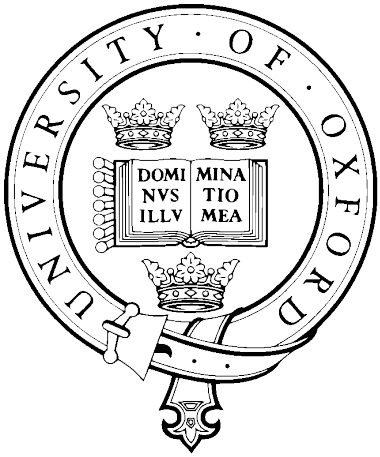
Erica A. Boschin
The Queen's College

DYNAMICS OF COGNITIVE CONTROL AND FLEXIBILITY IN THE ANTERIOR CINGULATE AND PREFRONTAL CORTICES



Thesis submitted for the
Degree of Doctor of Philosophy

Michaelmas Term
2013

TABLE OF CONTENTS

ABSTRACT **3**

DETAILED ABSTRACT **4**

ACKNOWLEDGMENTS **10**

GENERAL INTRODUCTION **14**

Theoretical and Behavioural Background 14
Neural Substrates of Cognitive Control and Behavioural Flexibility 30

ACC *38*
dlPFC *48*
FPC *57*

Scope of the Thesis 66

CHAPTER 1: Human Neuropsychological Investigations **71**

CHAPTER 2: Transcranial Magnetic Stimulation of the dlPFC **97**

CHAPTER 3: Adaptation Effects Comparisons Across Tasks **123**

CHAPTER 4: Application of the Linear Ballistic Accumulator Model **167**

CHAPTER 5: Deep Brain Stimulation and Intracranial Recordings in the ACC **201**

CHAPTER 6: Non-human Primate Neuropsychological Investigations **239**

GENERAL DISCUSSION **265**

Summary and Implications of the Behavioural and Modeling Findings 265
Summary and Implications of the Investigation into the Neural Substrates 279
Limitations of the Current Research 306
Final Comments 310

REFERENCES **312**

APPENDIX **341**

Erica Boschin
The Queen's College

DPhil Thesis
Michaelmas Term 2013

Dynamics of Cognitive Control and Flexibility in the Anterior Cingulate and Prefrontal Cortices

ABSTRACT

The body of work hereby presented aims at better defining the specific mechanisms underlying cognitive control and flexibility, and to investigate the neural substrates that might support these dynamics. More specifically, the anterior cingulate (ACC), dorsolateral prefrontal (dlPFC) and frontopolar (FPC) cortices have been proposed to play a fundamental role in monitoring and detecting the presence of environmental contingencies that require the recruitment of cognitive control (such as competition between responses in the presence of conflicting information), implementing cognitive control, and supporting higher-order cognitive processing, respectively. This thesis investigates the effects of damage to these regions, and of interference with their activity, on these processes. It also argues for the importance of dissociating possible separate cognitive control components that might differently contribute to behavioural adjustments (such as caution and attention/task-relevant processing), and provides one of the first attempts to quantify them within the parameters of a mathematical model of choice response-time, the Linear Ballistic Accumulator (LBA). The results confirm the crucial role of the dlPFC in modulating behavioural adjustments, as both damage and interference with this region's activity significantly affect measures of conflict-induced behavioural adaptation. It is hypothesized that dlPFC might drive behavioural adjustments by encoding recent conflict history and/or supporting the automatization of a newly advantageous behavioural strategy during the early stages after a change in conflict levels. When a task does not involve competition between a habit and instructed behaviour, lesions or interference with ACC's activity do not appear to affect behaviour in a manner that is consistent with the classic conflict-monitoring framework. It is suggested that its role might be better described as a more general monitoring and confirmatory mechanism that evaluates both actual and potential outcomes of an action, in order to proactively guide adjustments away from contextually disadvantageous responses. Finally, lesions to the FPC do not affect abstract-rule integration, but do impair the early stages of acquisition of a new abstract rule, when a previously rewarded rule stops being rewarded, and specifically when acquisition is dependent on self-initiated exploration. This suggests a role for FPC in the evaluation of multiple concurrent options in order to aid the development of new behavioural strategies.

DETAILED ABSTRACT

Cognitive control and flexibility are crucial to the cognitive system's ability to maximize the outcome of actions, particularly in response to changes in context or environmental contingencies. The dynamics of these processes are often investigated as part of research into conflict-monitoring and rule-guided behaviour. Findings from this line of research have indicated the anterior cingulate (ACC), dorsolateral prefrontal (dlPFC) and frontopolar (FPC) cortices as regions that play a fundamental role in the modulation of control and flexibility, particularly within contexts involving competition between different options. However, the specific mechanisms that cognitive control implements to support behavioural adjustments, and therefore the roles of each of these areas in supporting specific aspects of cognitive control, are often unspecified. Contradictory findings have also emerged in the literature, particularly with regards to the role of the ACC in conflict-monitoring and in allocating functions to the FPC.

Chapter 1 looks at the effects of ACC and dlPFC damage in human neuropsychological patients during the performance of a version of the Wisconsin Card Sorting test that has previously been adapted to investigate the role of these regions in conflict-monitoring in the monkey. Differently from classic conflict tasks, which elicit competition between a habit and an instructed response, the WCST is a task that elicits competition between two abstract rules ('match by colour' vs. 'match by shape') that can be applied in order to perform the correct response. When healthy participants perform this task, speed of response on a high-conflict trial that follows another high

conflict trial (HH) is higher than on a high-conflict trial that follows a low-conflict trials (LH). This is an index of conflict-induced behavioural adaptation. Results show that, consistent with non-human primate findings, but not with the conflict-monitoring framework, ACC does not appear necessary for optimal performance on this task: patients with ACC lesions respond with speed comparable to their controls, both on high-conflict trials overall and on adaptation trials. Patients with dlPFC lesions, on the other hand, are found to be impaired in measures of conflict-induced behavioural adaptation, showing an abolition of the conflict-induced adjustments (i.e. the difference in speed of response on HH vs. LH trials) that are observed in healthy controls. This is consistent with both the non-human primate findings on the role of dlPFC on this task, as well as with the conflict-monitoring framework that indicates dlPFC as a crucial structure for driving behavioural adaptation.

Chapter 2 replicates the findings of Chapter 1, applying Transcranial Magnetic Stimulation (TMS) in young healthy participants during the conflict WCST analogue. TMS of the dlPFC abolishes the adaptation effect observed on non-TMS trials. It is found that applying TMS to two different sites of dlPFC (approximately area 9/46 and dorsal area 9) seems to have slightly different modulatory effects on the specific adaptation trials that are affected. TMS to more lateral and inferior parts of dlPFC (area 9/46), affects both HH trials (with a decreases in speed of response) and LH trials (with an increase in the speed of response), although it does not appear to significantly affect performance on one type of trial over the other. TMS to more superior parts of dlFPC (area 9), appears to

affect only LH trials (with an increases in the speed of response). This suggests that dlPFC might be modulating adaptation effects through different mechanisms, and that there might be some degree of functional subdivision within this area. It is hypothesized that one of the roles of dlPFC might be to maintain a memory trace of recent conflict-history in working memory (and/or support the encoding and automatization of new, contextually appropriate response strategies), in order to drive adaptation, and that, if this cannot be formed, or is lost, the cognitive system may reset onto 'default' response options determined by overall task demands. In the case of a task that does not involve the possibility to exploit a habit to facilitate performance on low-conflict trials, an indirect route that supports rule-guided behaviour might be selected as the default, therefore facilitating responses on LH trials.

Chapter 3 illustrates a series of behavioural studies that aimed at investigating the different mechanisms underlying adaptation in the WCST, compared to more commonly used conflict tasks that are designed to elicit competition between habitual behaviours and task-specific instructed responses (such as the Simon and Flanker tasks). Results show that while adaptation effects in the WCST are restricted exclusively to differences in speed of response between HH and LH trials, with responses on low-conflict trials unaffected by the level of conflict on the previous trial, tasks that involve competition between a habit and instructed responses also (or exclusively) show reductions in the speed of response on low-conflict trials, when these follow high-conflict trials (HL trials). This might be an indication of two dissociable cognitive control mechanisms that drive

the modulation of responses on HH trials as opposed to HL trials: enhanced attention/task-relevant processing for the former, and enhanced caution or direct route suppression on the latter. It was also found that increasing the frequency of high-conflict trials in the WCST does not affect conflict effects (i.e. the difference in speed of response between high- and low-conflict trials), unlike in classic conflict tasks, where elimination (or reversal) of the conflict effect is commonly reported. This suggests that the reversal of the conflict effect might result specifically from modulation of a direct route to response that normally supports habitual behaviour, rather than by enhanced attention or task-relevant processing.

