, 24, -7, see Table 5), confirming that active response inhibition did not activate the OFC.

Table 5. Task by Scan interaction during the post-response interval of Experiment 2

Region	Laterality	BA	Direction of peak activity	# voxels	Tal x	Tal y	Tal z
Sup. frontal g.	R	8	Match=Non-match	5	3	36	49
Middle frontal g.	L	8	Match=Non-match	4	-25	-3	45
Precentral g.	R	6	Match=Non-match	7	55	-8	31
Precentral g.	B	3/1	Non-match	42	-41	-17	43
Inf. frontal g.	L	47	Non-match	8	-52	24	-7
Sup. Temporal	R	22	Non-match	4	58	0	-1

3.3.3. Effects of incentives (Incentive x Scan interaction) in Experiment 2

The Incentive x Scan exploratory ANOVA during preparation to respond revealed many of the same areas identified by the incentive analysis of Experiment 1 (see Tables 3 and 6). Such areas included OFC, DLPFC, medial prefrontal, supplementary motor and premotor areas, thalamus, medial temporal lobe areas and extrastriate cortex. Most importantly, it confirmed that the lateral

OFC reacted most strongly to samples predicting high likelihood of a negative outcome. As can be seen in Figure 13, this analysis identified lateral OFC activations bilaterally, localized to BA 11/47. The specificity of activation to Penalty cues was clearer in the left hemisphere (Talairach coordinates of the peak -31, 29, -10, Penalty > Reward at scan 5, t(16) = 2.20, p = 0.04), and approached significance in the right hemisphere (Penalty > Reward at scan 6, t (16) = 1.85, p = 0.08).

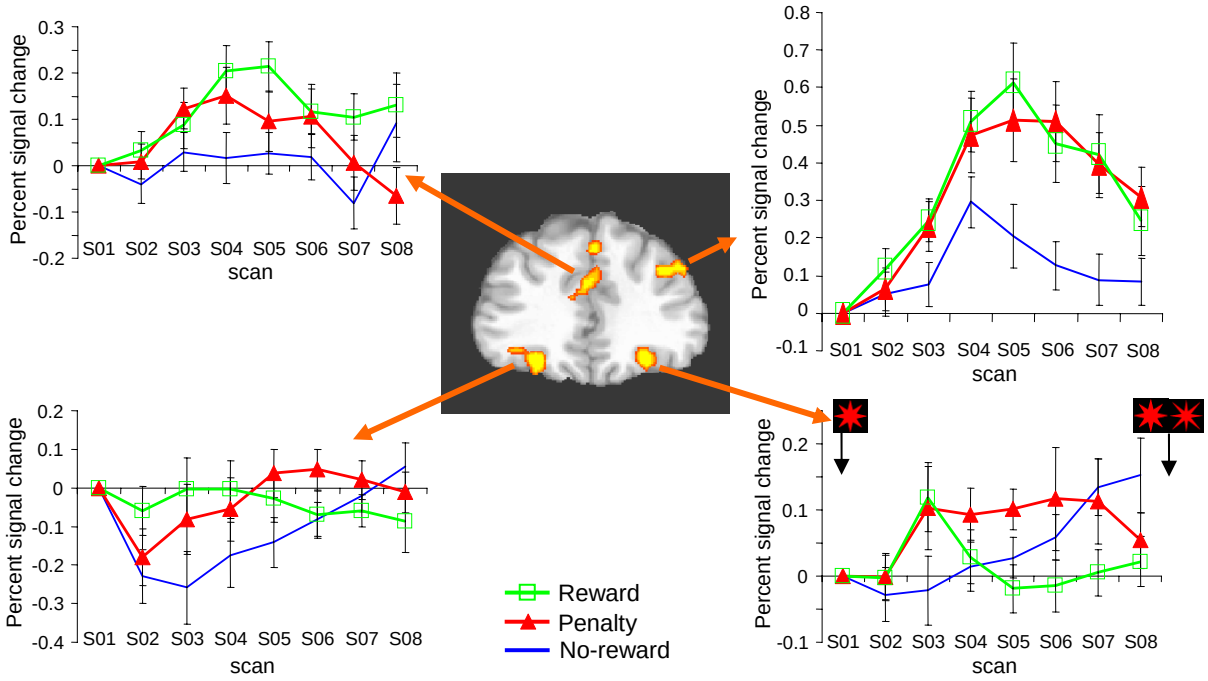

Figure 13. Prefrontal cortical activity after Reward vs. No-reward vs. Penalty cues

The figure depicts several prefrontal areas with significant Incentive by Scan interaction in Experiment 2. While the OFC regions of interest showed evidence for more sustained activation during Penalty trials relative to both Reward and No-reward (bottom graphs), the left DLPFC (top right) activated equally strongly in Reward and Penalty trials relative to No-reward. In the medial frontal cortex (top left), the Reward trials elicited the strongest activity.

Table 6. Incentive by Scan interaction during the preparation interval of Experiment 2

Region	Laterality	BA	Direction of peak activity	# voxels	Tal x	Tal y	Tal z
Sup. medial frontal cortex	L	8	Rew > Pun > Nrew	22	-6	26	44
Inf. frontal g.	R	9/44	Rew > Pun > Nrew	4	47	16	33
G. rectus	B	11	Rew > Pun > Nrew	10	0	43	-11
Precentral g.	R	6/8	Rew > Pun = Nrew	8	50	-12	46
Anterior cingulate	L	32	Rew > Pun = Nrew	17	-3	30	26
Sup. medial frontal cortex	L	6	Rew = Pun > Nrew	21	-3	3	68
Precentral g.	B	6	Rew = Pun > Nrew	48	-37	-2	57
Middle frontal g.	L	9/8	Rew = Pun > Nrew	46	-44	40	26
Inf. frontal g.	L	44	Rew = Pun > Nrew	13	-47	6	25
Lateral orbital cortex	B	11	Pun > Rew > Nrew	34	-31	29	-10
Inf. frontal g.	L	47	Pun > Rew = Nrew	8	-50	24	-10
Thalamus	L		Rew > Pun > Nrew	77	-6	-29	9
Caudate (head)	L		Rew > Pun > Nrew	6	-12	10	1
Caudate (body)	R		Rew > Pun = Nrew	4	19	-13	23
Putamen	R		Rew = Pun = Nrew	4	34	-5	-3
Sup. Temporal	L	38	Rew = Pun > Nrew	21	-52	14	-9
Sup. Temporal	R	38	Pun > Rew = Nrew	12	44	14	-20
Hippocampus	L		Pun > Rew > Nrew	4	-31	-25	-7
Paracentral lobule	R	6	Rew = Pun > Nrew	10	3	-36	62
Inf. Parietal	L	40	Rew = Pun > Nrew	5	-41	-46	50
Postcentral	B	4	Rew = Pun > Nrew	14	58	-8	28
SupraMarginal	R	40	Pun = Neu > Rew	52	55	-37	32
Precuneus	R	30	Rew > Pun > Nrew	38	13	-48	16
Fusiform	B	37	Rew > Pun > Nrew	80	-44	-54	-17

The OFC areas showing the incentive effect in this study showed substantial overlap with the OFC regions of interest identified in the same analysis of Experiment 1. We further verified that the effects were entirely equivalent by mapping the set of left OFC voxels which showed increased activity to No-reward in Experiment 1 onto the group of subjects of the current study. We then computed the average signal time course and conducted an Incentive x Scan ANOVA. This produced qualitatively similar results to the analysis of the ROI identified by the exploratory analysis of Experiment 2.

In this experiment, the activity in the more dorsal areas of the PFC followed a similar pattern to that found in Experiment 1: higher peak during preparation for Reward trials relative to No-reward [scan 5: $t(16) = 2.80$, $p = 0.01$]. However, in this case the Penalty trials elicited similar activation to that of Reward trials (scan 5: $t(16) = 1.27$, $p = 0.22$), but higher relative to No-reward trials [scan 5: $t(16) = 2.26$, $p = 0.04$]. Among other structures which showed effects of incentives in the first experiment, the medial prefrontal cortex was also robustly activated in this experiment (see Figure 13), and had the highest activity during Reward trials [scan 5, Reward > Punishment: $t(16) = 2.1$, $p = 0.05$]. In contrast, the medio-ventral striatum and the hippocampus were not identified in Experiment 1 at the chosen statistical threshold, but showed interactions of incentives and time in the preparation phase of this experiment [$F(14, 224) = 2.77$, $p < 0.001$, and $F(14, 224) = 2.38$, $p = 0.004$, respectively]. In planned t tests for the ventral striatum ROI, though the activity peaked highest during preparation for Reward trials, the Reward – Punishment t test of the peak difference was not significant ($p = 0.21$), whereas both Reward and Punishment peaked higher than No-reward (both p values smaller than 0.01). Similarly, activity in the hippocampus was higher during both Reward and Penalty than during No-reward (both $p < 0.02$), but was not significantly different between Reward and Penalty ($p > 0.13$).

3.4. Discussion

3.4.1. OFC activity is elicited when active control coincides with increased error rate

The analysis of task effects in this experiment revealed increased activity in control-related areas of the lateral and dorsal PFC during preparation for the Non-match relative to preparation of the

more automatic Match task (see Figure 11). These results are in agreement not only with our own results from the previous experiment, but also with a rich literature emphasizing the role of the lateral PFC in various aspects of active top-down control of behavior (Miller & Cohen, 2001). For example, (MacDonald et al., 2000) presented evidence for a specific involvement of the left lateral PFC during preparation for trials in which subjects anticipated the need to suppress an automatic, prepotent tendency in order to respond correctly, i.e. preparation for color-naming Stroop trials relative to the more automatic word-reading trials. In the left lateral PFC, the peak of the Task by Scan interaction was located relatively more posterior and more dorsal in Experiment 2 than in Experiment 1. One possible explanation for these differences lies in the slightly modified requirements of the Non-match/Match manipulation of Experiment 2. In this version, the increased pressure for speeded responses in the presence of slightly more complex task rules (i.e. the additional incentive level) might have led subjects to engage to a larger extent processes such as online maintenance of stimulus features (in preparation for fast recognition of the sample stimulus), and/or general preparation for inhibition of prepotent tendencies. Both type of processes have been associated with activation of the more typical areas of the DLPFC such as BA46 and BA9 (Barch et al., 1997; D'Esposito, Postle, Jonides, & Smith, 1999; Smith & Jonides, 1999); (Cohen & Servan-Schreiber, 1992; MacDonald et al., 2000). Interestingly, enhancement of control processes in response to monetary incentives were also noted in close vicinity to areas activated by the task manipulation (see 3.3.3 and Figures 12 and 13). This result was consistent with that of higher activation during Reward trials of Experiment 1, as well as with previous findings in tasks which involved monetary incentives in humans (Pochon et al., 2002; Ramnani & Miall, 2003; Taylor et al., 2004), or primary rewards in non-human primates (Wallis & Miller, 2003; Watanabe et al., 2002).

In contrast to the analysis of task effects in Experiment 1, this time the lateral OFC did show evidence for effects of task. As can be seen in Figure 11, in the left OFC regions preparation for Non-match peaked higher than preparation for Match trials late in the interval. Given that in this paradigm the Non-match trials resulted in erroneous performance more frequently than Match trials, this finding is consistent with our hypothesis that the OFC activity during tasks with increased demands for control (i.e. Non-match relative to Match, No-go vs. Go etc) is a correlate of variations in accuracy of responding, which in turn change the expectation for rewards and punishments of stimuli associated with different levels of control.

As we pointed out before, the involvement of OFC in motivational processes through outcome representation had been suggested by single unit and lesion studies in monkeys, by studies of patients with OFC lesions, as well as by human neuroimaging data. However, previous theories all had aspects of control processing which couldn't be ruled out. For instance, both Rolls (Rolls, 1998, 2000) and Elliott (Elliott, Dolan et al., 2000) have proposed that the OFC is critical when subjects have to reverse associations between stimuli or responses and rewards. However, neither author had clearly specified how the tasks used as evidence for those theories controlled for differential engagement of control processes in response to reversals in contingencies. This confound was even more obvious in view of the work of Roberts and Robbins, who have interpreted perseverative errors, seen in OFC-lesioned marmosets after reversals of reward contingencies, as evidence for the OFC being critical for top-down suppression of "affective associations" between stimuli and rewards. As was mentioned in the introduction to this chapter, postulating an interaction between OFC (the site of maintenance of reward/punishment representation, perhaps in interaction with other limbic structures) and

DLPFC (the source of top-down bias according to task-specific context) can account for all these findings without the need for the OFC to be actively involved in control of behavior.

3.4.2. The specificity of lateral OFC activity to negative outcomes

In this experiment, there were no significant differences in reaction time or accuracy of responding between the Penalty and the Reward conditions; hence, the increased activity during preparation for Penalty trials cannot be accounted simply by better preparation (with or without associated arousal effects). This apparent specificity of OFC activity for negative outcomes is in line with previous suggestions of a medio-lateral regional specialization of the human OFC, with lateral areas being involved in processing of punishments and medial ones more sensitive to rewards (O'Doherty et al., 2003; Rolls, 2004; Small et al., 2001). As was detailed in the discussion of Experiment 1, this type of function could also account for the specific deficit observed in monkeys with lesions of the OFC (Dias & Aggleton, 2000; Dias et al., 1996a, 1996b, 1997; Roberts & Wallis, 2000).

This experiment also generated results which allowed us to examine whether patterns of OFC activity might provide support for predictions generated by prospect theory (Kahneman & Tversky, 1988), or by decision affect theory (Mellers et al., 1997). The former postulates that the psychological value of an outcome (formerly called “utility” in classical economic theories) is assessed in terms of a gain or a loss relative to the status-quo. In contrast, decision affect theory allows multiple reference points to influence the value of an outcome, and postulates that outcomes are evaluated in terms of the anticipated pleasure and/or disappointment, taking into consideration *alternative* possible outcomes (Mellers et al., 1999; Mellers et al., 1997). In the current variant of the incentive manipulation, the utility of each reward trial (which is

proportional to the magnitude of the reward and the probability of occurrence) should have increased relative to Experiment 1, since the magnitude of each reward increases by 100%, while probability of obtaining the reward (due errors or late responses) decreases from approximately 90% to 70%. In other words, since in both experiments missing a reward does not change the status quo, the expected utility of Reward trials is basically reduced to the product between the magnitude of reward and its probability (approximately 0.45 in Experiment 1, and approximately 0.7 in the second). Therefore, if the OFC activity represented a signal inversely proportional to the expected utility (as suggested by the increase to No-reward cues observed in Experiment 1), the Reward trials in the current experiment should have elicited more of a decrease relative to Experiment 1. In contrast, the OFC representations might be more closely related to the anticipated disappointment (Mellers, 2001). As we mentioned earlier, this framework allows for multiple reference points to influence the expected impact of outcomes. In this experiment, the presence of a Penalty condition and a more demanding response deadline, translated in a bigger impact of any given Reward on the overall gains. Therefore, if OFC activity was indeed related to expected disappointment, it should increase from Experiment 1 to Experiment 2. Similarly, the expected utility of No-reward trials should not change. In terms of expected disappointment, the predictions for No-reward trials are not so clear. On one hand, the presence of Penalty trials and the overall increase in frequency of negative outcomes (i.e. errors), might tend to turn the No-reward trials into a better outcome (relative to those of Experiment 1). Conversely, the increased number of missed rewards (due to the adaptive deadline) would tend to emphasize their "missed opportunity for reward" aspect. Consequently, no change in OFC activity of No-reward trials from Experiments 1 to Experiment 2 would be consistent with either the prospect theory or the subjective anticipated pleasure frameworks, whereas changes in the No-reward activity would be

more consistent with the latter hypothesis. The patterns of activations in Experiments 1 and 2 during Reward trials were indeed consistent with the anticipated pleasure/disappointment framework and not with the expected utility theory: activity to Reward cues was higher, early in the sample-target interval, in Experiment 2 compared to the first experiment. In the peak voxel timeseries of both left and right OFC regions of interest, the No-reward cues of Experiment 2 lead to late rise in activity similar to that observed in Experiment 1. However, in the right OFC area, an initial decrease in activity was also observed (see Figure 13); this initial decrease in activity to No-reward was also present, bilaterally, at the inspection of the averaged signal changes. As we explained before, this initial decrease in No-reward related activity would also be consistent with predictions of decision affect theory. Obviously, more research is necessary to precisely characterize these effects during reward/punishment anticipation and all the factors that participate in these evaluative processes. However, it is interesting to note that recent results in patients with OFC lesions suggest that the human OFC is involved in the "counterfactual effects" of feelings at the time of outcome presentation. Specifically, the OFC seems to be critical for experiencing regret in the context of alternative outcomes (Camille et al., 2004). This effect was noted by measuring the reported feelings and the skin conductance changes of subjects who experienced monetary gains from gambles in which they were either given full feedback (i.e. the amount that "could have been won") or not. Healthy subjects reported experiencing feeling of regret when they were aware of an alternative better outcome than the one obtained, whereas subjects with OFC lesions did not endorse such feelings. The changes in skin conductance at the time of outcome presentation were consistent with the self-report of each group. In our data, the changes between Experiments 1 and 2 in the activity elicited by Reward and No-reward cues suggest that this type of counterfactual effect is also manifested in the OFC at the time of

outcome anticipation. Overall, this type of results also fit with existing hypotheses which propose an involvement of the OFC in the use of reward and punishment feedback in service of decision making (Bechara et al., 1994; Bechara, Damasio, Damasio, & Lee, 1999; Bechara, Damasio, Tranel, & Anderson, 1998; Bechara et al., 1997; Bechara, Tranel, Damasio, & Damasio, 1996; Damasio, 1996). However, these theories postulate that the role of the OFC in such processes to mediate the triggering of bodily reactions in response to stimuli which previously led to emotional reactions, and that decisions are made based on the nature of these "somatic" signals. Despite the face validity of such a theory given the well documented correlation between emotions and autonomic reactivity, no direct evidence has been provided yet, in either animal models or humans, that such bottom-up signals are indeed necessary for decision-making (Rolls, 2000).