Chapter 4 describes how the proposed mechanisms that may underlie adaptation effects (and route modulation) during conflict tasks could be quantified by the parameters of the Linear Ballistic Accumulator model of choice response time. Reduction of speed on HL trials is associated with higher response-caution parameters in the model. Higher caution values are also associated with slower responses on LH trials, but only in the Simon task. Increased caution might be an index of proactive suppression of the direct route after it has been suppressed on high-conflict trials (LH), which, in turn, affects speed of response on HL trials. This was not observed in the parameter values for the best WCST model, which does not involve the exploitation of the direct route to response on any trial. However, differences in speed between HH and LH trials in the WCST cannot be clearly quantified as either caution or speed of evidence accumulation, indicating that mechanisms other than increased attentional

resources or enhanced rule-guided behaviour, via the indirect route, might underlie these adaptation effects.

In Chapter 5, the WCST and a 3-item version of the Simon task is administered to a small cohort of patients who are fitted with Deep Brain Stimulation (DBS) implants, in order to test the effects of intracranial stimulation of the ACC on performance during these two tasks, one involving conflict between a habit and instructed responses (Simon) and one only involving competition between instructed responses (WCST). Preliminary results seem to indicate that ACC stimulation does not affect speed of response on high-conflict, compared to low-conflict, trials of the WCST. This is consistent with the neuropsychological investigation described in Chapter 1, as well as the corresponding non-human primate study. Inspection of the numerical differences between high- and low-conflict trials in the two patients who completed the Simon task also shows no effect of stimulation. However, unlike in the aforementioned studies, there appears to be an effect of stimulation on adaptation in the WCST, with the adaptation effect abolished by stimulation. Local Field Potential (LFP) recordings in one patient reveal conflict-related beta oscillations in the ACC during the WCST, which appears to be anticipatory in nature and modulated by feedback-related activity on the previous trial. This indicates that the ACC might be monitoring the outcome of actions in order to proactively modulate oscillatory activity on subsequent trials.

Finally, Chapter 6 reports the result of a study investigating the effects of FPC lesion in the macaque monkey during the acquisition and application of a response strategy that involves combining two abstract rules. FPC animals were not impaired, compared to controls, when required to combine the two abstract rules, which is inconsistent with the proposed role of the FPC in rule-integration. However, FPC animals were impaired during the early stages of acquisition of the second abstract rule, once the former stopped being rewarded, and only when this was dependent on self-initiated exploration. Control animals showed a much more rapid increase in correct responses from the first to the second day of rule acquisition, indicating that an intact FPC is necessary for a quick and efficient development of a newly advantageous behavioural strategy, possibly aided by the evaluation of alternative options.

ACKNOWLEDGMENTS

Academic acknowledgments

First and foremost, I would like to thank my supervisor, Dr Mark Buckley, for his mentorship and guidance throughout my doctorate. No chapter of this thesis would have been possible without his help and support. I would also like to thank OFTNAI, for their generous financial support that allowed me to undertake this course of studies.

For Chapter 1, I would like to thank Dr Buckley and Dr Merima Maja Brkic for allowing me to extend and complete the project they had begun together prior to the start of my DPhil. For this chapter, I tested 2 ACC patients and 11 of the control subjects, which were combined to those tested by Dr Brkic, concluding the study. I carried out the data analysis, interpretation and write-up of the findings. For this study, I also want to thank Dr Jon Simons and Ms Rosemary Dunn at the MRC Cognition and Brain Sciences Unit in Cambridge, who made it possible for me to access their patient and control volunteers database, and to test on their premises. Finally, I would like to thank Professor Glyn Humphreys and Dr Nele Demeyere, for making recruitment of our final ACC patient possible, the Department of Psychology at the University of Birmingham, for allowing me to test on their premises, and Dr Dante Mantini, for teaching me how to use MRIcron and SPM for lesion tracing in order to generate the images included in this chapter and in the Appendix.

For Chapter 2, I would like to thank Dr Rogier Mars for his valuable advice in helping me develop the TMS version of the WCST analogue, teaching me how to practically carry out TMS experiments, and providing support and guidance throughout this project. For this study, I modified the WCST program written by Dr Buckley to adapt it to the requirements of my TMS paradigm. I carried out all testing, data analysis, interpretation, and write-up of the findings. I would like to thank the Decision and Action laboratory, in the Department of Experimental Psychology, for allowing me to carry out my experiment in their TMS laboratory and use their equipment.

For Chapter 3, I would like once again to thank Dr Buckley for teaching me how to program in Turbo Pascal, which allowed me to modify various scripts in order to design the tasks used in this study. I carried out all task design, testing, data analysis and interpretation for this chapter.

For Chapter 4, I would like to express my gratitude to Professor Birte Forstmann and Dr Leendert van Maanen, for providing me with the unique opportunity to spend some time at the University of Amsterdam and learn how to apply the LBA model to my behavioural data. I would like to thank Dr van Maanen, in particular, for writing the original adaptation of Donkin and colleagues' LBA script for our WCST analogue and for all the advice, opinions and guidance he provided throughout this project. For this chapter, I collected the behavioural data to carry out the model fits (from Chapter 3), I

re-adapted Dr van Maanen's LBA script to investigate adaptation effects in the WCST and our other tasks, I carried out all model fittings, interpretation and write-up.

For Chapter 5, I would like to thank Dr Alex Green, from the Neurosurgery Department at the John Radcliffe Hospital (Oxford), for providing us access to his DBS patients and arranging the testing sessions. I also want to thank Dr Jonathan Hyam and Dr Faizal Ahmad Bahuri, for being present during the testing sessions, assisting the patients and arranging the intracranial recording procedures for this study. I would like to thank Dr Bahuri for making time in his busy schedule to teach me how to use Spike2 and Matlab for data extraction and processing, and also FSLview to determine electrode positioning in the patients' scans. For this chapter, I would also especially like to thank Dr Mark Stokes for his extremely valuable guidance with the LFP data analysis, for teaching me how to use the Field Trip analysis toolbox and for providing the original Field Trip scripts I modified for my analysis. I would also like to thank Linda Katona and Colin McNamara for their opinions and advice on the electrophysiology analysis. For this chapter, I carried out all testing, behavioural data analysis and interpretation, LFP data analysis and interpretation (with the help from the aforementioned collaborators) and write-up.

Finally, for Chapter 6, I would like to thank Dr Buckley for giving me the opportunity to work on this non-human primate project, which proved to be a really valuable experience for my doctorate. He performed the surgeries and designed the

task used in this study. I would also like to thank Dr Carinne Piekema for training the animals during the first part of this study, and the Biomedical Services staff and veterinary technicians for their excellent work in the day-to-day care and husbandry of our animals. For this chapter, I concluded the experimental phase of the project, carried out all the data analysis, interpretation, and write-up.

Personal acknowledgments

I would like to thank my parents, first and foremost, for their continuous patience, support and understanding through this life-long pursuit of education, from my first day of school to the last day of my doctorate. This is the end of a journey for them as much as it is for me. Grassie ala Piaseta, for always being there to provide fun and friendship during the, unfortunately rare, times I made it back home over the course of my DPhil. Also, to its newest addition, Aria, whose birth I had to miss to finish this thesis. Thanks to 108 and all its friends, for being a true home to me over the last five years. You'll be a hard one to top. Last, but not the least, thanks to my online buddies for making it all more bearable over the final few months, especially when the thesis was decisively on the floor!

GENERAL INTRODUCTION

THEORETICAL AND BEHAVIOURAL BACKGROUND

COGNITIVE CONTROL

The main challenge for any living agent is to interact with its environment in such a way as to maximize the outcome of its actions in order to efficiently achieve its internal goals - from basic ones such as feeding and reproduction to higher-level, more abstract goals and tasks such as acquiring a new skill-set, choosing a career, investing money, winning at competitive sports, etc. The natural environment, however, is extremely rich in information. At any point in time, it provides an agent with information that is pertinent to its current internal goals, as well as superfluous and/or conflicting information. When the goals change from one point time to another (i.e. the agent shifts from one state to another), so does the relevancy and value of the information in the environment, and actions must be selected accordingly. As a goal becomes more complex or more abstract, the amount of information that could be of potential use increases, creating more uncertainty. Furthermore, the degree of fluctuation in the information available can make the environment extremely volatile, or relatively stable and predictable. Therefore, one of the most crucial features of the cognitive system is the ability to selectively attend to relevant information, and to

regulate and coordinate goal-directed behaviour, in a rich and ever-changing environment. This feature is generally referred to as *cognitive control*.

Historical background

Cognitive control emerges at the junction of several cognitive processes - such as memory, attention, decision-making and action planning – and, as such, encompasses several underlying mechanisms shared by other cognitive functions. Indeed, theoretical accounts of cognitive control, particularly in the early decades, have stemmed from a wide variety of research into different aspects of cognition.

The earliest psychological concept to resemble what would be later defined as cognitive control can be found in William James' theory of attention (James, 1890). James distinguished between active attention and passive attention mechanisms, with the former being controlled by internal goals, in a top-down fashion, and the latter being triggered in a bottom-up fashion by external stimuli. Active attention can be thought of as part of cognitive control, allowing the agent to attend to information that is relevant to the current internal goals. The term *cognitive control* began to appear in the literature as early as the 1960's, intended as a collection of individual traits that underlie different cognitive styles (see Gardner, 1959; Gardner et al., 1959; 1968), or as the application of previously acquired information to guide behaviour, as opposed to the process of acquiring said information (Hammon & Summers, 1972). Towards the

1970's, it began being used as a term denoting internally driven, as opposed to externally triggered, behaviour (Kendler, 1971); a use more aligned to James' concept of active attention.

However, as mentioned earlier, cognitive control is not limited exclusively to attentional processes. Michael Posner was one of the first authors to refer to cognitive control as a mechanism for the conscious, intentional modulation of attention and memory and for strategy implementation (Posner, 1975; 1987). Alan Baddeley and colleagues, from the 1970's through to the 1990's, developed and investigated one of the classic models of working-memory (Baddeley & Hitch, 1974; Baddeley et al., 1996; Baddeley, 1996), which postulated the existence of a Central Executive, a component that specifically deals with cognitively demanding tasks. Several of its functions, such as the coordination of different tasks and strategies and the capacity to selectively focus attention to relevant stimuli while inhibiting distractors, are some of the defining features of cognitive control. Norman & Shallice included a comparable component, the Supervisory Attentional System (SAS), in their model of action-selection (Norman & Shallice, 1980). The SAS was proposed as a system that oversees and controls the selection of action schemas depending on environmental circumstances (routine or non-routine), in order to explain how automatic versus controlled action schemas may be generated. Similarly to the Central Executive, the selection and coordination of different actions by the SAS in response to contextual changes is an aspect of behaviour that cognitive control is responsible for.

Components such as the Central Executive are always characterized by limited processing capacity, and this limited capacity is thought to underlie behavioural limitations, such as those experienced in dual-task performance: when two concurrent tasks are performed at the same time, decrements in performance are observed, due to the fact that components such as the Central Executive can only process a limited amount of information at any one time. However, Allport (1987) pointed out that these limitations might actually be due to interference between the two tasks, rather than limited processing capacity of such a module, resulting in the need for extra control, which, in turn, results in costs on response speed. Stemming from this viewpoint, in recent years, the study of cognitive control has found a fruitful niche of investigation within the area of conflict and conflict-monitoring research.

CONFLICT

Historical background of conflict tasks

Research into conflict-processing mechanisms investigates how the cognitive system responds to interference between competing information (i.e. conflict), when the subject is engaged in instructed, goal-directed behaviour. This branch of research finds its roots in classic interference tasks that have been widely applied in experimental psychology since the 1930's. The Stroop task (Stroop, 1935) is probably the oldest and

most renowned interference task. It requires participants to name the colour a written word is printed in. Findings from this task reveal that response times are longer and errors more frequent when the colour does not match the written word (e.g. 'red' is printed in blue ink), than when it does (e.g. 'red' is printed in red ink), or when the semantic meaning of the written word is not related to the colour (e.g. 'monkey' is printed in red ink). This is known as the Stroop effect, and it has been reliably replicated across studies (e.g. Egner, Delano, & Hirsch, 2007; Jacoby, Lindsay, & Hessels, 2003; Logan, 1979; Mead et al., 2002; Simon, 1990; Stuss, Floden, Alexander, Levine, & Katz, 2001; van Veen & Carter, 2005; Vendrell, Junqué, Pujol, & Jurado, 1995). The common explanation for this effect is that the mismatch between printed colour and written word generates interference (i.e. conflict) between the instructed response (naming the ink-colour) and the habitual tendency to read the written word.

Similar effects have been obtained across a variety of other interference tasks through the last century. Simon carried out a large amount of research on lower-level interference at the perceptual level, using auditory and visual stimuli (Simon & LeMay, 1969; Simon & Small, 1969; Simon, 1969). His research showed how interference between the spatial location of a stimulus and an arbitrary response associated with its identity (e.g. pressing a button on the left handside when presented with a high pitched sound, or a red dot) affects performance. If the stimulus is presented on the opposite side from the instructed response, effects similar to the ones observed in the Stroop task are obtained: response times are longer and errors more frequent. These findings

have also been widely replicated (e.g. Chen & Melara, 2009; Hedge & Marsh, 1975; Hommel, Proctor, & Vu, 2004; Kerns, 2006; Peterson, Kane, & Alexander, 2002; Praamstra, Kleine, & Schnitzler, 1999; Simon, 1990; Simon, Sly, & Vilapakkam, 1981; Wühr, 2005).

Finally, a large body of research has also investigated interference effects between stimuli and distractors. The classic task used to investigate this type of interference is the Eriksen Flanker task (Eriksen & Schultz, 1979), during which participants are required to respond to the identity of a target stimulus (i.e. a letter), while ignoring the surrounding distractors (i.e. other letters, or symbols). When the central stimulus is of the same identity as the distractors, responses are faster and more accurate than when the distractors are different from the target stimulus (e.g. Coles & Gratton, 1985; Gratton, Coles, & Donchin, 1992; Nieuwenhuis & Stins, 2006; Ullsperger, Bylsma, & Botvinick, 2005).

Within this type of experimental paradigms, the cognitive system is presented with all the fundamental challenges that require the implementation of cognitive control: the need to attend to task-relevant information in the presence of competing distractors and the need to coordinate and execute a behaviour, based on an internal-goal, whilst in the presence of irrelevant information that can prime responses that are not beneficial in the current context. These tasks, and a large number of variations and combinations of these tasks (Egner et al., 2007; Hazeltine, Poldrack, & Gabrieli, 2000;

Peterson et al., 2002; Sheth et al., 2012; Verbruggen, Notebaert, Liefooghe, & Vandierendonck, 2006) have therefore become widely applied to research into cognitive control.

Conflict-monitoring

While all these tasks effectively exemplify the type of situations during which cognitive control is required in order to perform the appropriate behaviour, the need to recruit such a mechanism cannot always be foreseen in advance. Furthermore, evidence suggests that control is not an absolute mechanism, which is either engaged or not engaged depending exclusively on overall task difficulty (in this case, interference). Rather, control can be increased or decreased via the influence of other factors such as practice (Cohen, Dunbar, & McClelland, 1990; Kornblum, Hasbroucq, & Osman, 1990; Shiffrin & Schneider, 1977), even in classic conflict task like the Stroop task (Milham, Banich, Claus, & Cohen, 2003; Heink, Bibi, Yanai & Tzeglov, 1997). The more a task, or skill, becomes familiar, the more efficiently it is performed, reducing the need for cognitive control. Frequency of interference has also been shown to play a role in performance: the more the participant is faced with interference, the better the performance (Hommel, 1994; Logan, 1979; Nieuwenhuis, Yeung, van den Wildenberg, & Ridderinkhof, 2003; Sturmer, Leuthold, Soetens, Schro, & Sommer, 2002). Two classic effects, which demonstrate how the engagement of cognitive control is relative rather than absolute, are speed-accuracy trade-offs (Pew, 1969; Wickelgren, 1977) and post-

error slowing (Laming, 1979; Rabbitt, 1966). In experimental settings, participants are usually instructed to respond as fast and as accurately as possible. However, when accuracy is emphasized in the instructions, response times tend to become slower, in order to ensure greater accuracy, than when speed is emphasized, even if the task demands in the two emphasis conditions are exactly the same. Post-error slowing is a similar effect, which results in higher response times in order to increase accuracy following an error, a situation that essentially implicitly emphasizes accuracy via negative feedback.

In view of the aforementioned issues, Botvinick and colleagues (2001) pointed out that one aspect of cognitive control that had been, until then, neglected by the literature was precisely the question of how the need for increased cognitive control is signalled to the cognitive system. This question led them to develop their conflict-monitoring and cognitive-control model (here onwards referred to as the CMCC model), which represented the start of a popular stream of research over the last two decades. The basic tenets of the CMCC framework suggest that a monitoring module detects the occurrence of conflict. The module then biases other modules within the cognitive system to actively implement cognitive control, including increased attention, resistance to distractors and the production of contextually appropriate goal-directed responses (see Figure 1). Research into the conflict-monitoring and cognitive control account has boomed over the last two decades, particularly due to the substantial support it has obtained from evidence in the field of neuroimaging (e.g. Kerns, 2006; Kerns et al.,

2004; MacDonald, 2000; van Veen, Cohen, Botvinick, Stenger, & Carter, 2001), as well as the applicability to the design of computational models (e.g. Brown & Braver, 2008; Davelaar, 2008; van Veen, Cohen, Botvinick, Stenger, & Carter, 2001; Ward & Ward, 2006).