3.4.3. Orbitofrontal activity and anxiety

The specificity of OFC activation to Penalty cues is in line with multiple results which have pointed towards this cortical area as a likely substrate of anxiety (Chua, Krams, Toni, Passingham, & Dolan, 1999; Crestani et al., 1999; Grachev & Apkarian, 2000; Grafman, Vance, Weingartner, Salazar, & Amin, 1986; Malizia, 1999; Malizia et al., 1998; Rauch et al., 1997; Rauch et al., 1995; Swedo et al., 1992; Tashiro et al., 2001; Vasa et al., 2004). However, upon detailed examination of the actual pattern of activations and deficits associated with OFC function, it quickly becomes apparent that this literature is rather mixed with respect to the precise OFC involvement. As such, studies of human patients have suggested that OFC lesions can lead to either increased (Grafman et al., 1986) or decreased anxiety (Vasa et al., 2004). In imaging studies, various indices of OFC activity have been associated both with increased

anxiety in healthy subjects (Chua et al., 1999; Grachev & Apkarian, 2000; Liotti et al., 2000) and with various types of pathological anxiety (Mataix-Cols et al., 2004; Rauch et al., 1997; Rauch et al., 1995; Swedo et al., 1989). In animal studies, sensitivity to threat stimuli has been related to OFC function (Crestani et al., 1999). On the other hand, some studies reported negative correlations of OFC activity with measures of anxiety, either in anxiogenic situations or in anxious patients who improved after treatment (Buchsbaum et al., 2001; Fredrikson, Wik, Annas, Ericson, & Stone-Elander, 1995; London et al., 2004). One possible explanation for the heterogeneity of these results is that the different measures of activity have used different types of contrasts, such as: rest vs. activity, emotionally-neutral vs. emotionally-valenced task, or anxious subjects vs. healthy controls in the same anxiety-provoking situation. In order to examine the relationship between anxiety and PFC function in our dataset, we performed an analysis of incentive effects after subdividing the subjects according to self-reported levels of trait anxiety (rated at the end of the scanning session using the Spielberger State-Trait Anxiety Inventory – Trait, STAI-T). The 17 subjects were divided into terciles based on the STAI-T scores, and the lowest and highest of the terciles were selected for analysis. Though this analysis was necessarily underpowered, the results are nevertheless interesting and worth noting (see Figure 14).

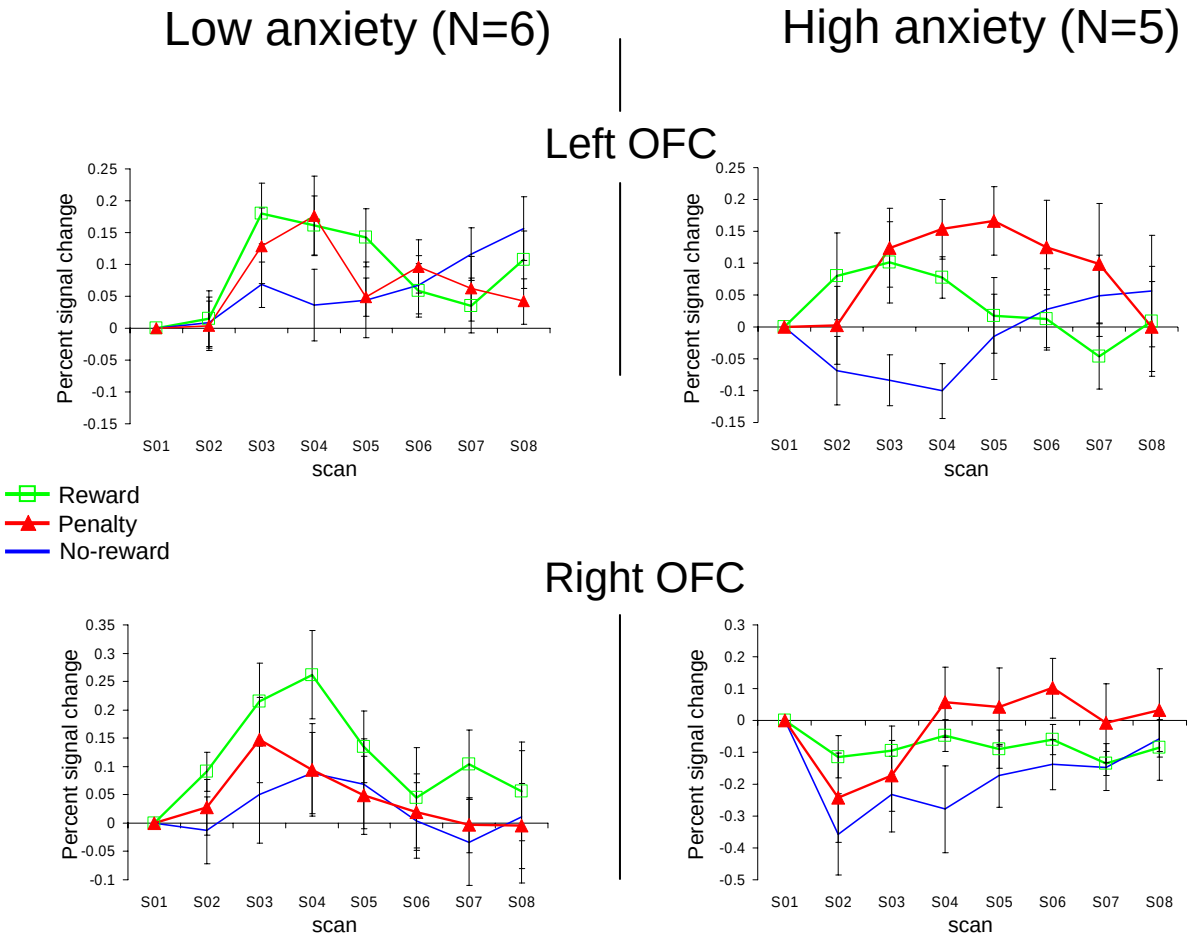

Figure 14. Effects of high vs. low trait anxiety on the lateral OFC incentive-related activity

Graphs of activity to Reward, Penalty and No-reward cues in two subgroups of subjects: the lowest scoring (left column graphs, N=6) and the highest scoring (right column graphs, N=5) terciles. In both left (top graphs) and right OFC (bottom graphs), high trait anxiety seems to result in lower modulation of activity by incentive type, and in more sustained activity to Penalty cues.

In both left and right OFC, the maximal effect of preparation to avoid penalties was more evident in high trait anxiety than in low anxiety subjects. It would be premature to speculate at any length on the precise condition differences which generated this effect, since the Group (high vs. low anxiety) by Incentive by Scan ANOVA did not reach significance in either region of interest (both p values greater than 0.2). Furthermore, it is possible that the focus on negative consequences of actions (characteristic of anxiety) increased the tonic neural activity in these

areas and decreased the degree to which activity could be further modulated by trial-to-trial changes in type and probability of outcomes. Despite the preliminary nature of these results, they nevertheless suggest that this paradigm might be appropriate for the study of the neural substrates of anxiety, and justify future studies in larger subject samples.

CHAPTER 4

The OFC hyperactivity in OCD

4.1.Introduction and rationale

It has been repeatedly proposed that obsessive-compulsive disorder (OCD) is associated with dysfunction in processes subserved by the fronto-striatal-thalamic-cortical loops (Rapoport, 1991; Rauch, 2000; Saxena et al., 1998). These pathogenetic models of the disorder often emphasize the critical position of the OFC in these circuits, and point out the consistency between this theoretical framework and findings of hyperactivity in the OFC of OCD patients at rest (Alptekin et al., 2001; Kwon et al., 2003; Saxena et al., 2003; Saxena et al., 1999; Swedo et al., 1989) and during symptom provocation (Rauch et al., 1994; Rauch et al., 2002; Saxena et al., 1999). Some results also suggest that OFC hyperactivity decreases with successful treatment, or that successful treatment is predicted by lower OFC activity levels (Brody et al., 1998; Rauch et al., 2002; Saxena et al., 2003; Saxena et al., 1999; Swedo et al., 1992). In parallel with neuroimaging findings, the phenomenological similarity between obsessive-compulsive behaviors and deficits in suppression of inappropriate behaviors has been frequently pointed out. Consequently, neuropsychological studies have focused on this type of processes and reported impaired performance of OCD patients compared to healthy controls on various tasks thought to tap into the capacity to inhibit inappropriate or prepotent response tendencies. Such deficits include: lack of suppression of visual saccades (Rosenberg, Dick, O'Hearn, & Sweeney, 1997),

decreased semantic negative priming (Enright & Beech, 1993) and increased Stroop interference (Hartston & Swerdlow, 1999). Though other findings suggest that these impairments reflect deficits in guiding behavior based on internal representations rather than poor inhibitory control (Maruff, Purcell, Tyler, Pantelis, & Currie, 1999), this behavioral deficit of OCD patients has continued to receive considerable attention. Furthermore, the link between OCD and abnormal cognitive control has been also fueled by the long standing view that the OFC is critical for inhibitory control. In parallel with the findings suggesting OFC abnormalities in OCD, many studies have also reported basal ganglia hyperactivity (Brody et al., 1998; Horwitz et al., 1991; Perani et al., 1995; Rauch et al., 1994; Rauch et al., 2001; Saxena et al., 2002; Saxena et al., 1999), along with functional and structural abnormalities of other prefrontal regions such as the anterior cingulate cortex (Brody et al., 1998; Perani et al., 1995; Swedo et al., 1989; Szeszko et al., 2004) and of the thalamus (Fitzgerald, Moore, Paulson, Stewart, & Rosenberg, 2000; Gilbert et al., 2000; Perani et al., 1995; Rauch et al., 1994; Rosenberg, Benazon, Gilbert, Sullivan, & Moore, 2000). In an effort to integrate all these findings in a comprehensive model of OCD pathogenesis, it has been hypothesized (Saxena et al., 1998) that at the origin of OCD symptomatology is an imbalance between the direct and the indirect pathways in the fronto-striato-thalamic loops (Alexander, DeLong, & Strick, 1986; Cummings, 1995; Nakano, 2000). In this model (Saxena et al., 1998), a relative “imbalance” between the activity through the direct and indirect pathways would result in pathological reentrant activity in the fronto-striato-thalamic loops involving areas of the prefrontal cortex such the OFC. This framework is closely related to another major view which places emphasis on basal ganglia dysfunction in OCD, and proposes that these structures are the repository of automatic, stereotyped behaviors, which become maladaptive because of basal ganglia dysfunction. Despite their immediate appeal, these

theories have several fundamental weaknesses. One is the lack of direct evidence for either of them, since most of the evidence used in formulating them has been generated in experiments at rest or during symptom provocation, and therefore could not clearly differentiate between causal mechanisms, epiphenomena and compensatory activity. Even the few studies which compared the brain activity of OCD patients and controls in relationship to cognitive performance had limited interpretability either because of the use of a block design (Nordahl et al., 1989), which was confounded by processes associated with incorrect performance, or because indices of brain activity were measured separately from cognitive performance (Kwon et al., 2003). Furthermore, on closer examination, the logic behind linking the imaging findings with neuropsychological studies and phenomenology of OCD is not clear: most of the structures referred to in these theories have been found hyperactive in imaging studies of OCD, but OC symptoms are caused by lesions of the same regions (Berthier et al., 1996), implying that either the imaging correlates of hypofunction can appear either as hypo or hypermetabolism, or that these structures are actually implicated in some sort of compensatory activity brought about by the symptomatic state.

The above considerations formed the motivation for the work described in this chapter. Specifically, it appears that a necessary step along the path for understanding OCD is characterizing the previously identified brain dysfunctions in terms of their specificity for OCD, and in terms of their relevance as a phenotype underlying vulnerability for symptomatology. With respect to OFC dysfunction, while being frequently referred to as having high relevance for OCD pathogenesis, some results suggest that it is associated with anxiety symptoms in several disorders, including OCD, simple phobia and PTSD (Rauch et al., 1997). Furthermore, as is the

case with many findings of brain dysfunction in this disorder, it is still unclear to what extent they represent an underlying vulnerability factor as opposed to an epiphenomenon.

The main goal of the experiment described in this chapter was to provide evidence that the OFC hyperactivity in OCD represents a fundamental underlying neural dysfunction and is not brought about only by factors which induce OCD-specific symptoms. A secondary goal was to verify predictions generated by the overall framework regarding OFC involvement in executive function, elaborated in Chapters 2 and 3. Specifically, if the lateral OFC is involved in representing possible negative outcomes and this function is exaggerated in OCD patients, then their OFC activity should be increased in conditions that deviate from the perceived task default, and/or have potential for negative outcomes.

To examine the hypothesis that OFC hyperactivity is not a manifestation specific to the symptomatic state, we conducted an event-related fMRI experiment in which OCD patients and group-matched controls were performing a disease-neutral cognitive task (Ursu et al., 2003). If the OFC of patients is part of a neural circuit which is fundamentally disturbed in OCD, as suggested by the neuropsychological literature, one would expect to find evidence of hyperactivity despite the fact that patients have to focus on stimuli that have no relevance for obsessive-compulsive phenomenology. Furthermore, given the proposed function of the OFC in representation of potential outcomes, the anticipatory OFC hyperactivity in OCD should be manifested in response to cues signaling high likelihood of relatively negative outcomes (in this case, signaling the occurrence of infrequent probes which violate the dominant rule defining a target).

4.2.Methods

4.2.1. Subjects

Fifteen adult healthy volunteers (8 males) and 15 adult patients (7 males) with OCD participated in the experiment. Patients were diagnosed according to DSM-IV criteria for OCD and were excluded if criteria were met for major depressive episode within the last month, lifetime psychotic symptoms, substance abuse within the previous 6 months, history of major neurological disorders or head trauma, or the presence of first-degree relatives with psychotic disorders. The comparison subjects were screened for history of head trauma, neurological or psychiatric disorders, substance abuse in the past 6 months, current psychoactive medication and history of psychotic disorders in first-degree relatives. Informed consent was obtained from all subjects, who were paid for participation. All procedures were approved by the Institutional Review Board of the University of Pittsburgh. Neither demographics nor handedness, evaluated with the Edinburgh inventory (Oldfield, 1971) were significantly different between groups (see Table 7).

Seven of the patients had a history of major depressive disorder, two of post-traumatic stress disorder (PTSD), one of generalized anxiety disorder, one of panic disorder, and one of anorexia. At the time of study, 6 patients met criteria for generalized anxiety disorder, one for PTSD and one for social phobia. All but two patients were medicated at the time of the study: 4 were taking sertraline, 2 fluvoxamine, 3 fluoxetine, 2 clomipramine, 1 paroxetine, 1 citalopram, 1 bupropion, and 1 venlafaxine. Six of the patients were also treated with clonazepam (2 patients), lorazepam (1 patient), gabapentin (1 patient) and valproate (2 patient). Immediately after the scanning session, 14 of the 15 patients were evaluated using the state version of the Spielberger State-Trait Anxiety Inventory (STAI-S) (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1980) and the Yale-

Brown Obsessive-Compulsive Scale (Y-BOCS, (Goodman, Price, Rasmussen, Mazure, Delgado et al., 1989; Goodman, Price, Rasmussen, Mazure, Fleischmann et al., 1989), see Table 7). All psychiatric

Measure	Group	
	Obsessive-compulsive disorder (n=15)	Controls (n=15)
Number of males, females	7, 8	8, 5
Age	32.06 (8.06 22-45)	30.85 (7.96 18-45)
Handedness (right, left)	13, 2	13, 2
Education (years)	15.8 (2.46 12-20)	16.56 (1.93 14-20)
YBOCS total	20.67 (5.05 9-28)	---
YBOCS (obsessions)	10.46 (2.94 4-14)	---
YBOCS (compulsions)	10.0 (3 4-14)	---
STAI-S [a]	40.0 [b] (9.4 22-62)	---

interviews and clinical assessments were conducted by the same rater (S.U.).

Table 7. Demographics and clinical measures of the patient and control groups

Group means are reported, with standard deviation (SD) and range in parentheses. Demographic measures, evaluated with t tests (for mean age) and χ^2 tests (for gender composition) were not different between groups (all p values > 0.4).
YBOCS = Yale-Brown Obsessive-Compulsive Scale, STAI-S = Spielberger Trait-State Anxiety Inventory – State.
[a]one patient was not scored on the STAI-S inventory. [b]score mean was within one SD of normative scores for the general population, with 2 individual scores falling outside 1SD of the normative scores.

4.2.2. Behavioral task and testing procedures

Subjects were scanned while performing the AX-CPT, a modified Continuous Performance Test. Every 12 seconds, single letters were presented centrally for 0.5 seconds. Subjects were instructed to press a “target” button (i.e. using the index finger of the dominant hand) whenever

the letter was an X which had been preceded by an A, and a "non-target" button (middle finger of the dominant hand) after all other stimuli, including X letters preceded by other letters than an A. Starting with the first stimulus of a block, stimuli could therefore be parsed in pairs of sequential stimuli (cue and probe) which constituted a trial, each followed by a 11.5 seconds inter-trial interval. This design ensured a constant 12 second inter-stimulus interval. Within each block of 10 trials, 70% were "AX" sequences, while A-nonX ("AY"), nonA-X ("BX") and nonA-nonX ("BY") trials occurred each with 10% frequency.

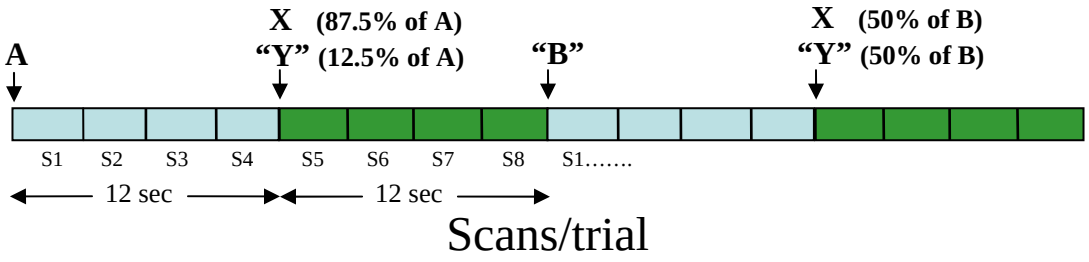

Figure 15. Task design used to compare OCD patients and controls (Experiment 3).

A trial consisted of the presentation of a cue letter (A or "B") and a probe (X or "Y"). The letter A was a cue in 80% of the trials, and one of four other letters (i.e. "B") was used in the rest of the trials (20%). Of the A cue trials, 87.5% (7 out of 8) were followed by target Xs, and 12.5% (1 out of 8) by different non-targets ("Y"s). During the "B" cue trials, half the times the cue was followed by a non-target X (i.e. the letter X was a probe, but it required a rare non-target response), and the other half were non-X (also non-targets). The letter A was never used as a probe and the Xs were never used as cues.

Subjects were instructed to respond as quickly and as accurately as possible. Half of the twelve blocks of trials administered contained degraded stimuli (an individually titrated percentage of the pixels were randomly removed), and the order of the standard vs. degraded blocks was counterbalanced across subjects and groups. This manipulation had the role of increasing the number of errors for a previously published analysis (Ursu et al., 2003). However, only data from the six blocks of regular, non-degraded stimuli were used in the current analyses.

Before scanning, subjects were trained on the task outside the scanner until they performed with approximately 80-90% accuracy. Patients naïve to MRI scanning procedures were exposed to a simulator within one week prior to scanning to minimize state anxiety related to the scanning procedure. Stimuli were presented on a Macintosh computer using PsyScope software. They were made visible to the subjects by projecting them on a translucent screen placed in the scanner approximately 50cm from the center of the bore, and viewed via an MR-safe prismatic mirror installed on the scanner head coil, directly above the subjects' eyes. Responses during the scanning session were collected using specialized fiberoptic response device attached to the subjects' dominant hand.