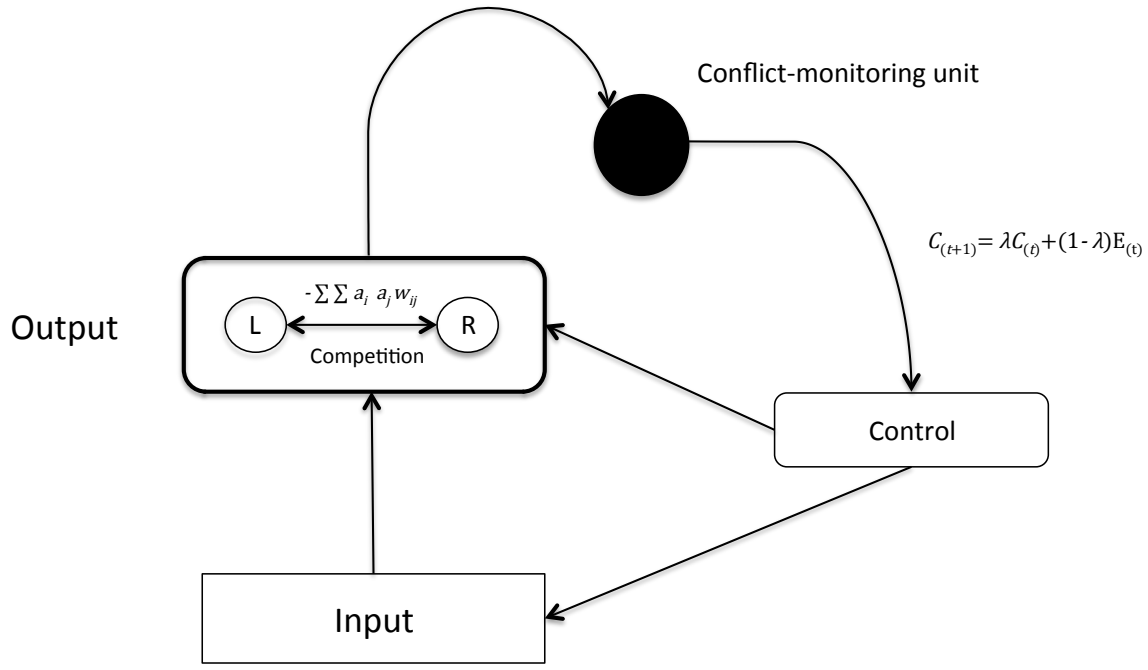


Figure 1. *Basic schematic representation of the Conflict Monitoring and Cognitive Control model of Botvinick and colleagues (2001). The amount of competition (i.e. conflict) between mutually inhibiting response units (e.g. left or right button press) in the output layer is measured as Hopfield energy. Energy is zero if no unit, or only one unit, is active at any one point in time. When both units are concurrently active, energy in the output layer rises. Hopfield energy levels are monitored by a conflict-monitoring unit. Activity in the conflict-monitoring unit increases as a function of the amount of Hopfield energy in the response layer. Energy on trial* t *is then converted into a control value* C*, to bias activation of units implementing cognitive control, which affects input processing and response biasing/inhibition.*

FLEXIBILITY AND ADAPTATION

The fact that conflict can be detected by a conflict-monitoring module in the first place, means that its presence is contingent on changes in the environmental context,

or task: if something is constantly present, then, by definition, there should be no need for detection. Stoffels (1996) was one of the first authors to investigate how presenting participants with conflict and non-conflict trials in separate blocks, as compared to mixing the two types of trials in the same block, affected performance. He observed that mixing conflict and non-conflict trials within the same block generated uncertainty, which affected non-conflict trials only, by slowing down response times. Stoffels attributed these findings to the fact that uncertainty (i.e. mixing conflict and non-conflict trials within a block) triggers the need for increased caution, and this reduces response speed to trials that would normally involve less effortful, and less costly, automatic responses (i.e. non-conflict trials) as opposed to trials that require controlled responses. Conflict tasks can therefore also be used to investigate the third fundamental challenge for the cognitive system that requires the implementation of control, beyond simply ignoring distractors and resisting interference: the ability to flexibly modulate behaviour depending on contextual changes; in this case, changes in conflict levels.

Conflict-induced behavioural adaptation

While one part of conflict-monitoring research focuses on performance differences on the current trial, depending on whether or not conflict is present, another part is concerned with adaptation mechanisms; that is, how performance is modulated on trials *after* conflict has been experienced. In conflict tasks, as we saw earlier, responses to trials containing interference are usually slower (and less accurate)

than to those who do not; as they require the recruitment of cognitive control in order to perform the accurate response, which is a costly and effortful process. We will refer to the response time difference between the two types of trials as *conflict effects*. These effects, also known as interference or congruency effects, are the focus of the part of conflict-monitoring research that looks at conflict detection. Another category of findings is represented by the so-called *sequential* or *adaptation effects*. First reported by Gratton and colleagues in the 1990's (Gratton et al., 1992), adaptation effects refer to the reduction in the size of conflict effects after conflict has been experienced on the previous trial. This phenomenon, like conflict effects, has also been consistently replicated in the literature (e.g. Akçay & Hazeltine, 2008; Botvinick, Braver, & Barch, 2001; Burle, Allain, Vidal, & Hasbroucq, 2005; Chen & Melara, 2009; Egner & Hirsch, 2005; Notebaert & Verguts, 2007; Purmann, Badde, & Wendt, 2009; Stürmer & Leuthold, 2003; Stürmer, Siggelkow, Dengler, & Leuthold, 2000; Ullsperger, Bylsma, & Botvinick, 2005; Wendt, Kluwe, & Peters, 2006; Wühr & Ansorge, 2005), and represents one example of how cognitive control can be regulated by changes in environmental contingencies (in this case, conflict levels).

The CMCC model suggests that cognitive control is recruited in order to prevent detrimental conflict effects on performance. This tenet of the model is directly relevant to explain adaptation effects. Once the conflict-monitoring module detects conflict, control is implemented, and this reduces the effect of conflict on subsequent trials (Botvnick et al., 2001). However, the precise mechanisms by which cognitive control

reduces conflict susceptibility are still very much undefined. Several competing and/or complementary accounts have been proposed.

Some accounts suggest that adaptation effects can be explained by mechanisms unrelated to cognitive control, such as task design limitations that encourage repetition or negative priming (Mayr, Awh, & Laurey, 2003) or costs incurred in the binding and unbinding of arbitrary stimulus-response associations, rather than being driven by increased cognitive control (Chen & Melara, 2009). However, while these mechanisms might play a role, adaptation effects have been reported even in the absence of trials that may trigger priming effects (e.g. Akçay & Hazeltine, 2008; Fernandez-Duque & Knight, 2008; Kerns, 2006; Kerns et al., 2004; Notebaert & Verguts, 2007; Ullsperger, Bylsma, & Botvinick, 2005), and in tasks that do not involve any arbitrary stimulus-response associations (e.g. Mansouri, Buckley, & Tanaka, 2007).

One of the most common explanations for adaptation effects is that conflict generates competition between two or more responses and that, once cognitive control is engaged, the irrelevant response is suppressed, and the relevant response is enhanced, therefore reducing interference on the subsequent trial (e.g. Botvinick et al., 2001; Yeung, Botvinick, & Cohen, 2004). If this were the case, the reduction in the conflict effects should involve an increase in speed of response and accuracy on subsequent conflict trials. The first conflict trial in a sequence will always incur performance costs because cognitive control is not yet recruited. Once control is

recruited, suppression of the irrelevant response and enhancement of the relevant response should lead to better performance on conflict trials. However, many studies do not specify whether this is the case (Chen & Melara, 2009; Kunde, 2003; Purmann et al., 2009; Sturmer et al., 2002; Verbruggen et al., 2006; Wühr & Ansorge, 2005) or report adaptation effects that are entirely due to reductions in speed of response on non-conflict trials (e.g. Horga et al., 2011; Stoffels, 1996). This suggests that cognitive control is, in some contexts, exerted primarily or exclusively through an increase in caution, rather than enhanced attention or rule-guided behaviour. Furthermore, some studies have indicated that conflict and competition between responses are not necessarily synonymous. Burle and colleagues (2005) showed that competition between two responses is occasionally present even on non-conflict trials, but that adaptation effects are dependent on the level of conflict across trials, rather than competition. This suggests that it is not competition that triggers changes in cognitive control.

A somewhat similar account to the one just described also suggests that adaptation effects are due to the suppression of irrelevant responses. However, according to this framework, adaptation is due to the modulation of more general response mechanisms, rather than specific competing responses. More specifically, adaptation effects might stem from the interplay between two different routes to response that support habitual/automatic versus instructed/rule-based responses, respectively (Hommel et al., 2004; Stoffels, 1996; Sturmer et al., 2002). The dual-route to response framework is particularly compelling for a number of reasons.

First, of all it is generalizable to all classic conflict tasks. In all classic conflict tasks, conflict is indeed elicited by creating competition between a habitual behaviour (such as reading words, in the case of the Stroop task, or directing attention to a stimulus location, in the case of the Simon task, or processing all information presented, in the case of the Flanker task) and an instructed response (such as naming the colour, responding to the identity, or focusing on a detail). Therefore, the idea that adaptation effects might be due to the modulation of mechanisms that support these two types of responses is a particularly convincing one.

Secondly, it is consistent with the predictions that stem from the CMCC model and can also provide a more specific explanation of why reductions in speed on non-conflict trials are often observed as a part of adaptation effects: the detection of conflict triggers adjustments in the modulation of the two routes, with suppression of the route supporting habitual/automatic responses (which results in slower responses to non-conflict trials) and enhanced biasing of the route supporting instructed responses resulting in less susceptibility to conflict.