4.2.3. fMRI data acquisition and analysis

Images were acquired with a 1.5-T GE Signa scanner with a standard head coil. Structural images were obtained with a standard T1-weighted pulse sequence. Functional volumes of twenty adjacent coronal slices (in plane resolution: 3.75 mm, 3.8mm thick), were acquired every 3 seconds using a two-shot T2*-weighted spiral scanning sequence (Noll, Cohen, Meyer, & Schneider, 1995), with the following parameters: repetition time (TR) 1500ms; echo time (TE) 35ms; flip angle 50°; field of view 24cm. Functional data were movement corrected (Woods et al., 1992). The estimated mean movement was smaller than 1mm for the six movement parameters in both groups, and there was no main effect of group or Group x Parameter interaction in a multivariate analysis, all $p<0.05$. Then data were linearly detrended on a voxel-by-voxel basis, and then co-registered to a common reference structural scan using a 12-parameter automatic algorithm (Woods et al., 1992). Images were then smoothed with an 8-mm FWHM three-dimensional Gaussian filter and pooled across participants to increase signal-to-noise and accommodate for residual differences in individual anatomy.

Eight stimulus-locked scans were acquired during each 24s between subsequent cues. Event-related analyses of the blood-oxygenation-level dependent (BOLD) responses after cues used a voxel-wise ANOVA model with Subject as random factor, Group (patients vs. controls) as between-group factor, and scan (1 through 4) as one of the repeated measures factors (the second one was cue type). The dependent variable was the signal value in each of the four scans that followed probe presentations (Carter et al., 1998; Ursu et al., 2003). Region of interest (ROI) were defined using the same procedure as described in section 2.2.3, with the only difference that the correction for the number of slices which included prefrontal cortex was adapted to the slice placement and dimensions of the current acquisition. In each ROI, where necessary, direction of effects was confirmed using the same criteria as those described in section 2.2.3.

4.3.Results

4.3.1. Behavioral results

The behavioral performance of the two groups was contrasted by conducting random effects ANOVAs of mean reaction times and accuracy rates. The accuracy of responding to the cues (A vs. B) showed only a main effect of cue [$F(1, 28) = 5.60$, $p < 0.03$], reflecting that neither patients or controls made any errors to B cues, and made a very small but statistically significant number of errors to A cues (approximately 1%) There was no interaction with group ($p < 0.64$). Therefore, in all subsequent behavioral and imaging analyses, we only included trials with correct responses to the cue. The reaction time ANOVA revealed a significant main effect of cue and group, with responses to B cues being slower in both groups (main effect of cue: $F(1, 28) = 6.46$, $p < 0.02$) and patients slower than controls (main effect of group: $F(1, 28) = 5.43$, $p < 0.03$) but no interaction of Group and Cue ($p > 0.5$).

We also examined the effects of cue and probe letter on responses to probes, both in terms of reaction time and accuracy. The results of Group x Probe type (AX, BX, AY, BY) ANOVAS of mean reaction time and accuracy rates are summarized in Figure 16. In summary, patients were overall slower than controls, but not differentially slower for any of the probe types (Group by Probe interaction: $F(3, 84) = 1.13$, $p > 0.34$). Responding to both non-target Xs (BX probes) and to non-X probes that followed the frequent A cues (AY probes) was slower in both groups. The difference in accuracy of responding to different types of probes was not modulated by group (Group by Probe interaction: $F(3, 84) = 1.34$, $p > 0.25$). Finally, the effect of Cue on reaction time and accuracy of responses to probes (collapsed across probe type) was examined in two Group by Cue ANOVAS. Similar to results of the other behavioral analyses, the interaction of Group by Cue type was not statistically significant in either reaction time to probes [$F(1, 28) = 0.94$, $p > 0.34$] or accuracy to probes [$F(1, 28) = 2.52$, $p > 0.12$].

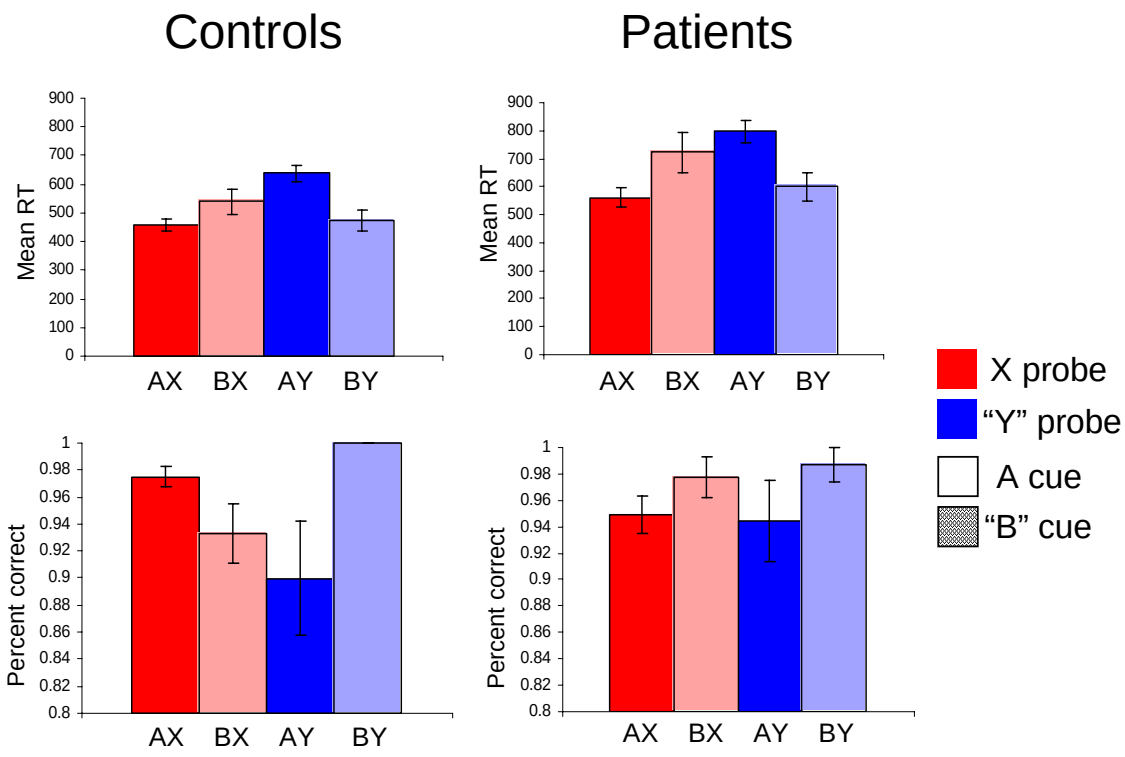

Figure 16. Reaction times and accuracy rates of responding to probes in the AX-CPT task

Both patients and controls responded slower to BX and AY probes, and patients were overall slower than controls. The patient group's accuracy of responses was overall similar to that of controls. Despite apparent differences in accuracy rates across cue type and probe type, these differences did not interact with group.

4.3.2. Effects of target vs. non-target predictors on brain activity

An exploratory Group (patients vs. controls) by Cue (A vs. B) by Scan (S1 through S4) of the fMRI data revealed a region with significant interaction in the right lateral OFC, with the peak voxel (i.e. the voxel of maximum F value) at Talairach coordinates 36, 32, -10 (see Figure 17). In this ROI, the pattern of signal change suggested hyperactivity in patients relative to controls after B cues. Indeed, planned contrasts in the peak voxel showed that the difference between activity to B and A cues was significant in patients (scan 3: $t(14) = 2.69$, $p = 0.02$), whereas in controls it was not (scan 3: $t(14) = 0.36$, $p > 0.7$). The increased activity of patients was also manifested as a

trend for an increase in peak signal change to B cues when compared directly to the peak B signal change of controls (scan 3: $t(13) = 1.55$, $p = 0.13$).

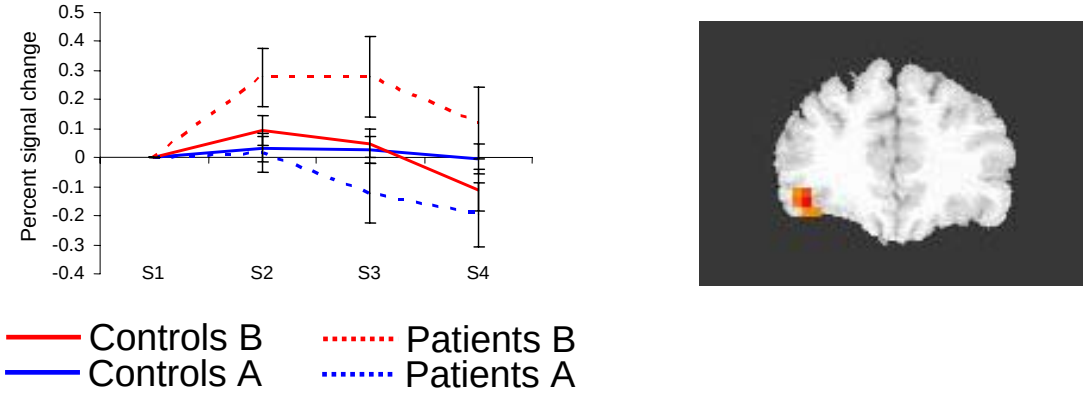

Figure 17. Lateral orbitofrontal region with significant Group by Cue by Scan interaction

The patient group showed increased activity to B cues relative to A cues, whereas controls did not. Comparing directly the peak signal change to B cues (at scan 3) of controls and patients showed a trend for increased activity in patients.

We also tested the correlations between the peak signal changes elicited by B cues in patients and the individual clinical ratings, obtained immediately after the scanning session. In order to decrease the influence of outliers in the estimate of average signal change, we computed the mean signal change across patients in the scans covering the cue-probe interval, and eliminated all scans with signal change more than 2 standard deviations above and below the mean. Then we computed, for each patient, the average signal change at scan 3 in the peak voxel of the OFC region of interest. Those average values were correlated with the symptoms scores in 14 patients who had both STAI-S and YBOCS scores. STAI-S correlated with B cue activity on both parametric (Spearman's r) and non-parametric (Kendall's tau) tests ($p = 0.02$ and $p = 0.003$ respectively). In contrast, the YBOCS scores did not correlate with the peak of B cue activity ($p > 0.3$ for both tests). Examination of the estimates of mean signal change revealed that 6 of the

14 patients had negative values. Therefore, we also conducted a non-parametric test of the correlation between state anxiety scores and fMRI activity excluding the subjects with negative mean signal changes. In this analysis, the non-parametric test of the correlation between STAI-S scores and fMRI activity showed only a trend for significance (Kendall's tau test, $p = 0.08$). Finally, we explored the OFC region of interest in terms of its activity during the post-response interval (normalized to the first scan of the interval, S5). A Group by Probe (AX, AY, BX, BY) by Scan (S5-S8) ANOVA did not yield a significant Group by Probe by Scan interaction [$F(9, 252) = 1.15$, $p > 0.32$].

In the exploratory analysis of the cue-related activity, other areas with significant Group by Cue by Scan interaction included pregenual and dorsal anterior cingulate cortex (BA 32 and 24), left lateral PFC (BA 9/8), bilateral superior and left inferior temporal gyrus.

4.4.Discussion

The results presented in this chapter confirmed that OFC hyperactivity in OCD patients relative to controls is manifested independently of OCD-specific stimuli. The location of the area of hyperactivity was similar to that previously reported in resting or symptom provocation studies. However, in this experiment, the stimuli to which patients had to respond were letters, which were not the focus of their obsessive-compulsive concerns (confirmed upon debriefing at the end of the scanning session). Consequently, this provides more direct evidence that this hyperactivity reflects an underlying functional deficit linked to the vulnerability for symptoms, rather than a simple correlate of the symptomatic state. This has been suggested by the few studies which compared both the brain activity and the cognitive performance of OCD patients with that of healthy controls (Kwon et al., 2003; Nordahl et al., 1989). However, the use of an event-related

design in this study provided interpretive advantage over those studies which involved the performance of a simple auditory discrimination task in a block design (Nordahl et al., 1989) or simply correlated PET data with performance on neuropsychological and cognitive tasks administered outside the scanner (Kwon et al., 2003). As such, our results rule out the possible role of exaggerated reactions of patients to erroneous trials, since all analyses were conducted on trials in which performance was correct. With regards to the possible symptom correlate of the OFC hyperactivity, our results have to be interpreted in the context of previous research. Several studies have noted correlations between OFC activity and obsessive-compulsive symptoms….. ,decreases of resting OFC metabolism after successful therapy (Saxena et al., 2002), and an inverse relationship between pre-treatment OFC activity and response to medication, i.e. better response in subjects with lower pre-treatment OFC activity (Rauch et al., 2002). However, other studies provided evidence that the OFC hyperactivity of patients is also related to the severity of anxiety symptoms (Swedo et al., 1992; Swedo et al., 1989). In these data, the relationship between the OFC and exaggerated concerns with potential negative outcomes of patients was supported by the specific patient hyperactivity to non-target predictors (B cues). This interpretation is consistent with the results presented in Chapters 2 and 3, which strongly support a role for this OFC area in representation of negative outcomes. In the AX-CPT task, subjects were instructed that the letter X will be a target when preceded by an A, but not if preceded by any other letter. Therefore, the infrequent “B” cues (non-A) predicted with 50% likelihood the occurrence of a non-target X, which in the context of the high frequency of target responses to X letters (70% of trials were AX sequences) signaled the potential of an erroneous response, relative to A cues, which were certain predictors of target Xs. Unfortunately, a direct test of the possibility that the B cue hyperactivity is related to anxiogenic mechanisms was not possible

because in this study patients were not evaluated in terms of trait anxiety. While the correlation between state anxiety and B cue-related activity of patients indirectly supports this view, several considerations prevent any strong claims regarding this possibility. First, 6 of the 14 measures of fMRI signal change used in the correlation analysis were negative, and activity decreases in fMRI are hard to interpret (Gusnard & Raichle, 2001; Raichle et al., 2001; Shulman et al., 1997). Secondly, the relevance of correlations between state anxiety scores and B cue activity depends on the high correlation between state and trait anxiety (Spielberger et al., 1980). However, the reliability of this state-trait relationship has been questioned by some researchers (Ramanaiah, Franzen, & Schill, 1983). Future studies will have to explore this relationship more directly, either in populations of patients with other anxiety disorders or in groups of healthy subjects pre-selected based on trait anxiety scores, and using tasks more sensitive to representation of outcomes such as the ones used in Experiments 1 and 2 of this thesis.

Eliciting the OFC hyperactivity in OCD patients in response to a disease-neutral task provides evidence that the OFC of patients is indeed dysfunctional, and support the claim that previously observed OFC hyperactivations at rest (Alptekin et al., 2001; Baxter, Schwartz, Guze, Bergman, & Szuba, 1990; Brody et al., 1998; Busatto et al., 2000; Kwon et al., 2003; Saxena et al., 2002; Saxena et al., 1999; Swedo et al., 1992; Swedo et al., 1989) and during symptom provocation (Mataix-Cols et al., 2004; Rauch et al., 1994; Rauch et al., 2002; Saxena et al., 1999) are a indicative of an underlying neural abnormality. However, establishing precisely the type of symptoms that this dysfunction is responsible for will have to be the focus of future research.

Based on the results of the first two experiments included in this thesis, and considering the phenomenology of OCD (for which a central element are exaggerated concerns with

potential *future negative consequences* of actions), these results also provide support for the hypothesis that the OFC is involved in representing potential negative outcomes (in the case of patients performing the AX-CPT task, high likelihood of a stimulus to which the automatic "target" response was not appropriate and could lead to errors).

Together with the preliminary analyses of the effects of trait anxiety on OFC activity presented in Chapter 3, these results suggest that high levels of anxiety might interact both with tonic OFC activity and with phasic activation in response to individual stimuli. If that were true, the changes in baseline occurring as a result of tonic effects of anxiety might influence phasic activations which are the main source of differences in event-related fMRI designs (for example, some of the differences in signal change between high and low anxiety subjects might reflect effects of overlap on a different baseline). These results point out the complexities of interpreting OFC activations with respect to a putative role in normal and pathological anxiety, and suggest that careful control of the baseline used for estimation of both tonic and phasic brain activity may be necessary in future studies of neural substrates of anxiety, perhaps through the use of combined block and event-related designs (Visscher et al., 2003).

CHAPTER 5

Summary Discussion

The research presented in this thesis was aimed at elucidating the contribution of the human orbitofrontal cortex to executive functions. The projects described in the previous three chapters used functional magnetic resonance imaging to ask three as yet unanswered fundamental questions regarding the involvement of the OFC in behavioral control. In Chapter 2, we demonstrated that, counter to claims maintained by existing theories, the human OFC is not engaged by active top-down control processes, but rather in representing the *potential outcome* of actions. Chapter 3 confirmed the OFC's role in representation of outcomes, and provided convergent evidence that the lateral OFC activity is specific for potential negative outcomes. At the same time, results of that project reconciled the involvement of the OFC in representation of outcomes with previous results interpreted as evidence for a role in control, by showing that OFC is engaged under conditions in which high engagement of control processes coincides with changes in outcome (i.e. increased error rates). Finally, Chapter 4 explored the relevance of these findings for human psychopathology, in particular for obsessive-compulsive disorder. We showed that in this psychiatric population, characterized by excessive concerns with future negative consequences of actions, the OFC is hyperactive in response to stimuli which, while being unrelated to the symptoms of the disease, are predictors of possible negative outcomes.

This chapter will review the main findings of the three projects mentioned above and elaborate on the significance of these findings, with an emphasis on issues that were either raised

or unresolved by each set of results. Finally, the last section will present a systems view which will attempt to integrate the proposed OFC contribution with that of other brain structures in guiding goal directed human behavior.

5.1. The OFC represents possible outcomes but does not actively implement control

Chapters 2 and 3 presented results of two event-related fMRI experiments in which subjects prepared, in any given trial, to perform one of two tasks with different levels of demands for control, and anticipated receiving various amounts of a monetary reward for correct performance. In both experiments, the monetary incentives had the effect of engaging control processes which resulted in improved performance. Across experiments, the increased demands for control of one of the tasks were reflected in increased activity in lateral and dorsal areas of the PFC, regardless whether the increased task complexity which engaged control also resulted in degradation of performance (i.e. increased error rates). In contrast, activity in the OFC was only modulated by demands for control when the latter coincided with changes in performance outcomes (i.e. increased error rates). Furthermore, the engagement of control in response to incentives was reflected in increased activity in the same areas which activated in response to demands for control. In contrast, the OFC (in particular its lateral aspects) was robustly activated in a condition in which a change in monetary outcome was anticipated but in which subjects also showed evidence of reduced engagement of control processes. These convergent results (summarized in Figure), showed that the OFC activity paralleled variations of motivational states brought about by changes in likelihood of outcomes (i.e. incentives or performance errors). In contrast, the lateral and dorsal PFC were consistently engaged when top-down control processes were activated.

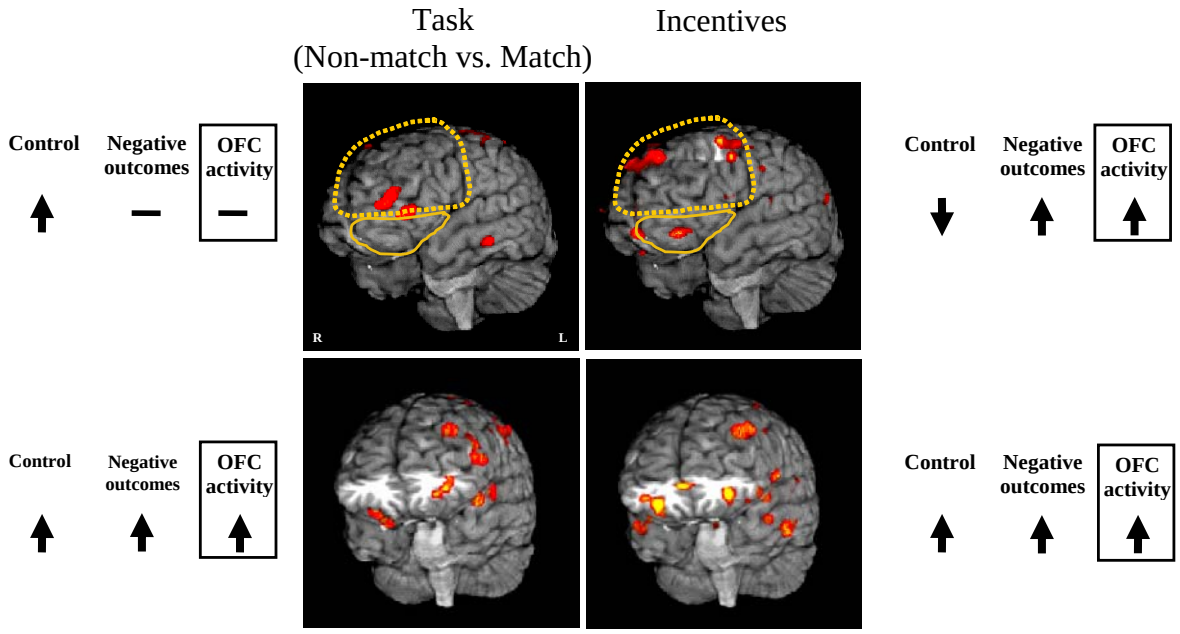

Figure 18. Summary of findings from Experiments 1 and 2.

Statistical maps generated by the interaction of Time and Task (left) or Incentives (right) in Experiment 1 (top) and Experiment 2 (bottom). The activation maps from Experiment 1 (top) have the approximate boundaries of lateral and orbital PFC highlighted. Next to each map, the predictions generated by the two hypothesis regarding the OFC role in behavioral control are contrasted with the actual signal change observed in the left lateral OFC (black box). The BOLD signal change in the left OFC always follows the changes in representation of potential negative outcomes, but not the changes in active cognitive control (top maps).

Across experimental designs, task-related activations were largely overlapping in dorsal and lateral prefrontal cortex. However, in the left lateral PFC, the area showing increased activity during preparation for Non-match was located more caudally in Experiment 2 relative to that identified by the same contrast in Experiment 1. Two main possibilities could explain this variation. The first explanation has to do with the inherent technical limitations of the fMRI methodology. It is possible that the observed variability in location of statistically significant task effects reflects limits of the test-retest reliability of the fMRI signal. Several reasons make this account unlikely. Firstly, previous studies have shown that the fMRI activations in response to working memory paradigms have acceptable test-retest reliability in group-averaged data

(Manoach et al., 2001). Furthermore, the reliability of working memory-related activations has been shown to be comparable to that of activations elicited by simple motor tasks (Noll et al., 1997), and in our two experiments the task effects in motor and premotor areas were highly reproducible (see Tables 1, 2, 4, 5). A second possibility is that the variation may be due to differences in the type of demands posed on representation of task-specific context by each paradigm. In the first experiment, the fixed response deadline may have allowed subjects to address the increased demands of the Non-match task by preparing for a "cognitive branching" strategy (i.e. identifying the sample, followed by reactivation of the task set), which has been shown to activate more anterior aspects of the lateral PFC (Koechlin, Basso, Pietrini, Panzer, & Grafman, 1999). In Experiment 2 however, the more demanding response deadline and the higher "stakes" of each incentive trial may have resulted in stronger representations of the task context in order to improve performance under conditions when trading speed for accuracy was disadvantageous. In favor of this explanation are previous results showing that task context representation in service of top down control activates the typical DLPFC areas (i.e. BA 46/9) (Barch et al., 1997; Cohen et al., 1997; D'Esposito et al., 1995; Jonides, Smith, Marshuetz, Koeppe, & Reuter-Lorenz, 1998; Kerns et al., 2004; MacDonald et al., 2000).

Another issue which emerged from these experiments was the functional distinction between the OFC and the cortex of the inferior convexity (IC), which in human includes mostly BA47, ventral aspects of BA10 rostrally and of BA45 caudally. In our experiments, the analyses of task effects identified significant interactions with time in the IC, both during preparation and post-response intervals (Tables 1, 2, 4 and 5). While in some analyses this interaction seemed to be due to relative signal decreases (e.g. in some of the right IC regions of interest), in the left IC activity was consistently higher during preparation for Non-match, suggestive of a role in

implementation of cognitive control required by that task. This interpretation is in line with findings from both animal and human experiments, though some inconsistencies still remain. For example, some studies have shown that monkeys with lesions of the IC have behavioral deficits more similar to those caused by DLPFC than by lesions of the ventromedial cortex (Bachevalier & Mishkin, 1986; Kowalska, Bachevalier, & Mishkin, 1991). Other authors, however (Mishkin & Manning, 1978), emphasize the fact that lesions centered around the principal sulcus of monkeys cause mainly deficits in performance of spatial tasks and only minimal impairments in objects-based tasks (such as delayed object alternation and delayed object matching), whereas lesions of the IC impair the latter tasks but not the former. In humans, lesions of the inferior lateral PFC, which frequently include the inferior convexity, have recently been argued to be most closely associated with deficits of active inhibitory control (Aron, Fletcher, Bullmore, Sahakian, & Robbins, 2003; Aron et al., 2004). Furthermore, imaging data have emphasized the role of the inferior aspects of the lateral PFC in inhibitory function (Bunge et al., 2002; Garavan et al., 1999) and in task switching (Konishi, Nakajima, Uchida, Kameyama et al., 1998; Konishi, Nakajima, Uchida, Sekihara et al., 1998; Sohn et al., 2000). While some of these inconsistencies might be accounted for by the imperfect homology of cortical regions across species, it is just as possible that the involvement of the IC cortex in active control processes reflects an important functional distinction from the rest of the prefrontal cortex on the ventral surface of the frontal lobes. In fact, the anatomical data in non-human primates support such a distinction. The cytoarchitecture and connectivity of the monkey IC are markedly different from that of the orbital cortical surface, which includes mainly dysgranular and even agranular cortex, and more similar to the granular cortex of the dorsolateral PFC (Barbas, 1992; Cavada, Company, Tejedor, Cruz-Rizzolo, & Reinoso-Suarez, 2000). Also, the connectivity of the inferior convexity is more

heavily biased towards premotor cortices whereas the OFC has relatively more connections with basal forebrain and limbic structures (Barbas, 1992; Ghashghaei & Barbas, 2001; Porrino, Crane, & Goldman-Rakic, 1981). These features may explain why, recently, areas of the human inferior convexity have been found active during goal selection (Arana et al., 2003; O'Doherty et al., 2003), strengthening the argument that the IC may actually be part of the same functional complex as the DLPFC.

A close examination of the connectivity data regarding the ventral OFC in monkeys leads to an interesting hypothesis. The cortex of the IC seems to include mainly the lateral portion of Walker's area 12 (corresponding to the most ventral aspects of BA 47 in humans). This area appears to already have a well defined granular layer, similar to the DLPFC. However, detailed reviews of existing connectivity data (Ongur & Price, 2000) have pointed out that the IC has extensive connections with the dysgranular orbital cortex (Walker areas 11 (caudo-medial), medial 12 and anterior 13). However, as it was outlined earlier in this document, the strong connectivity between the IC and other association cortices as well as with premotor areas, and the relatively weaker links with limbic and basal forebrain structures distinguish it from the rest of the ventral PFC. It is therefore possible that the IC serves as an integrator of motivational information in the service of complex abstract planning, by establishing and/or maintaining representations of associations between response rules and potential outcomes. This would account for the demonstrated involvement of inferior lateral PFC in the control of behavior when non-human primates have to detect a new complex stimulus dimension associated with reward (Dias & Aggleton, 2000; Dias et al., 1996a, 1996b) as well as the imaging results involving the inferior frontal gyrus in task switching (Konishi, Nakajima, Uchida, Kameyama et al., 1998;

Sohn et al., 2000), and No-go performance (Elliott, Rubinsztein, Sahakian, & Dolan, 2000; Garavan et al., 1999; Konishi, Nakajima, Uchida, Sekihara et al., 1998; Rubia et al., 2001).

5.2. The OFC activity is specific to predictors of negative outcomes

In both experiments presented in Chapters 2 and 3, variations in the level of incentives available for correct performance robustly modulated the activity in several areas of the PFC. However, not all PFC activations showed the same pattern of response. The lateral and dorsal PFC was more active when possible outcomes (i.e. incentives) had the effect of improving performance, consistent with a role for those cortical areas in implementation of control mechanisms. In contrast, the dysgranular cortex of the lateral OFC seemed to respond to the nature of the expected outcomes regardless whether they led to improvements in performance or not. As such, of the whole range of monetary incentives used in each experiment, the lateral OFC responded most strongly to the absence of positive incentives in Experiment 1, which coincided with indices of poorer performance, and to the presence of negative incentives in Experiment 2, which coincided with improvements in performance.

These results consistently spoke against a role of the OFC in implementation of control. However, they could not address all the complexities of the processes involved in translating changes of motivational context into task specific context. For instance, one important question that will need further investigation regards the precise mechanism through which the prospect of a different outcome results into increased efficacy of control processes. Previous research has shown that when rewards are available upon performing of a motor response, the neural activity increases in limbic and paralimbic structures as well as in lateral prefrontal areas frequently linked to implementation of control mechanisms. These effects have been apparent in single unit

studies of awake behaving non-human primates (Hikosaka & Watanabe, 2000; Leon & Shadlen, 1999; Roesch & Olson, 2003; Wallis & Miller, 2003; Watanabe et al., 2002) as well as in human neuroimaging studies (Pochon et al., 2002; Taylor et al., 2004; Thut et al., 1997). However, some of these results suggest that the effects of rewards on control-related activity may be progressively stronger in more caudal areas of the prefrontal cortex (Roesch & Olson, 2003). The effects of incentives described in Chapter 2 (Experiment 1) could, in retrospect, offer support for this possibility. In the left hemisphere, the increased activity to Reward trials peaked at the border of Brodmann's areas 8, 6 and 9 (Talairach coordinates -40, 14, 53), relatively more caudally than the maximal effect of task (BA 10/46, Talairach coordinates -36, 53, -4). However, in the following experiment, the overlap between incentive and task effects was significant (Talairach coordinates of the peak voxel in the left DLPFC areas: -44, 40, 26 and -37, 40, 28, respectively), though the ROI showing the incentive effect extended more caudally relative to the area with effects of task. Given the subtle differences between specific task requirements of each of our experiments, we cannot speculate as to how task-induced and incentive-induced enhancements in control could have interacted in each case. Further research will be necessary to clearly formulate hypotheses and test predictions regarding possible neural mechanisms underlying these interactions. For instance, a series of experiments can be envisioned in which the demands for control and speed/accuracy tradeoff would be kept constant, but in which levels of incentives would be parametrically manipulated, with the goal of examining whether and how these factors covary with the degree of overlap between task-induced and incentive-induced DLPFC activity.

Though not the focus of the experiment, another issue raised by the data was the role of other structures in processing of incentives. The medial prefrontal cortex, has been repeatedly

associated with processing of reinforcers or their predictors (Critchley et al., 2001; O'Doherty et al., 2003; Pochon et al., 2002; Rogers et al., 2004). As noted in section 2.3.3, in Experiment 1 the Incentive by Scan interaction was significant in the voxel of maximum F value, but the peak difference between signal changes to Reward and No-reward cues did not reach statistical significance. This could be due to the fact that the effects of rewards in the medial PFC may interact with the probability (Critchley et al., 2001; O'Doherty et al., 2003; Rogers et al., 2004) and with the magnitude (Elliott, Friston, & Dolan, 2000) of rewards. In our task, the rewards were relatively small, and they had a very high likelihood of occurrence, since producing a response that was both correct and faster than one second was only moderately challenging for subjects. Therefore, it is possible that these task features resulted in medial frontal wall activations which were not as strong as might have been expected. This may also account for the lack of activation in other regions such as the ventral striatum, which has previously been reported to respond to anticipation of monetary rewards (Breiter et al., 2001; Knutson, Adams, Fong, & Hommer, 2001) or primary rewards (Pagnoni et al., 2002).

A second structure frequently reported as sensitive to manipulations of primary and secondary reinforcers is the amygdala complex. Though its involvement in emotional processes has been frequently studied using aversive stimuli (Davidson, 2002; LeDoux, 1993; Morris & Dolan, 2004; Whalen et al., 1998), evidence has accumulated suggesting that this limbic structure may be equally important for processing of reward-related information (Baxter, Parker, Lindner, Izquierdo, & Murray, 2000; Bechara et al., 1999; Fried et al., 2001; Gottfried, O'Doherty, & Dolan, 2003), These results suggest that the amygdala is critical for processing of any emotionally-valenced information. This possibility is supported by data from our experiment, in which one of the regions of activation included the amygdala and related areas

such as the sublenticular extended amygdala (see Table 3 and Figure 5). In this ROI, activity peaked higher during anticipation of Rewards relative to No-reward trials. Obviously, given the limited resolution of existing fMRI datasets, it is not clear whether the reward vs. avoidance-related amygdala activations originate in different subnuclei or are completely overlapping, but generated by activations of distinct neuronal networks. Future studies in animal models and focused, high resolution imaging experiments in humans will be needed to directly address and distinguish between these possibilities.

While analyses of incentive effects in Experiment 2 replicated well the orbitofrontal foci of activation, they also resulted in some interesting differences in other reward-related areas. Most notably, while this analysis did not replicate the effects in the amygdala, it found effects of incentives in the left posterior hippocampal formation and in the substantia nigra (SN). In both these areas, the Incentive (Reward vs. Penalty vs. No-reward) by Scan (S1-S8) was significant both in the exploratory analysis and in the normalized signal changes. Those effects were driven by increases in both Reward and Penalty trials relative to No-reward (see Figure 19). Though the relationship between the hippocampal formation and motivation has not been the main focus of human neuroimaging studies, hippocampal activations have been previously noted in response to manipulations of reward (Berns, McClure, Pagnoni, & Montague, 2001; Elliott, Friston et al., 2000). Furthermore, data from animal studies have suggested a role for the hippocampus in motivated behaviors (Depue & Collins, 1999; Tracy, Jarrard, & Davidson, 2001).

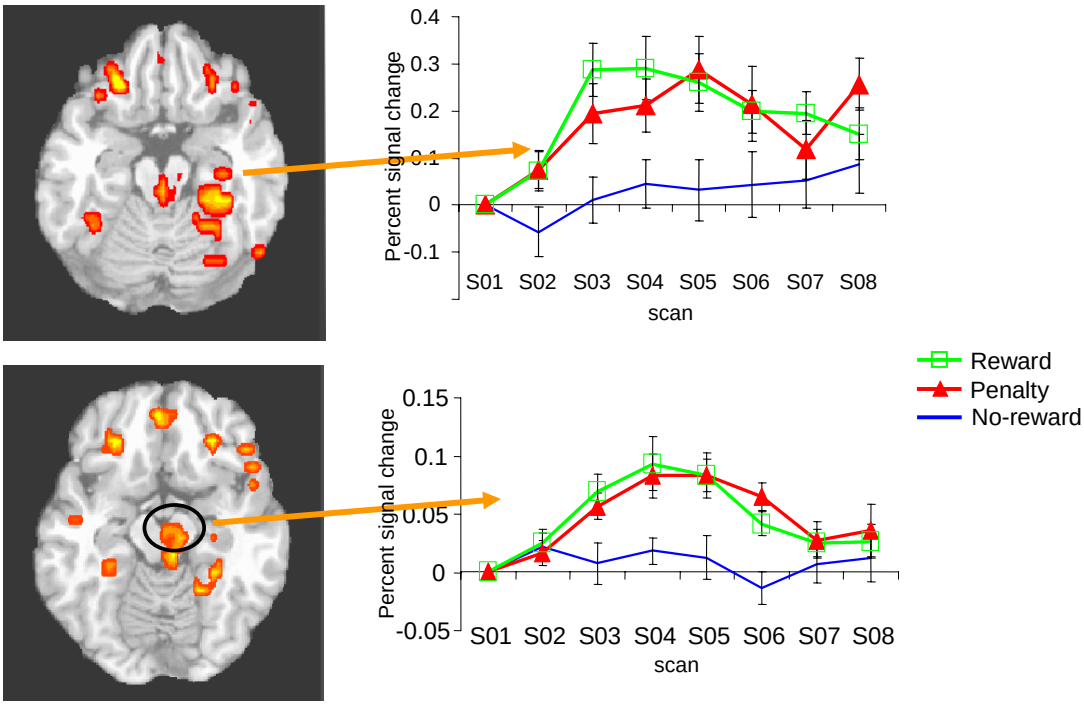

Figure 19. Incentive effects in the hippocampus and substantia nigra

In Experiment 2, both left posterior hippocampal formation and substantia nigra showed increased activity in response to cues signaling Reward or Penalty relative to No-reward cues.

Another interesting question that arose from the analysis of incentive effects on OFC activity is related to the precise neural mechanisms of expected feelings of pleasure and disappointment. As has been detailed in section 3.4.2, the pattern of OFC activity to Reward cues in the two experiments is consistent with new theories of decision making which propose that decisions are made in order to maximize the feelings of pleasure and minimize the disappointment that subjects anticipate as a result of multiple possible outcomes (Mellers et al., 1999; Mellers, 2001; Mellers et al., 1997). In the context of our paradigms, the increase in lateral OFC activity to Reward trials noted in Experiment 2 relative to Experiment 1 is consistent with a correlate of anticipated feelings of disappointment as a result of increased likelihood of missing the opportunity for a reward. However, the same theoretical framework necessarily implies that

some mechanisms must also exist which compute the expected *positive* impact of potential outcomes, i.e expected feelings of elation or pleasure (Mellers, 2000). The neural substrate of such a processes can only be speculated at this point. Obvious candidates, based on previous neuroimaging experiments, would be the medial OFC (Critchley et al., 2001; O'Doherty et al., 2001; O'Doherty et al., 2002; Small et al., 2001) and the cortex of the medial wall (Critchley et al., 2001; Elliott et al., 2003; O'Doherty et al., 2003; Pochon et al., 2002; Rogers et al., 1999; Rogers et al., 2004; Small et al., 2001). Our results provided evidence of increased medial prefrontal activity to Reward cues relative to No-reward in both experiments presented in Chapters 2 and 3. In contrast, the effects of reward anticipation in the medial OFC were not so clear, as the patterns of signal changes to anticipation of reward and penalty trials were quite complex. It is worth noting, at this point, that single unit recordings in humans have generated preliminary results suggesting that the proportion of cells in the medial OFC of humans which react specifically to negatively valenced information is higher than the proportion of cells specific to either positively valenced or neutral information (Kawasaki, Adolphs, Oya, Kovach, & Howard, 2004). Therefore, it would be premature to make any strong statements regarding the possible specificity of the medial OFC for processing rewards versus punishments.