Third, the interplay between action (intended as effortful, rule-guided and controlled behaviour) and habits (intended as relatively effortless, automatic behaviour), and its effects on cognition, has been a very extensively investigated topic since the early days of cognitive science (Cohen et al., 1990; Dickinson, 1985; Neisser,

1976), and can be seen as another approach to describing cognitive control. Neisser (1976) was one of the first theorists to argue that cognition is better described as the continuous interaction of bottom-up, automatic processes, and top-down, controlled processes, rather than exclusively one or the other. Norman & Shallice's SAS module (1980) specifically refers to the modulation of actions and habits. The interplay between actions and habits could account for the effects of frequency and practice on conflict effects. The more the instructed response strategy (or action) leading to correct responses on high-conflict trials is applied, the more it can be consolidated and become less dependent on cognitive control for its execution, possibly to the point of automaticity (i.e. becoming habitual). This would be reflected in quicker responses on high-conflict trials.

CONCLUSIVE REMARKS ON THE THEORETICAL BACKGROUND

To summarize, the detection of conflict in a task has been shown to trigger the recruitment of cognitive control, and this drives adaptation mechanisms that can prevent the detrimental effects of conflict on performance (and therefore reduce conflict effects) from one trial to the next. The actual mechanisms underlying adaptation effects are still unclear, but, from a review of the literature, it emerges that there may be two dissociable components that contribute to the reduction of conflict effects following exposure to conflict: caution, which may be reflected in slower responses after exposure to conflict, and enhancement of attention or processing of task-relevant

information, which may be reflected in faster responses to conflict trials after exposure to conflict. Furthermore, adaptation effects might result from the modulation of two routes to response, one supporting automatic, habitual responses, and one supporting rule-guided, instructed behaviour. All classic conflict tasks rely on competition between these two routes, as conflict is usually elicited between a habitual response and an instructed response. Increased caution could result in suppression of the habitual response route, and increased attention/task-relevant processing could result in biasing of the instructed response route. However, conflict and adaptation effects have also been reported using tasks that do not create competition between a habitual response and an instructed one, but only between two instructed responses, such as in the case of the Wisconsin Card Sorting Test (Mansouri et al., 2007). This raises several questions. First, does the cognitive system comprise general-purpose, centralized conflict-monitoring and cognitive control mechanisms that are necessary for conflict resolution and adaptation? If so, would they play a similar role when competition is created between instructed responses as opposed to habits and instructions? Also, can the role of cognitive control in behavioural adaptation be empirically dissociated into several components?

NEURAL SUBSTRATES OF COGNITIVE CONTROL AND BEHAVIOURAL FLEXIBILITY

So far, we reviewed the behavioural evidence concerning the recruitment and implementation of cognitive control. One of the main limitations of behavioural evidence in the study of cognition is the fact that it relies exclusively on response times and accuracy measures to infer the underlying cognitive processes. However, as response times represent the latency of all the processes that unfold from the moment the information (or stimulus) is inputted, to the moment a response is generated as an output, they can only provide an indirect measure of the processes happening between input and output. In order to better understand how information is processed between the input and output stages, it is therefore essential to identify neural correspondences to cognitive processes. Therefore, in order to answer the question of whether there are conflict-monitoring and cognitive control/flexibility modules in the cognitive system, whether they are essential to performance, and how their functions are carried out, it is necessary to identify brain regions that appear to be involved in these cognitive processes.

THE PREFRONTAL CORTEX IN COGNITIVE CONTROL AND FLEXIBILITY

A region of the frontal lobe known as the prefrontal cortex (PFC) has been identified as most likely candidate structure for supporting cognitive control and

behavioural flexibility. The PFC lies anterior to the premotor and somatosensory areas of the brain (Fuster, 2008; Nolte, 2009). It is the largest and evolutionary youngest region of the primate brain and it constitutes approximately 30-40% of the total cortical volume in the human brain (Brodmann, 1912; Semendeferi, Lu, Schenker, & Damasio, 2002).

The idea that PFC supports cognitive control initially stemmed from neuropsychological studies of the consequences of frontal lobe damage on executive function (Walsh, 1978). While widespread lesions to the frontal lobe can impinge upon very different aspects of behaviour (including emotion, personality and cognition), deficits specific to executive function have been attributed to lesions involving predominantly the anterior part of the frontal lobe, the PFC. Neuropsychological studies allowed for the identification of candidate regions for the executive function modules that had been hypothesized in models of memory and action. For example, Baddeley (1996), found that damage to this region resulted in the inability to direct and re-direct processing resources in a manner that is appropriate for the current context, resulting in what he defined as Dysexecutive Syndrome, the impairment of the functions ascribed to the Central Executive proposed in their model of working memory. Similarly, Norman and Shallice (1980) noted that the functions ascribed to the SAS module, described in their theory of action planning, are also impaired by prefrontal damage, resulting in poor action planning in situations where automatic responses are not advantageous, difficulties in learning new tasks and the inability to correct behaviour after errors. With

increased investigation into the role of PFC, it is now clear that damage to this area can result in substantial impairments in a wide range of executive functions, such as attention (Chao & Knight, 1995; Dias, Robbins, & Roberts, 1996; Godefroy & Rousseaux, 1996; Kane & Engle, 2002; Knight, Grabowecky, & Scabini, 1995), working-memory (Bauer & Fuster, 1976; Goldman‐Rakic, 1987; Petrides, 2000; Thompson-Schill et al., 2002), and decision-making and planning (Goldman‐Rakic, 1987; Milner, 1982; A. Owen, Downes, & Sahakian, 1990; Shallice & Burgess, 1991; Treichler, 1973). These effects have been reported extensively in both the animal and human literature (for an extensive review of the neuropsychological literature across species see Fuster, 2008). All of these executive functions play a crucial role in cognitive control and therefore strongly implicate PFC in this process.

Studies of the structure and connectivity, as well as the functional properties, of PFC in the healthy brain have complemented the neuropsychological findings by highlighting characteristics of this region that make it particularly suitable to support the requirements of goal-directed behaviour in a rich and changing environment.

The first requirement is flexibility. Like in the rest of the cortex, most neurons in PFC are excitatory pyramidal cells with high numbers of dendritic spines and long axons that allow long-range connections with many other cortical and subcortical areas (Nolte, 2009). This makes the PFC ideally conformed to support the integration and processing of a large and diversified amount of information. However, differently from other

cortical cells, PFC cells display highly adaptable response profiles. Prefrontal neurons receive and respond to information from all sensory systems as well as motor and reward areas of the brain (Masterman & Cummings, 1997; Miller & Cohen, 2001; Yeterian, Pandya, Tomaiuolo, & Petrides, 2012), rather than being sensitive purely to specific sensory modalities or types of information (as it is the case for other cortical areas). They are able to integrate larger numbers of inputs at any one time than cells in non-frontal areas (Elston, 2000) and they have been shown to encode different aspects of an environmental context in the form of a 'temporary trace', which changes when the context changes (Duncan, 2001).

All the information PFC has access to needs to then be processed in order to fulfil the second requirement of goal-directed behaviour, which is the ability to interact with the environment based on longer-term internal goals, rather than simply responding to immediate environmental variability. The PFC does not only receive input from a large number of motor, limbic, sensory and multimodal association areas, but also projects back to them via several fibre pathways. Miller and Cohen (2001), in their integrative theory of PFC function, suggested that this region may be involved in maintaining patterns of activity representing goals and the means to achieve them (i.e. the primary role of cognitive control) by processing all the information provided by the environment at any point in time and then guiding the flow of activity to perform a specific task along contextually relevant pathways.

PFC neurons have been shown to encode abstract rules (Asaad, Rainer, & Miller, 2000; Asaad, Rainer, & Miller, 1998; Bunge, Kahn, Wallis, Miller, & Wagner, 2003; Wallis, Anderson, & Miller, 2001; White & Wise, 1999), to resist to intervening distractors during a task (Miller & Desimone, 1994; Miller, 1996) and to maintain information across delays (Bunge et al., 2003; Fuster & Alexander, 1971; Goldman-Rakic, 1995; Mansouri et al., 2007; Miller, 1996). Abstract rules represent the building blocks for the formation and maintenance of internally generated goals and the means to achieve them. Resisting distractors allows the agent to exploit prior knowledge and previously encoded information over reflexively responding to what is afforded by the environment at a given point in time. Finally, the ability to maintain information and goals across temporal delays is seen as a fundamental feature of PFC to bridge the gap between perception and action (Fuster, 1990). Bringing all these features together, a hierarchical structure of PFC has recently been proposed, with increasingly abstract and temporally distant concepts, goals and plans being represented along a rostro-caudal gradient (Badre & D'Esposito, 2007; Badre, 2008; Botvinick, 2008). This ensures that the wide range of varied information that is available to PFC at any point in time can be exploited in favour of the agent's current internal goals, from simple to more abstract ones.

When looking at the neuropsychology literature on PFC damage, Fuster (2008) further distinguished between lateral prefrontal and medial prefrontal syndromes, the former being characterized by impairments in sustained attention, working-memory and

planning, and the latter being characterized by impairments in the monitoring of current and past performance, and the correction of errors. This indicates that there may be a further functional subdivision within the PFC of the frontal lobe, with parts of PFC being more involved in actively controlling behaviour, and parts being more involved in monitoring the results of behaviour, or the environmental contingencies, in order to trigger behavioural adjustments, when needed. Lateral PFC comprises Brodmann areas 8, 9, 10 and 46, whereas medial PFC comprises the medial portions of Brodmann areas 6, 8, 9 and 10, as well as the entirety of areas 24 and 32. While lateral PFC areas have dense reciprocal connections particularly with regions involved in the execution of actions (e.g. motor and pre-motor cortical areas and the basal ganglia) and high-order association areas - making it a prime candidate for the processing and integration of information to control behaviour - medial PFC areas also have dense reciprocal connections particularly with areas involved in processing emotional and reward information (e.g. the amygdala) and memory (e.g. the hippocampus) – making it a candidate structure for the monitoring of performance and feedback. Furthermore, these two subdivisions of the prefrontal cortex also present extensive reciprocal connections with each other (Wood & Grafman, 2003). This suggests that, rather than being a functionally homogeneous region, the PFC may comprise a network of several, functionally distinct subdivisions (see Petrides & Pandya, 2002; Petrides, 2005) that presents the ideal structural and connectivity profile to implement flexible, goal-directed behaviour using information about feedback obtained from current and recent performance.

Proposed functional subdivisions within the frontal lobe for the support of cognitive control and flexibility

The three proposed functional subdivisions within the PFC network that this thesis will focus on are the anterior cingulate cortex (ACC - in the medial-prefrontal area), the dorsolateral PFC (dlPFC – in the lateral prefrontal area) and the frontopolar cortex (FPC – occupying both lateral and medial PFC). All three of these areas present extremely dense local connectivity with each other, as well as with other prefrontal areas, such as motor areas and the orbitofrontal cortex (OFC), and posterior sensory, somatosensory and multimodal association areas. Despite these similar bidirectional projections, they also present substantial cytoarchitectonic differences (in the size, density and organization of cells within the cortical layers) and different proposed functional properties have been proposed for them.

As illustrated earlier, the CMCC model hypothesizes that one way cognitive control could be recruited is thanks to the presence of a monitoring module that detects the presence of conflict, and this alerts the cognitive system to the need to actively implement cognitive control (Botvinick et al., 2001). This type of activity seems to fit the functional profile of the medial-prefrontal areas discussed above, which are thought to perform monitoring functions. Indeed, Botvinck and colleagues postulate that the ACC is the neurological equivalent of their conflict-monitoring module (e.g. Botvinick, 2007;

Carter & Veen, 2007). The rich connections between this area and dlPFC areas make the latter the primary candidate to be the neural substrate of the modules whose role is to actively implement cognitive control in response to ACC signalling (e.g. Carter & Veen, 2007; Kerns, 2006). This correspondence is illustrated in Figure 2. Finally, the FPC has been proposed to play an important role in behavioural flexibility and goal-directed behaviour, by supporting complex, higher-order processing and decision-making (e.g. Badre, 2008; Koechlin & Summerfield, 2007) and/or by driving exploration and the evaluation of alternative options and responses (e.g. Boorman, Behrens, Woolrich, & Rushworth, 2009; Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006).

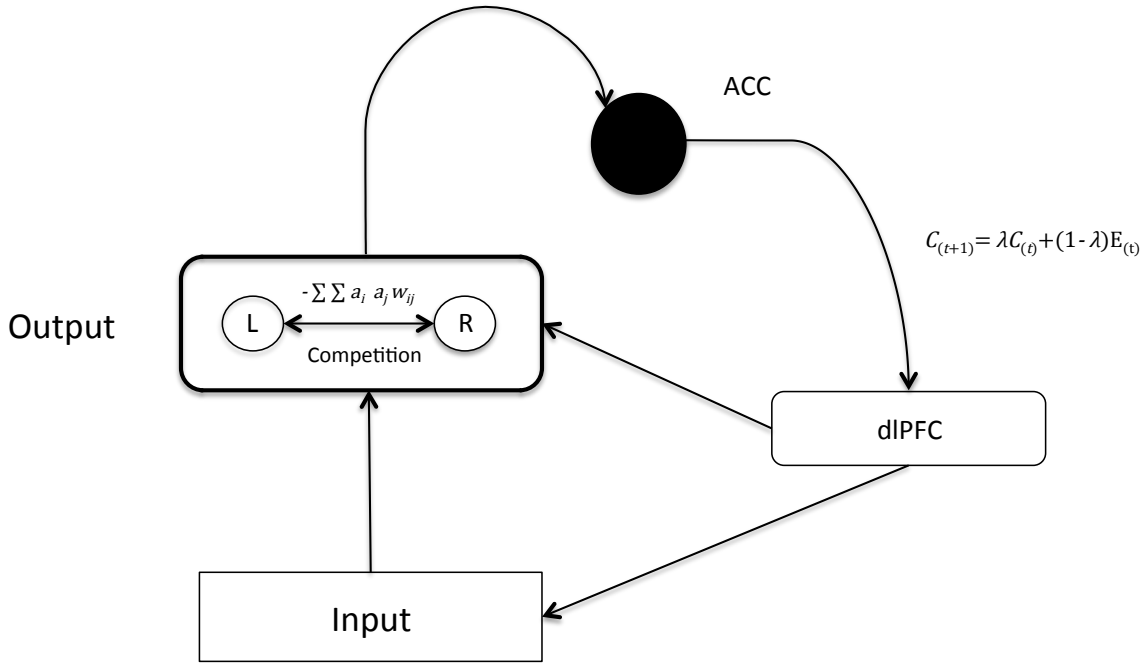


Figure 2. *Proposed correspondence between units in the CMCC model by Botvinick and colleagues (2001) and prefrontal regions. The ACC is proposed to support the functions of the conflict-monitoring unit. The dlPFC is proposed to be involved in the implementation of cognitive control in response to ACC signalling.*

ACC: THE MONITOR

ACC anatomy and connectivity

As mentioned earlier, medial PFC includes the medial portions of Brodmann areas 6, 8, 9 and 10, as well as the entirety of areas 24 and 32 (Fuster, 2008). The latter two regions constitute the anterior cingulate cortex (ACC – see Figure 3), a region extending dorsally over the corpus callosum, also sometimes identified as part of the 'limbic cortex', a more evolutionary ancient structure within the frontal lobe (Papez, 1937; Petrides & Pandya, 1994; Nolte, 2009). Area 24 has been further subdivided into three horizontally distinct areas, 24a, 24b and 24c, the former two occupy the cingulate gyrus and the latter occupying the cingulate sulcus (Petrides & Pandya, 1994; Vogt, Nimchinsky, Vogt, & Hof, 1995). In the human brain, area 24 is surrounded by area 32, which extends over the paralimbic gyrus and is distinguished by having a higher neuronal density than area 24. The presence of a double cingulate sulcus in the human brain is reported in about 30-50% of cases (Paus, Tomaiuolo, & Otaky, 1996; Vogt et al., 1995).

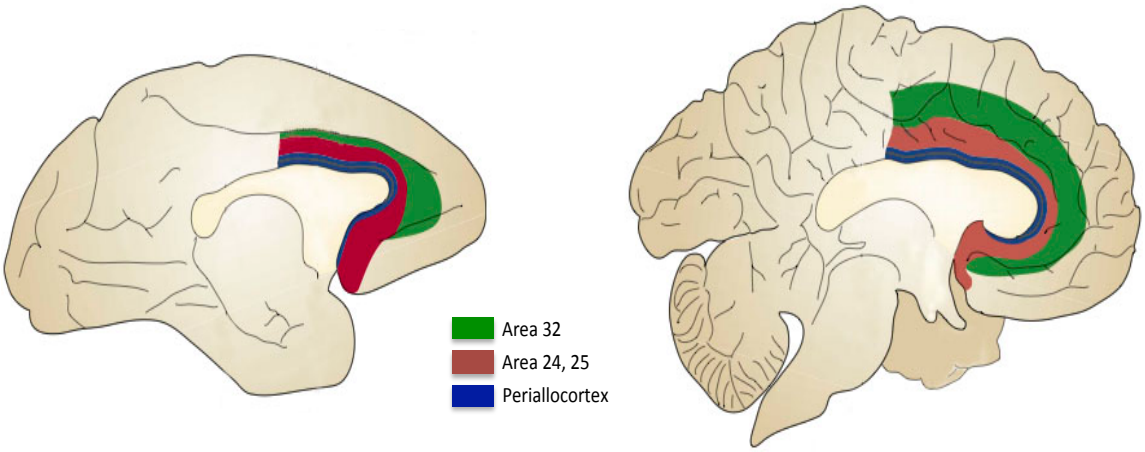


Figure 3. *The Anterior Cingulate cortex (ACC) in the monkey (left) and human (right) brain comprises Brodmann areas 32, 24 and subcallosal area 25 (which lays in the most rostro-ventral portion of area 24. The thin layer of Peralliocortex, shown here in blue, represents the boundary region between the neocortex and the Corpus Callosum. Figure adapted from Paus (2001, pp. 418).*

The ACC is characterized by the presence of pyramidal cells, as common in all cortical areas, and also of spindle neurons, a rare type of cell which has so far only been reported in a few species (great apes and mammals with large brains, such as elephant and whales) and which are localized to very specific areas of the brain: the ACC, the anterior insula and the dlPFC (Allman, Hakeem, & Watson, 2002; Fajardo et al., 2008; Nimchinsky et al., 1999). The ACC is the area presenting the highest density of dopamine receptors in the human cortex (Paus, 2001).