An alternative interpretation of the increased OFC activity to penalty cues of Experiment 2 is that it reflects a greater emotional impact of potential losses, given that subjects frequently report dreading a monetary loss more than they value the pleasure of a gain of equal size (Kahneman & Tversky, 1988). Matching the indices of behavioral performance during Reward and Penalty trials, through the use of the individualized dynamic response deadline, speaks against this interpretation, as does the increased activity during No-reward relative to Reward trials in Experiment 1, when no actual losses were expected. However, it could be argued that the

lack of behavioral difference between Penalty and Reward trials is actually a ceiling effect, induced by the increased pressure for speeded responses. In order to address this possibility, we conducted an analysis of physiological data (i.e. heart rate) collected in a subgroup of 10 subjects while they performed the task in the scanner (see Appendix C). Overall, that analysis looked for evidence of heart rate deceleration in anticipation of targets, which have been repeatedly described in cognitive tasks with fixed inter-stimulus intervals (Jennings, van der Molen, & Brock, 1997; Jennings, van der Molen, & Somsen, 1998; Jennings, van der Molen, & Steinhauer, 1998; van der Molen, Boomsma, Jennings, & Nieuwboer, 1989). No evidence for increased heart rate deceleration in preparation for targets of Penalty trials was found; in fact, there was a trend for a greater deceleration in anticipation of targets of Reward trials relative to both other trial types.

5.3. The OFC hyperactivity in OCD

The data presented in Chapter 4 were aimed at providing more direct evidence that the OFC hyperactivity previously reported in obsessive compulsive disorder (OCD) represents a fundamental neural deficit rather than an epiphenomenon of the symptomatic state. In doing so, we sought to confirm that this hyperactivity would follow a prediction generated by the theoretical framework that OFC is critical for representing possible negative outcomes. A group of OCD patients and a matched control group of healthy subjects were scanned in an event-related fMRI protocol while they were performing a cognitive task which used letters as stimuli. The frequency of various letter cues and probes was such that one type of cue was a strong predictor of a positive outcome (i.e. occurrence of the letter X which constituted a target), while another cue type predicted a high likelihood of a relatively negative outcome (i.e. the infrequent

occurrence of an X which did not constitute a target). Relative to control subjects, the patients' OFC activity did not differ in response to cues which predicted a high likelihood of target X's, but was increased in the right lateral OFC in response to cues which were predictors of the non-target X event.

The fact that the increase in OFC activity of OCD patients was elicited by stimuli which are not relevant to OC symptomatology provide support for the view that this hyperactivity is the correlate of an underlying dysfunction of this cortical area, and is inconsistent with the possibility that it represents merely an epiphenomenon of symptoms. Furthermore, its specificity for predictors of non-target events is in line with the theoretical framework established by the previous two experiments presented in this thesis. As such, it is consistent with a role of the OFC in representing negatively valenced possible outcomes.

One issue raised by the findings presented in Chapter 4 concerned the spatial resolution of fMRI and its implications for integration of these results with those of Chapters 2 and 3. The results of the first two experiments suggested that there might be significant differences between the type of executive function subserved by the OFC and the neighboring cortex of the inferior convexity (IC). At the same time, the OFC focus of hyperactivity in OCD patients was slightly more lateral than the OFC foci identified in Experiments 1 and 2. Therefore, we computed the relative distances between peaks of task and incentive effects in Experiments 1 and 2 relative to the peak of the OFC activation which showed differences between OCD patients and controls. The results of this analysis are presented in Appendix D. In summary, it can be observed that the distances between the focus of hyperactivity in the OCD group seems to be roughly equidistant from the peaks of incentive and task effects of Experiment 1. Those distances were both considerably larger than the estimate of spatial smoothing (FWHM) of the statistical maps.

Based on these results alone, it would be hard to interpret the OCD peak of hyperactivity as being more closely located to one or the other of the two effects (task and incentives) found in Experiment 1. However, the distance between the location of OCD hyperactivity focus and the location of incentive effects in Experiment 2 was comparable to the effective spatial resolution of the statistical maps; at the same time, the relative separation from the locus of task effects was much clearer. This analysis suggested, overall, that the location of the OFC hyperactivity in OCD is more similar to that of incentive effects than to the location of task effects in the prefrontal cortex. Considering that the comparison of OCD with controls used a regular spiral fMRI acquisition protocol, it is also possible that effects were also present in more medial areas of the OFC, but were masked by signal loss due to magnetic susceptibility artifacts (Noll et al., 1995). Therefore, future studies are needed, using the newly developed imaging method used in the first two experiments, to clarify the issues of precise localization of altered OFC activity in OCD patients.

Another question raised by this experiment is the issue of laterality of OFC activity. In most of the analyses described in Experiments 1 and 2, the clearest effects of negative outcomes were noted in the left OFC (see Tables 3, 4, 6 and Figures 5, 12, 13). In contrast, the increase in OFC activity to B cues in OCD patients was noted in the right hemisphere. Several factors may have contributed to these differences. One obvious possible influence could have been the technical limitations mentioned in the preceding paragraph. It is possible that, even if patients' left OFC was also hyperactive, this effect was missed if its location in the left hemisphere was more medial and therefore more susceptible to signal dropout. Another source of difference may reside, though, in asymmetries of the functional specializations of the two hemispheres and interactions of these asymmetries with the pathological processes involved in OCD. Previous

results have suggested that the right hemisphere might have a bias for processing emotional information, particularly if negatively valenced (Davidson, 1992). This would be consistent with some evidence of right-lateralized hyperactivity in patients, and with correlations between the right OFC activity and symptomatology (Kwon et al., 2003; Mataix-Cols et al., 2004). However, bilateral abnormalities and correlations of left lateralized effects and treatment outcomes are also quite frequent in functional imaging studies of OCD (Brody et al., 1998; Swedo et al., 1992). It is very difficult to address all the complexities of these issues at this time. One possibility is that the vast majority of the existing evidence linking OFC dysfunction to OCD has been obtained in functional imaging experiments which used blocked designs (for a review, see (Saxena et al., 1998)), and therefore might have biased results toward identifying differences in tonic brain activity. In contrast, our experiments used event-related designs which focus on the differences in phasic activity between processes or subject groups.

A third issue that cannot be addressed by our study of OCD patients is the possibility that the OFC hyperactivity of patients reflects a compensatory mechanism rather than a pathological one (Stein et al., 1999). This possible interpretation of abnormally increased activations is very difficult to rule out, especially in subjects with a disease such as OCD, who frequently can adapt to experimental situations relatively well, as was the case in our study. While our data cannot speak directly to this problem, other lines of research may help shed some light on this issue. Overall, several pieces of evidence suggesting intrinsic abnormalities of the OFC are inconsistent with interpreting OFC hyperactivity in terms of compensatory activity. For instance, some results suggest that anxiety correlates with biochemical markers of neuronal dysfunction in the human OFC (Grachev & Apkarian, 2000). Furthermore, findings of decreased GABA-A binding in OFC of patients with anxiety disorders (Malizia et al., 1998), together with the fact that disrupting the

function of GABA-A receptors in animal models results in increased anxiety (Crestani et al., 1999) suggest a causal relationship between impaired inhibitory OFC circuitry and anxiety.

5.4. Conclusions and future directions

In summary, the data presented in this thesis were aimed at furthering our understanding regarding the functional organization of the human prefrontal cortex, with emphasis on the role of its orbitofrontal areas in executive functions. The results presented in Chapters 2 and 3 provided converging evidence for three important aspects regarding the OFC involvement in executive control of behavior. Firstly, the results showed that the OFC is not involved in active top-down control processes. Secondly, they demonstrated that the contribution of lateral OFC areas to guiding behavior is through representation of the motivational context established by potential negative outcomes of actions. Finally, these data suggested that the outcome representations supported by the OFC may be influenced by the full range of possible outcomes, consequently contributing to anticipated feelings of pleasure and disappointment which may play a major role in decision making. In this context, it is possible that the executive control of behavior arises from interactions within a distributed network of structures, in which the orbitofrontal cortex integrates sensory information with affective information (processed in limbic-related structures such as the amygdala and the basal ganglia), in order to represent the emotional and motivational impact of potential outcomes. This “motivational context” can further be integrated with representations of task-specific context (presumably in the lateral and dorsal PFC), in the service of biasing the output mechanisms towards optimal behavior. Obviously, further research is necessary to test the explanatory power of this complex model, and several directions of this research can be envisioned already. For instance, it is important to

precisely characterize the mechanisms of translating changes in motivational context into changes in active control mechanisms. In this domain, one important question that will need to be addressed is whether all types of control processes are equally susceptible to the influence of motivational factors. At the other spectrum of interactions between motivation and control, much remains to be done to explore the precise neural underpinnings of representations of motivational context. At the behavioral level, it is known that perceived impact of outcomes are modulated by factors such as beliefs and individual dispositional traits (e.g. personality factors), and not much is known about their neural substrates. Identifying these substrates is critical for future efforts aimed at understanding behavioral control. The results presented here suggest that neuroimaging methodologies, in the context of clear theoretical constraints, can make very important contributions toward such efforts.

BIBLIOGRAPHY

Adams, B. L., Warneke, L. B., McEwan, A. J., & Fraser, B. A. (1993). Single photon emission computerized tomography in obsessive compulsive disorder: a preliminary study. *Journal of Psychiatry & Neuroscience, 18*(3), 109-112.

Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience, 9*, 357-381.

Alptekin, K., Degirmenci, B., Kivircik, B., Durak, H., Yemez, B., Derebek, E., & Tunca, Z. (2001). Tc-99m HMPAO brain perfusion SPECT in drug-free obsessive-compulsive patients without depression. *Psychiatry Res, 107*(1), 51-56.

Arana, F. S., Parkinson, J. A., Hinton, E., Holland, A. J., Owen, A. M., & Roberts, A. C. (2003). Dissociable Contributions of the Human Amygdala and Orbitofrontal Cortex to Incentive Motivation and Goal Selection. *J. Neurosci., 23*(29), 9632-9638.

Aron, A. R., Fletcher, P. C., Bullmore, E. T., Sahakian, B. J., & Robbins, T. W. (2003). Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. *Nat Neurosci, 6*(2), 115-116.

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends Cogn Sci, 8*(4), 170-177.

Bachevalier, J., & Mishkin, M. (1986). Visual recognition impairment follows ventromedial but not dorsolateral prefrontal lesions in monkeys. *Behavioural Brain Research, 20*(3), 249-261.

Barbas, H. (1992). Architecture and cortical connections of the prefrontal cortex in the rhesus monkey. *Adv Neurol*, 5791-5115.

Barbas, H., & Pandya, D. N. (1989). Architecture and intrinsic connections of the prefrontal cortex in the rhesus monkey. *J Comp Neurol, 286*(3), 353-375.

Barch, D. M., Braver, T. S., Nystrom, L. E., Forman, S. D., Noll, D. C., & Cohen, J. D. (1997). Dissociating working memory from task difficulty in human prefrontal cortex. *Neuropsychologia, 35*(10), 1373-1380.

Baxter, L. R., Jr., Schwartz, J. M., Bergman, K. S., Szuba, M. P., Guze, B. H., Mazziotta, J. C., Alazraki, A., Selin, C. E., Ferng, H. K., & Munford, P. (1992). Caudate glucose metabolic rate changes with both drug and behavior therapy for obsessive-compulsive disorder. *Archives of General Psychiatry, 49*(9), 681-689.

Baxter, L. R., Jr., Schwartz, J. M., Guze, B. H., Bergman, K., & Szuba, M. P. (1990). PET imaging in obsessive compulsive disorder with and without depression. *Journal of Clinical Psychiatry, 51 Suppl*, 61-69; discussion 70.

Baxter, L. R., Jr., Schwartz, J. M., Mazziotta, J. C., Phelps, M. E., Pahl, J. J., Guze, B. H., & Fairbanks, L. (1988). Cerebral glucose metabolic rates in nondepressed patients with obsessive-compulsive disorder. *American Journal of Psychiatry, 145*(12), 1560-1563.

Baxter, M. G., Parker, A., Lindner, C. C., Izquierdo, A. D., & Murray, E. A. (2000). Control of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex. *J Neurosci, 20*(11), 4311-4319.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition, 50*(1-3), 7-15.
Bechara, A., Damasio, H., Damasio, A. R., & Lee, G. P. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *Journal of Neuroscience, 19*(13), 5473-5481.
Bechara, A., Damasio, H., Tranel, D., & Anderson, S. W. (1998). Dissociation Of working memory from decision making within the human prefrontal cortex. *Journal of Neuroscience, 18*(1), 428-437.
Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science, 275*(5304), 1293-1295.
Bechara, A., Tranel, D., & Damasio, H. (2000). Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain, 123*(Pt 11), 2189-2202.
Bechara, A., Tranel, D., Damasio, H., & Damasio, A. R. (1996). Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. *Cereb Cortex, 6*(2), 215-225.
Benkelfat, C., Nordahl, T. E., Semple, W. E., King, A. C., Murphy, D. L., & Cohen, R. M. (1990). Local cerebral glucose metabolic rates in obsessive-compulsive disorder. Patients treated with clomipramine. *Archives of General Psychiatry, 47*(9), 840-848.
Berlin, H. A., Rolls, E. T., & Kischka, U. (2004). Impulsivity, time perception, emotion and reinforcement sensitivity in patients with orbitofrontal cortex lesions. *Brain, 127*(Pt 5), 1108-1126.
Berns, G. S., McClure, S. M., Pagnoni, G., & Montague, P. R. (2001). Predictability modulates human brain response to reward. *J Neurosci, 21*(8), 2793-2798.
Berthier, M. L., Kulisevsky, J., Gironell, A., & Heras, J. A. (1996). Obsessive-compulsive disorder associated with brain lesions: clinical phenomenology, cognitive function, and anatomic correlates [published erratum appears in Neurology 1996 Sep;47(3):855]. *Neurology, 47*(2), 353-361.
Breiter, H. C., Aharon, I., Kahneman, D., Dale, A., & Shizgal, P. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron, 30*(2), 619-639.
Breiter, H. C., & Rauch, S. L. (1996). Functional MRI and the study of OCD: from symptom provocation to cognitive-behavioral probes of cortico-striatal systems and the amygdala. *Neuroimage, 4*(3 Pt 3), S127-138.
Brody, A. L., Saxena, S., Schwartz, J. M., Stoessel, P. W., Maidment, K., Phelps, M. E., & Baxter, L. R., Jr. (1998). FDG-PET predictors of response to behavioral therapy and pharmacotherapy in obsessive compulsive disorder. *Psychiatry Research, 84*(1), 1-6.
Buchsbaum, M. S., Hollander, E., Haznedar, M. M., Tang, C., Spiegel-Cohen, J., Wei, T. C., Solimando, A., Buchsbaum, B. R., Robins, D., Bienstock, C., Cartwright, C., & Mosovich, S. (2001). Effect of fluoxetine on regional cerebral metabolism in autistic spectrum disorders: a pilot study. *Int J Neuropsychopharmacol, 4*(2), 119-125.
Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., & Gabrieli, J. D. (2002). Immature frontal lobe contributions to cognitive control in children: evidence from fMRI. *Neuron, 33*(2), 301-311.
Busatto, G. F., Zamignani, D. R., Buchpiguel, C. A., Garrido, G. E., Glabus, M. F., Rocha, E. T., Maia, A. F., Rosario-Campos, M. C., Campi Castro, C., Furuie, S. S., Gutierrez, M. A., McGuire, P. K., & Miguel, E. C. (2000). A voxel-based investigation of regional cerebral

blood flow abnormalities in obsessive-compulsive disorder using single photon emission computed tomography (SPECT). *Psychiatry Research, 99*(1), 15-27.
Camille, N., Coricelli, G., Sallet, J., Pradat-Diehl, P., Duhamel, J.-R., & Sirigu, A. (2004). The Involvement of the Orbitofrontal Cortex in the Experience of Regret. *Science, 304*(5674), 1167-1170.
Carmichael, S. T., & Price, J. L. (1994). Architectonic subdivision of the orbital and medial prefrontal cortex in the macaque monkey. *J Comp Neurol, 346*(3), 366-402.
Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science, 280*(5364), 747-749.
Casey, B. J., Trainor, R. J., Orendi, J. L., Schubert, A. B., Nystrom, L. E., Giedd, J. N., Castellanos, F. X., Haxby, J. V., Noll, D. C., Cohen, J. D., Forman, S. D., Dahl, R. E., & Rapoport, J. L. (1997). A developmental functional MRI study of prefrontal activation during performance of a Go-No-Go task. *Journal of Cognitive Neuroscience, 9*(6), 835-847.
Cavada, C., Company, T., Tejedor, J., Cruz-Rizzolo, R. J., & Reinoso-Suarez, F. (2000). The anatomical connections of the macaque monkey orbitofrontal cortex. A review. *Cereb Cortex, 10*(3), 220-242.
Chua, P., Krams, M., Toni, I., Passingham, R., & Dolan, R. (1999). A functional anatomy of anticipatory anxiety. *Neuroimage, 9*(6 Pt 1), 563-571.
Clark, L., Cools, R., & Robbins, T. W. (2004). The neuropsychology of ventral prefrontal cortex: Decision-making and reversal learning. *Brain and Cognition, 55*(1), 41-53.
Clarke, H. F., Dalley, J. W., Crofts, H. S., Robbins, T. W., & Roberts, A. C. (2004). Cognitive Inflexibility After Prefrontal Serotonin Depletion. *Science, 304*(5672), 878-880.
Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., & Smith, E. E. (1997). Temporal dynamics of brain activation during a working memory task [see comments]. *Nature, 386*(6625), 604-608.
Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: a connectionist approach to behavior and biology in schizophrenia. *Psychological Review, 99*(1), 45-77.
Cools, R., Clark, L., Owen, A. M., & Robbins, T. W. (2002). Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *J Neurosci, 22*(11), 4563-4567.
Cools, R., Clark, L., & Robbins, T. W. (2004). Differential Responses in Human Striatum and Prefrontal Cortex to Changes in Object and Rule Relevance. *J. Neurosci., 24*(5), 1129-1135.
Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res, 29*(3), 162-173.
Crestani, F., Lorez, M., Baer, K., Essrich, C., Benke, D., Laurent, J. P., Belzung, C., Fritschy, J. M., Luscher, B., & Mohler, H. (1999). Decreased GABAA-receptor clustering results in enhanced anxiety and a bias for threat cues. *Nat Neurosci, 2*(9), 833-839.
Critchley, H. D., Mathias, C. J., & Dolan, R. J. (2001). Neural activity in the human brain relating to uncertainty and arousal during anticipation. *Neuron, 29*(2), 537-545.
Cummings, J. L. (1995). Anatomic and behavioral aspects of frontal-subcortical circuits. *Annals of the New York Academy of Sciences, 769*, 1-13.

Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences, 351*(1346), 1413-1420.

Daum, I., Channon, S., Polkey, C. E., & Gray, J. A. (1991). Classical conditioning after temporal lobe lesions in man: impairment in conditional discrimination. *Behav Neurosci, 105*(3), 396-408.

Daum, I., Schugens, M. M., Channon, S., Polkey, C. E., & Gray, J. A. (1991). T-maze discrimination and reversal learning after unilateral temporal or frontal lobe lesions in man. *Cortex, 27*(4), 613-622.

Davidson, R. J. (1992). Emotion and affective style: Hemispheric substrates. *Psychological Science, 3*(1), 39-43.

Davidson, R. J. (2002). Anxiety and affective style: role of prefrontal cortex and amygdala. *Biol Psychiatry, 51*(1), 68-80.

De Araujo, I. E., & Rolls, E. T. (2004). Representation in the human brain of food texture and oral fat. *J Neurosci, 24*(12), 3086-3093.

Delgado, M. R., Nystrom, L. E., Fissell, C., Noll, D. C., & Fiez, J. A. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *J Neurophysiol, 84*(6), 3072-3077.

Depue, R. A., & Collins, P. F. (1999). Neurobiology of the structure of personality: dopamine, facilitation of incentive motivation, and extraversion. *Behav Brain Sci, 22*(3), 491-517.

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

D'Esposito, M., Detre, J. A., Alsop, D. C., Shin, R. K., Atlas, S., & Grossman, M. (1995). The neural basis of the central executive system of working memory. *Nature, 378*(6554), 279-281.

D'Esposito, M., Postle, B. R., Jonides, J., & Smith, E. E. (1999). The neural substrate and temporal dynamics of interference effects in working memory as revealed by event-related functional MRI. *Proc Natl Acad Sci U S A, 96*(13), 7514-7519.

Dias, R., & Aggleton, J. P. (2000). Effects of selective excitotoxic prefrontal lesions on acquisition of nonmatching- and matching-to-place in the T-maze in the rat: differential involvement of the prelimbic-infralimbic and anterior cingulate cortices in providing behavioural flexibility. *European Journal of Neuroscience, 12*(12), 4457-4466.

Dias, R., Robbins, T. W., & Roberts, A. C. (1996a). Dissociation in prefrontal cortex of affective and attentional shifts. *Nature, 380*(6569), 69-72.

Dias, R., Robbins, T. W., & Roberts, A. C. (1996b). Primate analogue of the Wisconsin Card Sorting Test: effects of excitotoxic lesions of the prefrontal cortex in the marmoset. *Behav Neurosci, 110*(5), 872-886.

Dias, R., Robbins, T. W., & Roberts, A. C. (1997). Dissociable forms of inhibitory control within prefrontal cortex with an analog of the Wisconsin Card Sort Test: restriction to novel situations and independence from "on-line" processing. *Journal of Neuroscience, 17*(23), 9285-9297.

Eblen, F., & Graybiel, A. M. (1995). Highly restricted origin of prefrontal cortical inputs to striosomes in the macaque monkey. *J Neurosci, 15*(9), 5999-6013.

Elliott, R., & Dolan, R. J. (1999). Differential neural responses during performance of matching and nonmatching to sample tasks at two delay intervals. *Journal of Neuroscience, 19*(12), 5066-5073.

Elliott, R., Dolan, R. J., & Frith, C. D. (2000). Dissociable functions in the medial and lateral orbitofrontal cortex: evidence from human neuroimaging studies. *Cereb Cortex, 10*(3), 308-317.

Elliott, R., Friston, K. J., & Dolan, R. J. (2000). Dissociable neural responses in human reward systems. *Journal of Neuroscience, 20*(16), 6159-6165.

Elliott, R., Newman, J. L., Longe, O. A., & Deakin, J. F. W. (2003). Differential Response Patterns in the Striatum and Orbitofrontal Cortex to Financial Reward in Humans: A Parametric Functional Magnetic Resonance Imaging Study. *J. Neurosci., 23*(1), 303-307.

Elliott, R., Newman, J. L., Longe, O. A., & William Deakin, J. F. (2004). Instrumental responding for rewards is associated with enhanced neuronal response in subcortical reward systems. *Neuroimage, 21*(3), 984-990.

Elliott, R., Rubinsztein, J. S., Sahakian, B. J., & Dolan, R. J. (2000). Selective attention to emotional stimuli in a verbal go/no-go task: an fMRI study. *Neuroreport, 11*(8), 1739-1744.

Elliott, R., Sahakian, B. J., Michael, A., Paykel, E. S., & Dolan, R. J. (1998). Abnormal neural response to feedback on planning and guessing tasks in patients with unipolar depression. *Psychol Med, 28*(3), 559-571.

Enright, S. J., & Beech, A. R. (1993). Reduced cognitive inhibition in obsessive-compulsive disorder. *British Journal of Clinical Psychology, 32*(Pt 1), 67-74.

Fellows, L. K., & Farah, M. J. (2003). Ventromedial frontal cortex mediates affective shifting in humans: evidence from a reversal learning paradigm. *Brain, 126*(Pt 8), 1830-1837.

Ferry, A. T., Ongur, D., An, X., & Price, J. L. (2000). Prefrontal cortical projections to the striatum in macaque monkeys: evidence for an organization related to prefrontal networks. *J Comp Neurol, 425*(3), 447-470.

Fitzgerald, K. D., Moore, G. J., Paulson, L. A., Stewart, C. M., & Rosenberg, D. R. (2000). Proton spectroscopic imaging of the thalamus in treatment-naive pediatric obsessive-compulsive disorder [see comments]. *Biological Psychiatry, 47*(3), 174-182.

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magnetic Resonance in Medicine, 33*(5), 636-647.

Fredrikson, M., Wik, G., Annas, P., Ericson, K., & Stone-Elander, S. (1995). Functional neuroanatomy of visually elicited simple phobic fear: additional data and theoretical analysis. *Psychophysiology, 32*(1), 43-48.

Fried, I., Wilson, C. L., Morrow, J. W., Cameron, K. A., Behnke, E. D., Ackerson, L. C., & Maidment, N. T. (2001). Increased dopamine release in the human amygdala during performance of cognitive tasks. *Nat Neurosci, 4*(2), 201-206.

Funahashi, S., Chafee, M. V., & Goldman-Rakic, P. S. (1993). Prefrontal neuronal activity in rhesus monkeys performing a delayed anti-saccade task. *Nature, 365*(6448), 753-756.

Fuster, J. M. (1997). *The prefrontal cortex : anatomy, physiology, and neuropsychology of the frontal lobe* (3rd ed.). Philadelphia: Lippincott-Raven.

Garavan, H., Ross, T. J., & Stein, E. A. (1999). Right hemispheric dominance of inhibitory control: an event-related functional MRI study. *Proceedings of the National Academy of Sciences of the United States of America, 96*(14), 8301-8306.

Gehring, W. J., Himle, J., & Nisenson, L. G. (2000). Action-monitoring dysfunction in obsessive-compulsive disorder. *Psychological Science, 11*(1), 1-6.

Ghashghaei, H. T., & Barbas, H. (2001). Neural interaction between the basal forebrain and functionally distinct prefrontal cortices in the rhesus monkey. *Neuroscience, 103*(3), 593-614.

Gilbert, A. R., Moore, G. J., Keshavan, M. S., Paulson, L. A., Narula, V., Mac Master, F. P., Stewart, C. M., & Rosenberg, D. R. (2000). Decrease in thalamic volumes of pediatric patients with obsessive-compulsive disorder who are taking paroxetine. *Archives of General Psychiatry, 57*(5), 449-456.

Glover, G. H., & Law, C. S. (2001). Spiral-in/out BOLD fMRI for increased SNR and reduced susceptibility artifacts. *Magn Reson Med, 46*(3), 515-522.

Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Delgado, P., Heninger, G. R., & Charney, D. S. (1989). The Yale-Brown Obsessive Compulsive Scale. II. Validity. *Archives of General Psychiatry, 46*(11), 1012-1016.

Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Fleischmann, R. L., Hill, C. L., Heninger, G. R., & Charney, D. S. (1989). The Yale-Brown Obsessive Compulsive Scale. I. Development, use, and reliability. *Archives of General Psychiatry, 46*(11), 1006-1011.

Gottfried, J. A., O'Doherty, J., & Dolan, R. J. (2003). Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science, 301*(5636), 1104-1107.

Grachev, I. D., & Apkarian, A. V. (2000). Anxiety in healthy humans is associated with orbital frontal chemistry. *Mol Psychiatry, 5*(5), 482-488.

Grafman, J., Vance, S. C., Weingartner, H., Salazar, A. M., & Amin, D. (1986). The effects of lateralized frontal lesions on mood regulation. *Brain, 109 (Pt 6)*, 1127-1148.

Gusnard, D. A., & Raichle, M. E. (2001). Searching for a baseline: functional imaging and the resting human brain. *Nature Reviews Neuroscience, 2*(10), 685-694.

Hartston, H. J., & Swerdlow, N. R. (1999). Visuospatial priming and stroop performance in patients with obsessive compulsive disorder. *Neuropsychology, 13*(3), 447-457.

Hikosaka, K., & Watanabe, M. (2000). Delay activity of orbital and lateral prefrontal neurons of the monkey varying with different rewards. *Cereb Cortex, 10*(3), 263-271.

Hornak, J., O'Doherty, J., Bramham, J., Rolls, E. T., Morris, R. G., Bullock, P. R., & Polkey, C. E. (2004). Reward-related reversal learning after surgical excisions in orbito-frontal or dorsolateral prefrontal cortex in humans. *J Cogn Neurosci, 16*(3), 463-478.

Horwitz, B., Swedo, S. E., Grady, C. L., Pietrini, P., Schapiro, M. B., Rapoport, J. L., & Rapoport, S. I. (1991). Cerebral metabolic pattern in obsessive-compulsive disorder: altered intercorrelations between regional rates of glucose utilization. *Psychiatry Research, 40*(4), 221-237.

Iversen, S. D., & Mishkin, M. (1970). Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. *Exp Brain Res, 11*(4), 376-386.

Jennings, J. R., van der Molen, M. W., & Brock, K. (1997). Mnemonic search, but not arithmetic transformation, is associated with psychophysiological inhibition. *J Exp Psychol Hum Percept Perform, 23*(1), 154-167.

Jennings, J. R., van der Molen, M. W., & Somsen, R. J. (1998). Changes in heart beat timing: reactivity, resetting, or perturbation? *Biol Psychol, 47*(3), 227-241.

Jennings, J. R., van der Molen, M. W., & Steinhauer, S. R. (1998). Preparing the heart, eye, and brain: foreperiod length effects in a nonaging paradigm. *Psychophysiology, 35*(1), 90-98.

Jones, E. G., & Powell, T. P. (1970). An anatomical study of converging sensory pathways within the cerebral cortex of the monkey. *Brain, 93*(4), 793-820.

Jonides, J., Smith, E. E., Marshuetz, C., Koeppe, R. A., & Reuter-Lorenz, P. A. (1998). Inhibition in verbal working memory revealed by brain activation. *Proc Natl Acad Sci U S A, 95*(14), 8410-8413.

Kahneman, D., & Tversky, A. (1988). Prospect Theory: An Analysis of Decision under Risk, *Gardenfors, Peter, ed; Sahlin, Nils Eric, ed; Decision, probability, and utility: Selected readings.. Cambridge; New York and Melbourne: Cambridge University Press 1988; 183 214. Previously published: 1979.*

Kawasaki, H., Adolphs, R., Oya, H., Kovach, C., & Howard, M. A. I. (2004). *Differential single-unit responses to emotional visual stimuli in the human prefrontal cortex.* Paper presented at the Annual meeting of the Cognitive Neuroscience Society, San Francisco.

Kerns, J. G., Cohen, J. D., MacDonald, A. W., 3rd, Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science, 303*(5660), 1023-1026.

Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J Neurosci, 21*(16), Rc159.

Koechlin, E., Basso, G., Pietrini, P., Panzer, S., & Grafman, J. (1999). The role of the anterior prefrontal cortex in human cognition. *Nature, 399*(6732), 148-151.

Konishi, S., Nakajima, K., Uchida, I., Kameyama, M., Nakahara, K., Sekihara, K., & Miyashita, Y. (1998). Transient activation of inferior prefrontal cortex during cognitive set shifting. *Nat Neurosci, 1*(1), 80-84.

Konishi, S., Nakajima, K., Uchida, I., Sekihara, K., & Miyashita, Y. (1998). No-go dominant brain activity in human inferior prefrontal cortex revealed by functional magnetic resonance imaging. *Eur J Neurosci, 10*(3), 1209-1213.

Kowalska, D. M., Bachevalier, J., & Mishkin, M. (1991). The role of the inferior prefrontal convexity in performance of delayed nonmatching-to-sample. *Neuropsychologia, 29*(6), 583-600.

Kringelbach, M. L., & Rolls, E. T. (2004). The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. *Prog Neurobiol, 72*(5), 341-372.

Kwon, J. S., Kim, J. J., Lee, D. W., Lee, J. S., Lee, D. S., Kim, M. S., Lyoo, I. K., Cho, M. J., & Lee, M. C. (2003). Neural correlates of clinical symptoms and cognitive dysfunctions in obsessive-compulsive disorder. *Psychiatry Res, 122*(1), 37-47.

Larsen, J. T., McGraw, A. P., Mellers, B. A., & Cacioppo, J. T. (2004). The Agony of Victory and Thrill of Defeat. Mixed Emotional Reactions to Disappointing Wins and Relieving Losses. *Psychological Science, 15*(5), 325-330.

LeDoux, J. E. (1993). Emotional memory systems in the brain. *Behav Brain Res, 58*(1-2), 69-79.

Leon, M. I., & Shadlen, M. N. (1999). Effect of expected reward magnitude on the response of neurons in the dorsolateral prefrontal cortex of the macaque. *Neuron, 24*(2), 415-425.

Liotti, M., Mayberg, H. S., Brannan, S. K., McGinnis, S., Jerabek, P., & Fox, P. T. (2000). Differential limbic--cortical correlates of sadness and anxiety in healthy subjects: implications for affective disorders. *Biol Psychiatry, 48*(1), 30-42.

London, E. D., Simon, S. L., Berman, S. M., Mandelkern, M. A., Lichtman, A. M., Bramen, J., Shinn, A. K., Miotto, K., Learn, J., Dong, Y., Matochik, J. A., Kurian, V., Newton, T., Woods, R., Rawson, R., & Ling, W. (2004). Mood disturbances and regional cerebral metabolic abnormalities in recently abstinent methamphetamine abusers. *Arch Gen Psychiatry, 61*(1), 73-84.

MacDonald, A. W., 3rd, Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science, 288*(5472), 1835-1838.

Machlin, S. R., Harris, G. J., Pearlson, G. D., Hoehn-Saric, R., Jeffery, P., & Camargo, E. E. (1991). Elevated medial-frontal cerebral blood flow in obsessive-compulsive patients: a SPECT study. *American Journal of Psychiatry, 148*(9), 1240-1242.

Macko, K. A., Jarvis, C. D., Kennedy, C., Miyaoka, M., Shinohara, M., Sololoff, L., & Mishkin, M. (1982). Mapping the primate visual system with [2-14C]deoxyglucose. *Science, 218*(4570), 394-397.

Malizia, A. L. (1999). What do brain imaging studies tell us about anxiety disorders? *J Psychopharmacol, 13*(4), 372-378.

Malizia, A. L., Cunningham, V. J., Bell, C. J., Liddle, P. F., Jones, T., & Nutt, D. J. (1998). Decreased brain GABA(A)-benzodiazepine receptor binding in panic disorder: preliminary results from a quantitative PET study. *Arch Gen Psychiatry, 55*(8), 715-720.

Manoach, D. S., Halpern, E. F., Kramer, T. S., Chang, Y., Goff, D. C., Rauch, S. L., Kennedy, D. N., & Gollub, R. L. (2001). Test-retest reliability of a functional MRI working memory paradigm in normal and schizophrenic subjects. *Am J Psychiatry, 158*(6), 955-958.

Maruff, P., Purcell, R., Tyler, P., Pantelis, C., & Currie, J. (1999). Abnormalities of internally generated saccades in obsessive-compulsive disorder. *Psychol Med, 29*(6), 1377-1385.

Mataix-Cols, D., Wooderson, S., Lawrence, N., Brammer, M. J., Speckens, A., & Phillips, M. L. (2004). Distinct neural correlates of washing, checking, and hoarding symptom dimensions in obsessive-compulsive disorder. *Arch Gen Psychiatry, 61*(6), 564-576.

Maunsell, J. H. R. (2004). Neuronal representations of cognitive state: reward or attention? *Trends in Cognitive Sciences, 8*(6), 261-265.

McClure, S. M., Berns, G. S., & Montague, P. R. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron, 38*(2), 339-346.

Mellers, B., Schwartz, A., & Ritov, I. (1999). Emotion-Based Choice. *Journal of Experimental Psychology: General, 128*(3), 332-345.

Mellers, B. A. (2000). Choice and the relative pleasure of consequences. *Psychol Bull, 126*(6), 910-924.

Mellers, B. A. (2001). Anticipated Emotions as Guides to Choice. *Current Directions in Psychol Sci, 10*(6), 210-214.

Mellers, B. A., Schwartz, A., Ho, K., & Ritov, I. (1997). Decision affect theory: Emotional reactions to the outcomes of risky options. *Psychological Science, 8*, 423-429.

Menon, V., Adleman, N. E., White, C. D., Glover, G. H., & Reiss, A. L. (2001). Error-related brain activation during a Go/NoGo response inhibition task. *Human Brain Mapping, 12*(3), 131-143.

Meunier, M., Bachevalier, J., & Mishkin, M. (1997). Effects of orbital frontal and anterior cingulate lesions on object and spatial memory in rhesus monkeys. *Neuropsychologia, 35*(7), 999-1015.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*(1), 167-202.

Milner, B. (1964). Some effects of frontal lobectomy in man. In J. M. Warren & K. Akert (Eds.), *The frontal granular cortex and behavior* (pp. 313-334). New York: McGraw-Hill.

Mishkin, M. (1964). Perseveration of central sets after frontal lesions in man. In J. M. Warren & K. Akert (Eds.), *The frontal granular cortex and behavior* (pp. 219-294). New York: McGraw-Hill.

Mishkin, M., & Manning, F. J. (1978). Non-spatial memory after selective prefrontal lesions in monkeys. *Brain Res, 143*(2), 313-323.

Morecraft, R. J., Geula, C., & Mesulam, M. M. (1992). Cytoarchitecture and neural afferents of orbitofrontal cortex in the brain of the monkey. *J Comp Neurol, 323*(3), 341-358.

Morris, J. S., & Dolan, R. J. (2004). Dissociable amygdala and orbitofrontal responses during reversal fear conditioning. *NeuroImage, 22*(1), 372-380.

Nakano, K. (2000). Neural circuits and topographic organization of the basal ganglia and related regions. *Brain & Development, 22 Suppl 1*, S5-16.