The ACC has reciprocal local connections to all other prefrontal areas (dorsal and ventro-lateral PFC, FPC, orbitofrontal cortex, and motor areas), to more posterior sensory and multimodal association regions in the temporal, parietal and occipital

cortices, as well as other paralimbic areas of the brain, particularly via the cingulate fasciculus (Yeterian et al., 2012). The ACC is connected to the basal ganglia via reciprocal projections, which, when damage, usually result in lack of motivation and apathy (Masterman & Cummings, 1997). The ACC also has dense projections to the spinal cord, the thalamus and the brainstem nuclei (Paus, 2001).

Due to its cytoarchitectonic and connectivity properties, ACC is considered to be a perfect structure for the convergence of cognitive, motor and emotional/reward processes (Paus, 2001), and, as such seems suited to support a monitoring function, both of performance and environmental contingencies that can lead to rewards.

As we will discuss in more detail later, the monkey and human ACC are fairly comparable as far as area 24 is concerned (Petrides & Pandya, 1994). The validity of comparisons between the monkey and the human area 32, however, has been the subject of debate (Cole et al., 2009; Shall & Emeric, 2010; Cole, Yeung, Freiwald, & Botvinick, 2010). In the monkey, this region seems to be smaller and restricted to more rostral parts of the ACC than in the human brain, where it surrounds area 24 to extend into a caudal portion termed area 32' (see Figure 4). Therefore, care must be taken when comparing findings from the two species.

ACC and conflict monitoring

In asking how cognitive control is recruited, Bovtinick and colleagues (2001) have argued that the ACC is involved in monitoring the presence of conflict in a task. If conflict is detected, the ACC signals the need to implement cognitive control to other structures, such as the dlPFC. This view has been supported by several other authors (e.g. Carter & Veen, 2007; Egner & Hirsch, 2005; Kerns et al., 2004; Yeung, Botvinick, & Cohen, 2004) and has received particular substantiation from the neuroimaging literature. ACC has been found to be highly active during conflict trials across a range of different, classic conflict tasks, such as the Stroop (Chen, Wei, & Zhou, 2006; Egner & Hirsch, 2005; Egner & Hirsch, 2005; Fan, 2003; Kerns et al., 2004; Kim, Kroger, & Kim, 2011; MacDonald, 2000; Mead et al., 2002), Simon (Horga et al., 2011; Kerns, 2006; Peterson et al., 2002; Wittfoth, Schardt, Fahle, & Herrmann, 2009) and Flanker (Durston, 2003; Fan, 2003; Hazeltine, Poldrack, & Gabrieli, 2000; Luks, Simpson, Dale, & Hough, 2007; van Veen, Cohen, Botvinick, Stenger, & Carter, 2001; Wittfoth et al., 2009) tasks. In a number of these studies, ACC activation has been observed with or immediately prior to dlFPC activation and has been correlated with good conflict-related adaptation on subsequent trials (Durston, 2003; Egner & Hirsch, 2005; Kerns, 2006; Kim, Chung, & Kim, 2012).

More supporting evidence has been provided by studies using electroencephalography (EEG), which have identified a number of event-related

potential (ERP) components that seem to be related to the presence of conflict in a trial and to originate from the ACC, such as the ERN (Carbonnell & Falkenstein, 2006; Coles, Scheffers, & Holroyd, 2001; Holroyd et al., 2004; Swick & Turken, 2002); N2 (Nieuwenhuis et al., 2003; Van Veen & Carter, 2002; West, Jakubek, Wymbs, Perry, & Moore, 2005; Yeung et al., 2004), N450 (Larson, Kaufman, & Perlstein, 2009; Swick & Turken, 2002; West et al., 2005), SP (Larson et al., 2009; West et al., 2005) and, to a lesser extent, P300 (Galashan, Wittfoth, Fehr, & Herrmann, 2008).

Despite the body of evidence originating from the fMRI and EEG literature, investigations into the role of the ACC in conflict monitoring in the macaque (via targeted lesions or single-cell recording) or in human neuropsychological investigations has proven less conclusive on the matter. Single-cell recordings in the monkey ACC have failed to identify cells that respond exclusively, if at all, to conflict (Emeric et al., 2008; Ito, Stuphorn, Brown, & Schall, 2003; Nakamura, Roesch, & Olson, 2005). Similarly, lesions to this area of the monkey brain do not seem to result in impairments in conflict tasks (Mansouri et al., 2007a). The human literature on the effects of ACC damage has shown that patients with lesions to this region either do not show impairments in classic conflict tasks (Fellows & Farah, 2005; Stuss et al., 2001; Swick & Jovanovic, 2002; Vendrell et al., 1995) or impairments restricted to the ability to use conflict information to drive behavioural adaptation on subsequent trials (diPellegrino et al., 2007; Sheth et al., 2012).

Explanations for the contrasting evidence

A number of explanations have been proposed to reconcile the fMRI and EEG evidence with the intracranial recordings and neuropsychological literature. One explanation that has been suggested is that there might be structural differences between the monkey ACC and the human ACC that make comparisons unreliable (Cole, Yeung, Freiwald, & Botvinick, 2009).

As mentioned earlier, area 24 of the ACC is cytoarchitectonically very similar between the human and the macaque brain (Petrides & Pandya, 1994). However, Cole and colleagues have argued that area 32 might not be as directly comparable (Cole et al., 2009; Cole et al. 2010). While, in the human brain, area 32 surrounds area 24, effectively forming a parallel horizontal layer dorsal to area 24, in the macaque brain, it is mostly restricted to a smaller zone in front of area 24 (see Figure 4). The large, caudal extension along the rostro-dorsal gradient of human ACC is also known as the Rostral Cingulate Zone (RCZ, RCZa = anterior, RCZp = posterior) and it occupies most of the caudal portions of areas 32 and 24. These have been named 32' and 24' (Vogt et al., 1995), to indicate a different subdivision from the more anterior 32 and 24. The RCZp/area 32' of the human brain has been shown to be sensitive to pre-response conflict in a meta-analysis of 38 fMRI studies carried out by Ridderinkhof and colleagues (2004). The homologue region in the monkey brain has been tentatively identified in the Cingulate Motor Area (CMA), more specifically in the ventral CMA (CMAv) (Picard &

Strick, 1996; Ridderinkhof & Ullsperger, 2004). However, while the human homologue of CMAv occupies area 32', it occupies area 23c in the monkey brain. Cole and colleagues argue that, due to the absence of a direct homologue of human 32' in the monkey brain, single-cell recording studies might have missed the true location of conflict processing in non-human primate brain. Alternatively, the same authors have suggested that the techniques used to investigate conflict-monitoring in humans might be able to uncover dynamics that cannot be observed with the methodologies applied to animal research, such as the sensitivity of fMRI to detect activation due to the competition between populations of interneurons (Cole et al., 2010). However, this does not explain why some lesion studies in humans that do include area 32' have found no effects of damage to this region on measures of conflict-monitoring or adaptation (for example Fellows & Farah, 2005; Baird et al., 2006), while others do (for example, diPellegrino et al., 2007; Sheth et al., 2012).

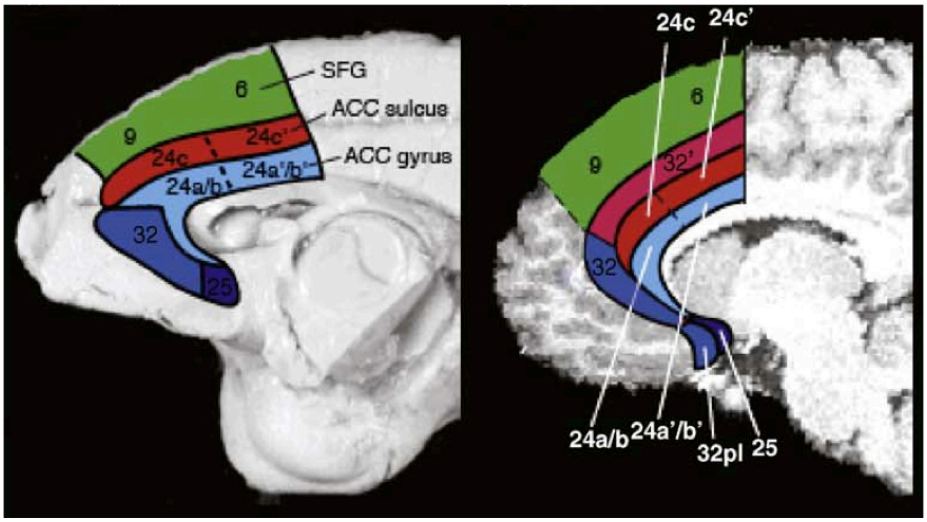


Figure 4. *Comparison between the human (right) and monkey (left) ACC. Area 32 in the macaque brain appears confined to the most rostral regions of the ACC, whereas it extends caudally to form a dorsal layer over area 24 in the human brain. The most caudal portion of this extension in the human brain has been termed area 32'. Figure adapted from Cole and colleagues (2009, pp. 568).*

Another possible explanation for the discrepancy between animal and human research, as well as between the results obtained from different human lesion studies, concerns the possibility that different subdivisions of the ACC might be involved in processing only specific types of conflict: emotional vs. cognitive, or perceptual-based vs. response-based.