Noll, D. C., Cohen, J. D., Meyer, C. H., & Schneider, W. (1995). Spiral K-space MR imaging of cortical activation. *Journal of Magnetic Resonance Imaging, 5*(1), 49-56.

Noll, D. C., Genovese, C. R., Nystrom, L. E., Vazquez, A. L., Forman, S. D., Eddy, W. F., & Cohen, J. D. (1997). Estimating test-retest reliability in functional MR imaging. II: Application to motor and cognitive activation studies. *Magnetic Resonance in Medicine, 38*(3), 508-517.

Nordahl, T. E., Benkelfat, C., Semple, W. E., Gross, M., King, A. C., & Cohen, R. M. (1989). Cerebral glucose metabolic rates in obsessive compulsive disorder. *Neuropsychopharmacology, 2*(1), 23-28.

O'Doherty, J., Critchley, H., Deichmann, R., & Dolan, R. J. (2003). Dissociating Valence of Outcome from Behavioral Control in Human Orbital and Ventral Prefrontal Cortices. *J. Neurosci., 23*(21), 7931-7939.

O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat Neurosci, 4*(1), 95-102.

O'Doherty, J., Rolls, E. T., Francis, S., Bowtell, R., McGlone, F., Kobal, G., Renner, B., & Ahne, G. (2000). Sensory-specific satiety-related olfactory activation of the human orbitofrontal cortex. *Neuroreport, 11*(4), 893-897.

O'Doherty, J. P., Deichmann, R., Critchley, H. D., & Dolan, R. J. (2002). Neural responses during anticipation of a primary taste reward. *Neuron, 33*(5), 815-826.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113.

Ongur, D., & Price, J. L. (2000). The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb Cortex, 10*(3), 206-219.

O'Reilly, R. C., Noelle, D. C., Braver, T. S., & Cohen, J. D. (2002). Prefrontal cortex and dynamic categorization tasks: representational organization and neuromodulatory control. *Cereb Cortex, 12*(3), 246-257.

Ownby, R. L. (1998). Computational model of obsessive-compulsive disorder: examination of etiologic hypothesis and treatment strategies. *Depress Anxiety, 8*(3), 91-103.

Pagnoni, G., Zink, C. F., Montague, P. R., & Berns, G. S. (2002). Activity in human ventral striatum locked to errors of reward prediction. *Nat Neurosci, 5*(2), 97-98.

Perani, D., Colombo, C., Bressi, S., Bonfanti, A., Grassi, F., Scarone, S., Bellodi, L., Smeraldi, E., & Fazio, F. (1995). [18F]FDG PET study in obsessive-compulsive disorder. A clinical/metabolic correlation study after treatment. *British Journal of Psychiatry, 166*(2), 244-250.

Phillips, M. L., Drevets, W. C., Rauch, S. L., & Lane, R. (2003). Neurobiology of emotion perception II: Implications for major psychiatric disorders. *Biol Psychiatry, 54*(5), 515-528.

Pitman, R. K. (1987). A cybernetic model of obsessive-compulsive psychopathology. *Comprehensive Psychiatry, 28*(4), 334-343.

Pochon, J. B., Levy, R., Fossati, P., Lehericy, S., Poline, J. B., Pillon, B., Le Bihan, D., & Dubois, B. (2002). The neural system that bridges reward and cognition in humans: an fMRI study. *Proc Natl Acad Sci U S A, 99*(8), 5669-5674.

Porrino, L. J., Crane, A. M., & Goldman-Rakic, P. S. (1981). Direct and indirect pathways from the amygdala to the frontal lobe in rhesus monkeys. *J Comp Neurol, 198*(1), 121-136.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America, 98*(2), 676-682.

Ramanaiah, N. V., Franzen, M., & Schill, T. (1983). A psychometric study of the State-Trait Anxiety Inventory. *J Pers Assess, 47*(5), 531-535.

Ramnani, N., & Miall, R. C. (2003). Instructed delay activity in the human prefrontal cortex is modulated by monetary reward expectation. *Cereb Cortex, 13*(3), 318-327.

Rapoport, J. L. (1991). Recent advances in obsessive-compulsive disorder [see comments]. *Neuropsychopharmacology, 5*(1), 1-10.

Rauch, S. L. (2000). Neuroimaging research and the neurobiology of obsessive-compulsive disorder: where do we go from here? *Biol Psychiatry, 47*(3), 168-170.

Rauch, S. L., Jenike, M. A., Alpert, N. M., Baer, L., Breiter, H. C., Savage, C. R., & Fischman, A. J. (1994). Regional cerebral blood flow measured during symptom provocation in obsessive-compulsive disorder using oxygen 15-labeled carbon dioxide and positron emission tomography [see comments]. *Archives of General Psychiatry, 51*(1), 62-70.

Rauch, S. L., Savage, C. R., Alpert, N. M., Fischman, A. J., & Jenike, M. A. (1997). The functional neuroanatomy of anxiety: a study of three disorders using positron emission tomography and symptom provocation. *Biological Psychiatry, 42*(6), 446-452.

Rauch, S. L., Savage, C. R., Alpert, N. M., Miguel, E. C., Baer, L., Breiter, H. C., Fischman, A. J., Manzo, P. A., Moretti, C., & Jenike, M. A. (1995). A positron emission tomographic study of simple phobic symptom provocation. *Archives of General Psychiatry, 52*(1), 20-28.

Rauch, S. L., Shin, L. M., Dougherty, D. D., Alpert, N. M., Fischman, A. J., & Jenike, M. A. (2002). Predictors of Fluvoxamine Response in Contamination-related Obsessive Compulsive Disorder. A PET Symptom Provocation Study. *Neuropsychopharmacology, 27*(5), 782-791.

Rauch, S. L., Whalen, P. J., Curran, T., Shin, L. M., Coffey, B. J., Savage, C. R., McInerney, S. C., Baer, L., & Jenike, M. A. (2001). Probing striato-thalamic function in obsessive-compulsive disorder and Tourette syndrome using neuroimaging methods. *Advances in Neurology, 85*, 207-224.

Ray, J. P., & Price, J. L. (1993). The organization of projections from the mediodorsal nucleus of the thalamus to orbital and medial prefrontal cortex in macaque monkeys. *J Comp Neurol, 337*(1), 1-31.

Ridderinkhof, K. R. (2002). Micro- and macro-adjustments of task set: activation and suppression in conflict tasks. *Psychol Res, 66*(4), 312-323.

Roberts, A. C., Robbins, T. W., & L., W. (Eds.). (1998). *The prefrontal cortex: executive and cognitive functions*. New York: Oxford University Press.

Roberts, A. C., & Wallis, J. D. (2000). Inhibitory control and affective processing in the prefrontal cortex: Neuropsychological studies in the common marmoset. *Cerebral Cortex, Vol 10*(3), 252-262.

Robins, L. N., Helzer, J. E., Weissman, M. M., Orvaschel, H., Gruenberg, E., Burke, J. D., Jr., & Regier, D. A. (1984). Lifetime prevalence of specific psychiatric disorders in three sites. *Archives of General Psychiatry, 41*(10), 949-958.

Roesch, M. R., & Olson, C. R. (2003). Impact of expected reward on neuronal activity in prefrontal cortex, frontal and supplementary eye fields and premotor cortex. *J Neurophysiol, 90*(3), 1766-1789.

Roesch, M. R., & Olson, C. R. (2004). Neuronal Activity Related to Reward Value and Motivation in Primate Frontal Cortex. *Science, 304*(5668), 307-310.

Rogers, R. D., Owen, A. M., Middleton, H. C., Williams, E. J., Pickard, J. D., Sahakian, B. J., & Robbins, T. W. (1999). Choosing between small, likely rewards and large, unlikely rewards activates inferior and orbital prefrontal cortex. *J Neurosci, 19*(20), 9029-9038.

Rogers, R. D., Ramnani, N., Mackay, C., Wilson, J. L., Jezzard, P., Carter, C. S., & Smith, S. M. (2004). Distinct portions of anterior cingulate cortex and medial prefrontal cortex are activated by reward processing in separable phases of decision-making cognition. *Biological Psychiatry, 55*(6), 594-602.

Rolls, E. T. (1998). The orbitofrontal cortex. In A. C. Roberts & T. W. Robbins & W. L. (Eds.), *The prefrontal cortex: executive and cogitive functions* (pp. 67-86). New York: Oxford University Press.

Rolls, E. T. (2000). The orbitofrontal cortex and reward. *Cerebral Cortex, 10*(3), 284-294.

Rolls, E. T. (2004). The functions of the orbitofrontal cortex. *Brain Cogn, 55*(1), 11-29.

Rolls, E. T., Critchley, H. D., Mason, R., & Wakeman, E. A. (1996). Orbitofrontal cortex neurons: Role in olfactory and visual association learning. *Journal of Neurophysiology, 75*(5), 1970-1981.

Rolls, E. T., Hornak, J., Wade, D., & McGrath, J. (1994). Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *Journal of Neurology, Neurosurgery and Psychiatry, 57*(12), 1518-1524.

Rolls, E. T., O'Doherty, J., Kringelbach, M. L., Francis, S., Bowtell, R., & McGlone, F. (2003). Representations of Pleasant and Painful Touch in the Human Orbitofrontal and Cingulate Cortices. *Cereb Cortex, 13*(3), 308-317.

Rolls, E. T., Yaxley, S., & Sienkiewicz, Z. J. (1990). Gustatory responses of single neurons in the caudolateral orbitofrontal cortex of the macaque monkey. *Journal of Neurophysiology, 64*(4), 1055-1066.

Rosenberg, D. R., Benazon, N. R., Gilbert, A., Sullivan, A., & Moore, G. J. (2000). Thalamic volume in pediatric obsessive-compulsive disorder patients before and after cognitive behavioral therapy. *Biological Psychiatry, 48*(4), 294-300.

Rosenberg, D. R., Dick, E. L., O'Hearn, K. M., & Sweeney, J. A. (1997). Response-inhibition deficits in obsessive-compulsive disorder: an indicator of dysfunction in frontostriatal circuits. *J Psychiatry Neurosci, 22*(1), 29-38.

Rosenkilde, C. E., Bauer, R. H., & Fuster, J. M. (1981). Single cell activity in ventral prefrontal cortex of behaving monkeys. *Brain Research, 209*(2), 375-394.

Rubia, K., Russell, T., Overmeyer, S., Brammer, M. J., Bullmore, E. T., Sharma, T., Simmons, A., Williams, S. C., Giampietro, V., Andrew, C. M., & Taylor, E. (2001). Mapping motor inhibition: conjunctive brain activations across different versions of go/no-go and stop tasks. *Neuroimage, 13*(2), 250-261.

Sawle, G. V., Hymas, N. F., Lees, A. J., & Frackowiak, R. S. (1991). Obsessional slowness. Functional studies with positron emission tomography. *Brain, 114*(Pt 5), 2191-2202.

Saxena, S., Brody, A. L., Ho, M. L., Alborzian, S., Maidment, K. M., Zohrabi, N., Ho, M. K., Huang, S. C., Wu, H. M., & Baxter, L. R., Jr. (2002). Differential cerebral metabolic changes with paroxetine treatment of obsessive-compulsive disorder vs major depression. *Arch Gen Psychiatry, 59*(3), 250-261.

Saxena, S., Brody, A. L., Ho, M. L., Zohrabi, N., Maidment, K. M., & Baxter, L. R., Jr. (2003). Differential brain metabolic predictors of response to paroxetine in obsessive-compulsive disorder versus major depression. *Am J Psychiatry, 160*(3), 522-532.

Saxena, S., Brody, A. L., Maidment, K. M., Dunkin, J. J., Colgan, M., Alborzian, S., Phelps, M. E., & Baxter, L. R., Jr. (1999). Localized orbitofrontal and subcortical metabolic changes and predictors of response to paroxetine treatment in obsessive-compulsive disorder. *Neuropsychopharmacology, 21*(6), 683-693.

Saxena, S., Brody, A. L., Schwartz, J. M., & Baxter, L. R. (1998). Neuroimaging and frontal-subcortical circuitry in obsessive-compulsive disorder. *British Journal of Psychiatry - Supplement*(35), 26-37.

Schultz, W., Tremblay, L., & Hollerman, J. R. (2000). Reward processing in primate orbitofrontal cortex and basal ganglia. *Cerebral Cortex, 10*(3), 272-284.

Shulman, G. L., Fiez, J. A., Corbetta, M., Buckner, R. L., Miezin, F. M., Raichle, M. E., & Petersen, S. E. (1997). Common blood flow changes across visual tasks: II.: Decreases in cerebral cortex. *Journal of Cognitive Neuroscience, 9*(5), 648-663.

Small, D. M., Zatorre, R. J., Dagher, A., Evans, A. C., & Jones Gotman, M. (2001). Changes in brain activity related to eating chocolate: from pleasure to aversion. *Brain, 124*(Pt 9), 1720-1733.

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

Sohn, M.-H., Ursu, S., Anderson, J. R., Stenger, V. A., & Carter, C. S. (2000). The role of prefrontal cortex and posterior parietal cortex in task switching. *Proceedings of the National Academy of Science of the United States of America, 97*(24), 13448-13453.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1980). Manual for the state-trait anxiety inventory. *Palo Alto: Consulting Psychology Press*.

Stein, D. J., Van Heerden, B., Wessels, C. J., Van Kradenburg, J., Warwick, J., & Wasserman, H. J. (1999). Single photon emission computed tomography of the brain with Tc-99m HMPAO during sumatriptan challenge in obsessive-compulsive disorder: investigating the functional role of the serotonin auto-receptor. *Progress in Neuro-Psychopharmacology & Biological Psychiatry, 23*(6), 1079-1099.

Swedo, S. E., Pietrini, P., Leonard, H. L., Schapiro, M. B., Rettew, D. C., Goldberger, E. L., Rapoport, S. I., Rapoport, J. L., & Grady, C. L. (1992). Cerebral glucose metabolism in childhood-onset obsessive-compulsive disorder. Revisualization during pharmacotherapy. *Archives of General Psychiatry, 49*(9), 690-694.

Swedo, S. E., Schapiro, M. B., Grady, C. L., Cheslow, D. L., Leonard, H. L., Kumar, A., Friedland, R., Rapoport, S. I., & Rapoport, J. L. (1989). Cerebral glucose metabolism in

childhood-onset obsessive-compulsive disorder. *Archives of General Psychiatry, 46*(6), 518-523.

Szeszko, P. R., MacMillan, S., McMeniman, M., Chen, S., Baribault, K., Lim, K. O., Ivey, J., Rose, M., Banerjee, S. P., Bhandari, R., Moore, G. J., & Rosenberg, D. R. (2004). Brain Structural Abnormalities in Psychotropic Drug-Naive Pediatric Patients With Obsessive-Compulsive Disorder. *Am J Psychiatry, 161*(6), 1049-1056.

Tanabe, T., Iino, M., & Takagi, S. F. (1975). Discrimination of odors in olfactory bulb, pyriform-amygdaloid areas, and orbitofrontal cortex of the monkey. *J Neurophysiol, 38*(5), 1284-1296.

Tashiro, M., Itoh, M., Kubota, K., Kumano, H., Masud, M. M., Moser, E., Arai, H., & Sasaki, H. (2001). Relationship between trait anxiety, brain activity and natural killer cell activity in cancer patients: a preliminary PET study. *Psychooncology, 10*(6), 541-546.

Taylor, S. F., Welsh, R. C., Wager, T. D., Luan Phan, K., Fitzgerald, K. D., & Gehring, W. J. (2004). A functional neuroimaging study of motivation and executive function. *Neuroimage, 21*(3), 1045-1054.

Thut, G., Schultz, W., Roelcke, U., Nienhusmeier, M., Missimer, J., Maguire, R. P., & Leenders, K. L. (1997). Activation of the human brain by monetary reward. *Neuroreport, 8*(5), 1225-1228.

Tracy, A. L., Jarrard, L. E., & Davidson, T. L. (2001). The hippocampus and motivation revisited: appetite and activity. *Behav Brain Res, 127*(1-2), 13-23.

Tremblay, L., & Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. *Nature, 398*(6729), 704-708.

Tremblay, L., & Schultz, W. (2000a). Modifications of reward expectation-related neuronal activity during learning in primate orbitofrontal cortex. *Journal of Neurophysiology, 83*(4), 1877-1885.

Tremblay, L., & Schultz, W. (2000b). Reward-related neuronal activity during go-nogo task performance in primate orbitofrontal cortex. *Journal of Neurophysiology, 83*(4), 1864-1876.

Ursu, S., Stenger, V. A., Shear, M. K., Jones, M. R., & Carter, C. S. (2003). Overactive action monitoring in obsessive-compulsive disorder: evidence from functional magnetic resonance imaging. *Psychol Sci, 14*(4), 347-353.

van der Molen, M. W., Boomsma, D. I., Jennings, J. R., & Nieuwboer, R. T. (1989). Does the heart know what the eye sees? A cardiac/pupillometric analysis of motor preparation and response execution. *Psychophysiology, 26*(1), 70-80.

Vasa, R. A., Grados, M., Slomine, B., Herskovits, E. H., Thompson, R. E., Salorio, C., Christensen, J., Wursta, C., Riddle, M. A., & Gerring, J. P. (2004). Neuroimaging correlates of anxiety after pediatric traumatic brain injury. *Biol Psychiatry, 55*(3), 208-216.

Visscher, K. M., Miezin, F. M., Kelly, J. E., Buckner, R. L., Donaldson, D. I., McAvoy, M. P., Bhalodia, V. M., & Petersen, S. E. (2003). Mixed blocked/event-related designs separate transient and sustained activity in fMRI. *Neuroimage, 19*(4), 1694-1708.

Volkow, N. D., Chang, L., Wang, G. J., Fowler, J. S., Ding, Y. S., Sedler, M., Logan, J., Franceschi, D., Gatley, J., Hitzemann, R., Gifford, A., Wong, C., & Pappas, N. (2001). Low level of brain dopamine D2 receptors in methamphetamine abusers: association with metabolism in the orbitofrontal cortex. *Am J Psychiatry, 158*(12), 2015-2021.

Waldvogel, D., van Gelderen, P., Muellbacher, W., Ziemann, U., Immisch, I., & Hallett, M. (2000). The relative metabolic demand of inhibition and excitation. *Nature, 406*(6799), 995-998.

Wallis, J. D., & Miller, E. K. (2003). Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur J Neurosci, 18*(7), 2069-2081.

Watanabe, M., Hikosaka, K., Sakagami, M., & Shirakawa, S. (2002). Coding and monitoring of motivational context in the primate prefrontal cortex. *J Neurosci, 22*(6), 2391-2400.

Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., & Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience, 18*(1), 411-418.

Woods, R. P., Cherry, S. R., & Mazziotta, J. C. (1992). Rapid automated algorithm for aligning and reslicing PET images. *Journal of Computer Assisted Tomography, 16*(4), 620-633.

Zald, D. H., & Kim, S. W. (1996). Anatomy and function of the orbital frontal cortex, I: anatomy, neurocircuitry; and obsessive-compulsive disorder. *J Neuropsychiatry Clin Neurosci, 8*(2), 125-138.

Zink, C. F., Pagnoni, G., Martin, M. E., Dhamala, M., & Berns, G. S. (2003). Human Striatal Response to Salient Nonrewarding Stimuli. *J. Neurosci., 23*(22), 8092-8097.

Zink, C. F., Pagnoni, G., Martin-Skurski, M. E., Chappelow, J. C., & Berns, G. S. (2004). Human Striatal Responses to Monetary Reward Depend On Saliency. *Neuron, 42*(3), 509-517.

APPENDIX A

Spiral-in acquisition protocol

In regular, spiral-out fMRI protocols, signal is acquired at a certain time (echo time, TE) after the slice-select RF pulse (i.e. the "excitation" pulse). This ensures a maximal difference between the signal from brain areas relatively richer in oxyhemoglobin (HbO2, presumably a result of increased flow of HbO2-rich blood in active areas, in excess of their metabolic needs) relative to relatively richer in deoxyhemoglobin (Hb). The "weaker" signal in the Hb-rich areas (i.e. "inactive" areas) is a result of local field inhomogeneities induced in the magnetic field by the paramagnetic properties of Hb. These inhomogeneities result in faster dephasing of protonic spins, which ultimately translates into weaker MR signal. Magnetic inhomogeneities are also caused by the close interface between the brain and other media with different magnetic properties, such as the air in the sinus cavities in the base of the skull. Based on the same principle described above, the signal in ventral areas of the brain is often decreased by these magnetic field inhomogeneities. One way of avoiding signal loss is to start acquiring the MR signal with a shorter echo time, before the inhomogeneities have had a chance to induce complete spin dephasing. The fMRI spiral-in acquisition protocols do just that, by sampling the edges of the k-space (the frequency-phase space that defines any MR signal) sooner after the excitation pulse (see Figure 19).

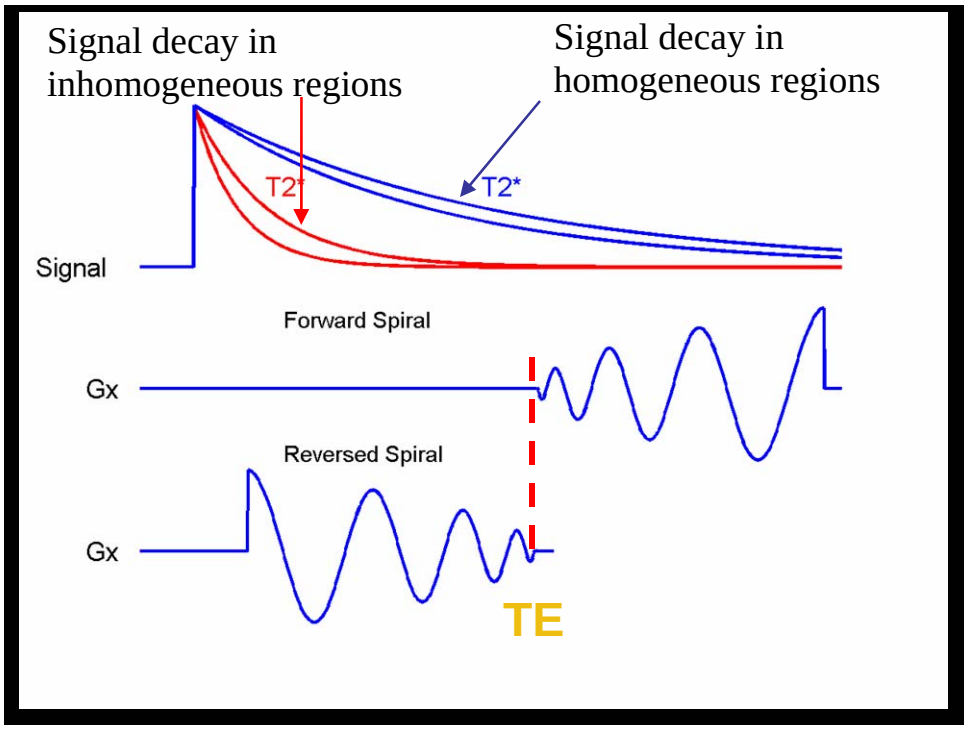

Figure 20. Principles of spiral-in (or "reverse") fMRI signal acquisition

The contrast of the fMR images depends mainly on the difference in rate of dephasing of spins between HbO2-"rich" and HbO2-"depleted" voxels (the slower decaying curve vs. the faster decaying one in the signal graphs). In a homogeneous magnetic field, both curves decay overall slower (blue curves) than in regions with magnetic field inhomogeneities (red curves). Therefore, in order to preserve the contrast generated by relative differences in HbO2, the spiral-in protocol acquires part of the data before it has a chance to decay in the susceptible regions (red curves of signal decay).

APPENDIX B

Analysis of response inhibition in Experiment 1

To parallel Waldvogel et al.'s analysis, we took advantage of the fact that in our task the targets contained the sample stimulus on the right or on the left. Therefore, for targets which contained the sample stimulus on the right side, correct Match responses were in effect Go responses, since they were automatic response with the right hand. For the same type of targets (i.e. with the sample on the right), Non-match responses were effectively No-go trials with respect to the right hand, since they required suppression of the automatic response with that hand. Comparing the activity to Match vs. Non-match responses to right-sided targets revealed that BOLD responses in the pre-supplementary motor areas (pre-SMA) were the same during execution and during suppression of responses of the dominant hand. In contrast, in the left motor cortex the signal changes were significantly weaker during suppression of automatic responses of the right (dominant) hand (see Figure 7).

These two types of trials were contrasted during the 8 scans following the response to targets (S9-S16). As per Waldvogel et al., pre-SMA and left motor cortex regions of interest were selected by identifying voxels with a significant main effect of scan, Bonferroni corrected for the number of voxels in the motor cortex and pre-SMA. Then a Condition (Non-match vs Match) by Scan (S9-S16) ANOVA was conducted on the average signal changes (relative to scan S9). In the pre-SMA there was no significant difference between Non-match and Match targets (Task x Scan interaction: $F(7, 126) = 1.12$, $p = 0.35$). In contrast, in the left motor cortex, the BOLD response was significantly reduced after Non-match compared to Match targets,

consistent with the suppression, during these Non-match targets, of the automatic right hand response [Task x Scan $F(7, 126) = 3.18$, $p = 0.004$].

APPENDIX C

Effects of Incentives on heart rate in Experiment 2

The analysis was conducted on a subgroup of 10 subjects for whom we obtained online recordings of heart rate (HR) for the entire duration of the experiment. We computed, within each incentive level, the average heart rate in each second of the 24 second interval between two consecutive samples stimuli. These averages are displayed in Figure 21. We looked for indices of differential arousal under each incentive level by selecting the average HR during the 10^{th} and 12^{th} second of the interval (the third to last and last second before probe onset, marked by the shaded area in the figure). This interval was selected based on previous research showing that the heart decelerations in preparation for performing a cognitive task occur over the last 3-4 beats before the anticipated onset of the stimulus that subjects have to respond to (Jennings, van der Molen, & Somsen, 1998; Jennings, van der Molen, & Steinhauer, 1998; van der Molen et al., 1989). At a heart rate of approximately 70 beats/min, this translates into an interval of approximately 3 seconds. For the selected timepoints, a two way Incentive (Reward vs. Penalty vs. No-Reward) by Timepoint (10^{th} vs. 12^{th} second) random effects ANOVA was conducted, with mean heart rate over each second as dependent variable. Overall, the ANOVA did not reach significance [$F(2, 18) = 1.25$, $p = 0.13$]. However, three other 2-way ANOVAs contrasting each pair of 2 levels of incentives showed trends for a bigger decrease in HR during preparation for Reward trials relative to either Penalty [$F(1, 9) = 2.44$, $p = 0.15$] or No-reward trials [$F(1, 9) = 3.19$, $p = 0.11$], while the decreases during Penalty and No-reward trials were highly non-significant [$F(1, 9) = 0.12$, $p = 0.74$].

We also computed the signal changes in this subgroup of 10 subjects for the OFC areas which showed effects of incentives in the entire sample of 17 subjects. As can be seen in Figure 22, the patterns of activation of the 10 subjects were very similar to those of the larger group, verifying that these subjects were representative of the larger sample in terms of incentive effects in the OFC.

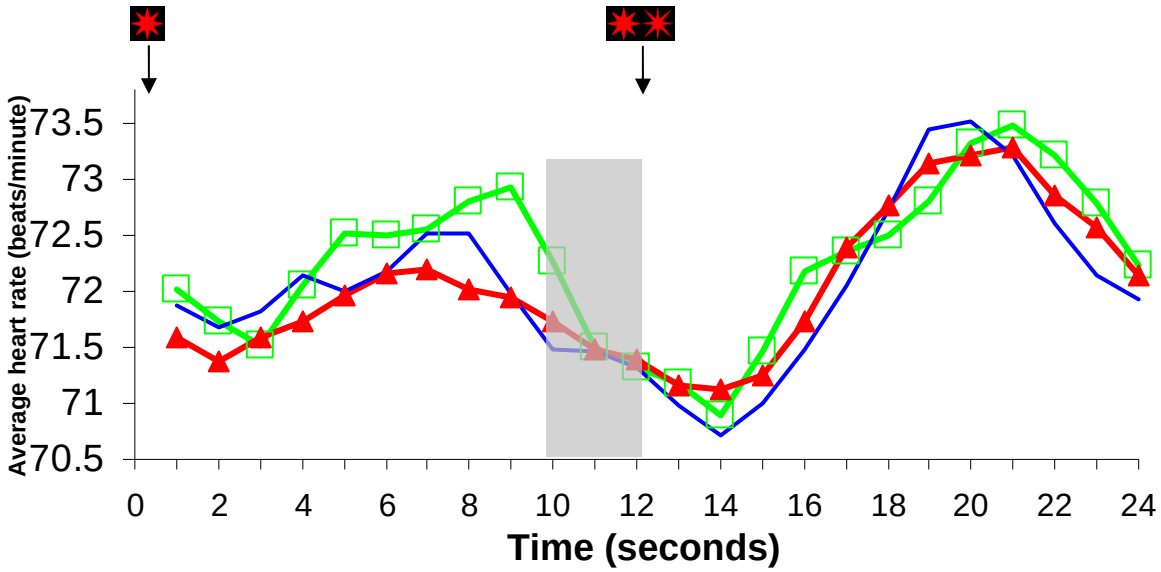

Figure 21. Effects of incentives on average heart rates in a subgroup of ten subjects

Grand averages of the heart rates are calculated for each second of each type of trial: Reward (green/square markers), Penalty (red/triangles), No-reward (blue/no marker). The shaded box outlines the interval over which the differences in HR deceleration has been tested via Incentive by Time ANOVAS, and showed trends for grater decelaration in anticipation of targets under the Reward condition relative to either Penalty or No-reward.

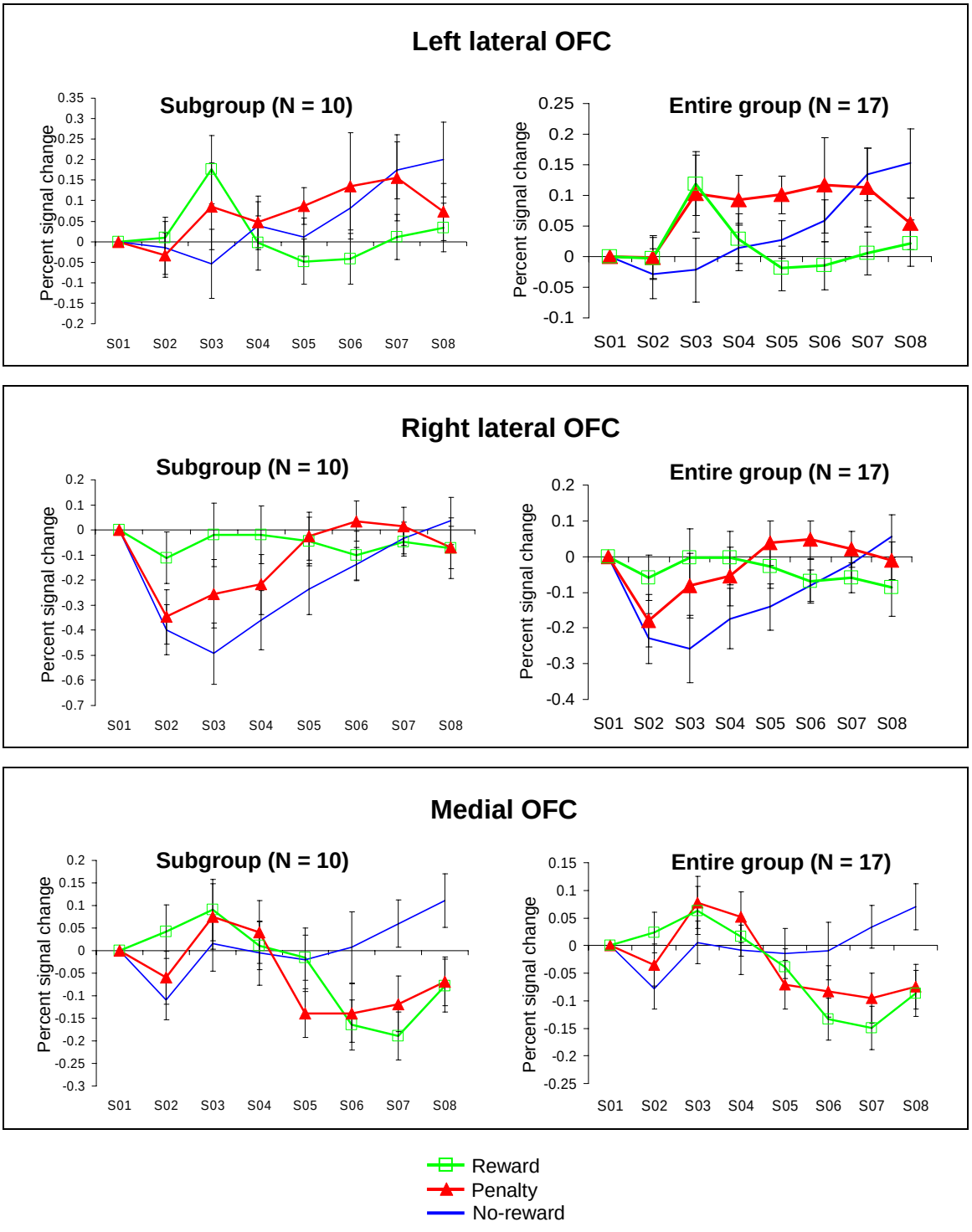

Figure 22. Incentive effects in the OFC of subjects with heart rate recordings during scanning

The figure depicts the patterns of signal change in the ten subjects in whom complete heart rate data were collected during scanning, for the OFC regions of interest which showed effects of incentives in the large group (see Figure 13).

APPENDIX D

Spatial relationship between the peak of OFC hyperactivity in OCD and maxima of Task and Incentive effects in Experiments 1 and 2

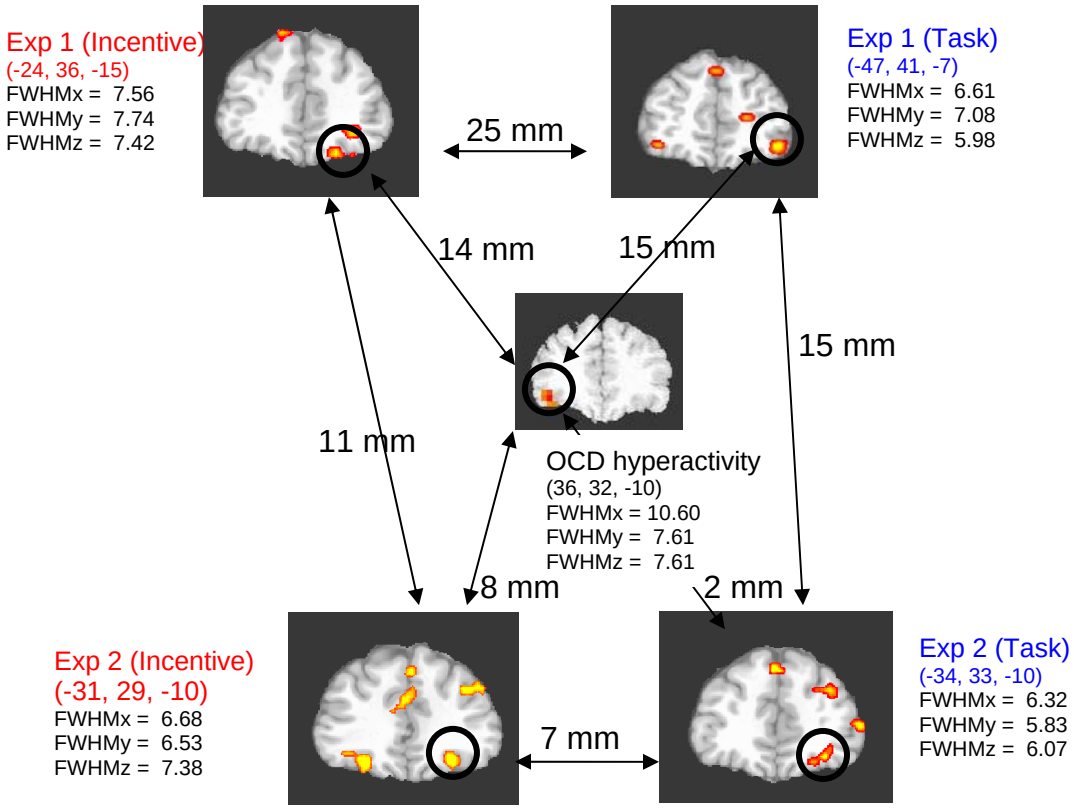

Figure 23. Peak to peak distances between main foci of activation in the three experiments

Distances between peaks of effects in the Talairach system of coordinates were computed for each pair of ROIs of interest. For any given pair or peak voxels A and B, given Talairach coordinates of the form (x_a, y_a, z_a) and (x_b, y_b, z_b), the direct distance between peaks was given by the formula: Sqrt [($|x_a|$ - $|x_b|$)^2 + ($|y_a|$ - $|y_b|$)^2 + ($|z_a|$ - $|z_b|$)^2], where $|x_a|$ is the absolute value of the x coordinate of peak A. This formula computed the peak-to-peak distances as if they were located on the same hemisphere. The smoothing of the statistical maps in which the ROIs were isolated was estimated using the 3dFWHM function of the AFNI software package (Cox, 1996).

The central figure represents the focus of hyperactivity in OCD patients, and the four other maps illustrate the two types of effects (incentives and task) in the first two experiments. For the Task effect in Experiment 1, the comparison was done with the most ventral ROI (the inferior convexity, BA 10/47) that was activated more during preparation for Non-match relative to Match. Overall, the distances to peaks of activations predicted by manipulations of outcomes (Incentive in Experiments 1 and 2, OFC Task effects in Experiment 2) are comparable to the size of the estimated smoothing. In contrast, the distance between the focus of OCD hyperactivity and the control-related activation (Task effect, Exp. 1) is greater than the spatial smoothing of the statistical maps.