The pure response theory of ACC function argues that ACC is only sensitive to response-based conflict (Milham et al., 2001; van Veen, Cohen, Botvinick, Stenger, & Carter, 2001). This theory suggests that tasks such as the Stroop task trigger conflict between the perceptual representations of the word (ink colour versus semantic meaning), whereas tasks like the Simon task trigger conflict that is exclusively restricted to competition between two responses (e.g. left versus right button-press). Therefore, discrepancies between different studies might be due to whether the task elicits perceptual-based conflict (as in the case of the Stroop task, in which case ACC should not be necessary for performance), or response-based conflict (as in the case of the Simon task, in which case ACC should be necessary for performance). However, other findings have indicated that both the perceptual-based Stroop task and the response-based Simon task activated the ACC (Peterson et al., 2002) and that the difference in activation due to the types of response conflict was instead evident in other cortical regions (Davelaar, 2008; Egner et al., 2007).

Another body of research indicates the possible presence of different subdivisions within human ACC that might respond differently to different types of conflict, emotional vs. cognitive or perceptual vs. response, along a rostro-caudal gradient, with a more rostral subdivision responding to response and/or emotional conflict and a more caudal subdivision responding to perceptual and/or cognitive conflict (see Bush, Luu, & Posner, 2000; van Veen & Carter, 2005) . Van Veen and Carter (2005) devised a version of the Stroop task that allowed the separation of semantic and response based conflict and they found that albeit both elicited ACC activation, they did so in non-overlapping regions within this area. More specifically, they showed that semantic conflict activated more caudal ACC areas, whereas response conflict activated more rostral regions of the ACC. Similar results were obtained by Kim and colleagues, with the same type of task (Kim et al., 2012, 2011). Maier and diPellegrino (2012) found that the rostral ACC lesions impaired conflict-induced adaptation on sequences of trials only when processing emotionally salient material, but not when processing emotionally neutral material. Such fine-grained subdivisions might not be as easily highlighted by human lesion studies involving rather widespread lesions. However, other studies have failed to show such clear-cut functional subdivisions within the ACC, with the same areas being involved in both perceptual and response based conflict (Fan, 2003; Weissman, Giesbrecht, Song, Mangun, & Woldorff, 2003; Wendt, Heldmann, Münte, & Kluwe, 2007), or lesions to rostral ACC affecting adaptation on cognitive-based tasks such as the Simon task (diPellegrino et al., 2007).

The issue with trying to tease apart subdivisions of ACC that are sensitive to one type of conflict over another is that different types of conflict may not be so easily dissociable. Every task will involve some reward/punishment component that provides emotional saliency to the context, even if the stimuli are not emotionally salient per-se. Also, every task involving conflict will involve some degree of perceptual and some degree of response conflict. Many tasks that attempt to tease the two apart use paradigms where two conflicting stimuli are mapped onto the same response, in order to elicit exclusively stimulus-based conflict, and two conflicting stimuli are mapped onto two different responses, in order to elicit response-based conflict (e.g. Soutschek, Taylor, Müller, & Schubert, 2013; van Veen & Carter, 2005). However, the latter will also include a degree of stimulus-based conflict and, in the former, it is possible that errors might be coded as correct responses simply because they are mapped onto the same response. At the moment, evidence for and against these types of functional subdivisions within the ACC is pretty equally divided (for a very recent review see Gasquoine, 2013). Resting-state fMRI analysis of ACC connectivity has revealed that while more rostral portions of ACC show higher connectivity with limbic and paralimbic regions and caudal portions with sensorimotor regions, these connectivity points seem to lie on a continuum with intermediate ACC portions being connected with lateral prefrontal areas implicated in complex cognitive functions (Margulies et al., 2007). Rostral and caudal ACC subdivisions, however, are also very densely interconnected with each other (Yeterian et al., 2012). It is therefore likely that a clear-cut dissociation of functions between ACC subdivisions, one that is based on task content, may not be

possible. Instead, these subdivisions might interact by providing different types of information to be used for more general processes.

dlPFC: THE CONTROLLER

dlPFC anatomy and connectivity

In the human brain, dlPFC (Figure 5) occupies parts of the superior and middle frontal gyri. It comprises Brodmann areas 46 (lying on the middle portion of the middle frontal gyrus), 9/46 (situated in the middle frontal gyrus) and 9 (occupying parts of the superior frontal gyrus).

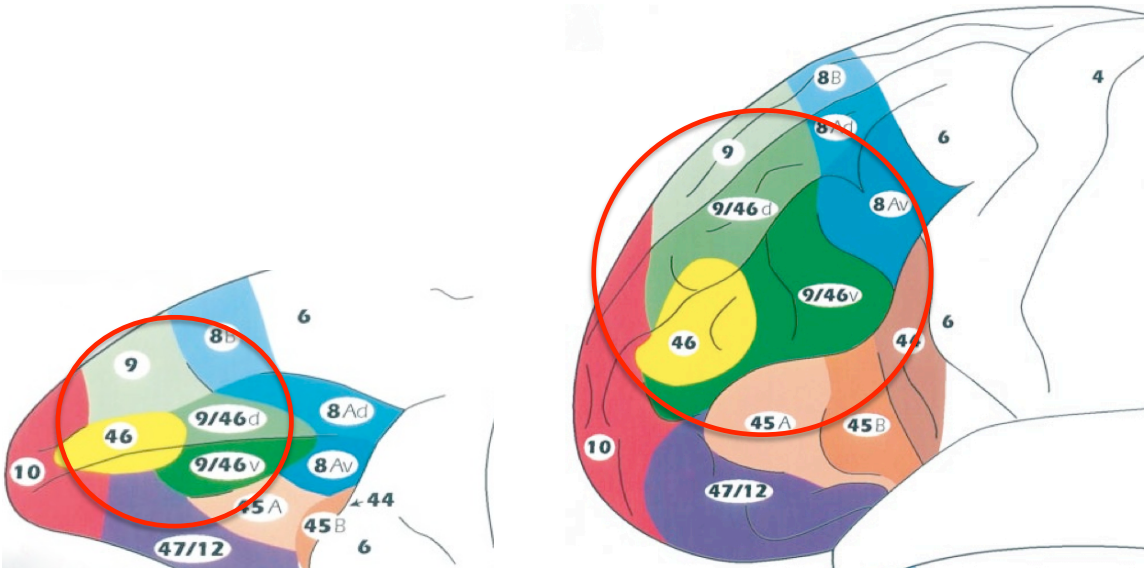


Figure 5. *Subdivisions of the lateral Prefrontal Cortex (PFC) in the human (right) and macaque (left) brain. The dorsolateral Prefrontal Cortex (dlPFC, indicated by the red circle) encompasses Brodmann areas 46, 9/46 and 9. Figure adapted from Petrides & Pandya (1999, pp. 1015).*

dlPFC is rich in pyramidal neurons, like other PFC areas. At the cytoarchitectonic level, area 9 has a less compact layer II, a less dense layer III, and a more dysgranular layer IV than area 46/9. Area 46 has a well developed layer IV and III and more uniform appearance than either of the other two dlPFC areas (Petrides & Pandya, 1994), and appears less myelinated (Rajkowska & Goldman-Rakic, 1995). Spindle neurons, like those found in the ACC, have been recently reported in area 9, although in lesser number than in the ACC (Fajardo et al., 2008).

The dlPFC is interconnected with a very large number of cortical and subcortical structures. It is connected to motor areas both in the frontal cortex (Supplementary Motor Areas, Premotor Areas and the Frontal Eye Fields) and in subcortical areas (Superior Colliculi). It receives inputs from all sensory, somatosensory and multimodal areas as well secondary-association areas via long association bundles in the superior longitudinal fasciculi and the orbitofrontal fasciculus. It has reciprocal connections with the ACC and other paralimbic areas via the cingulate fasciculus (Yeterian et al., 2012) and with the basal ganglia (Masterman & Cummings, 1997). The different subdivisions within dlPFC generally present very similar interconnections with other cortical areas, albeit with some differences: area 9/46 has less connectivity with auditory areas compared to the other two areas, but more connections with visual occipital association areas (Yeterian et al. 2012), whereas area 9 has less connectivity with parietal areas (Petrides & Pandya, 1999).

There has been debate across the decades regarding the specific delineation of areas 9 and 46 in the monkey brain as compared to the human brain (for a more detailed discussion, see Petrides & Pandya, 1994; Petrides, Tomaiuolo, Yeterian, & Pandya, 2012). Area 46 in the monkey brain has been restricted to the more rostral parts of the cortex that lies on the principal sulcus, with area 9/46 lying posterior to this, as opposed to dorsally, like it is the case in the human brain (Petrides et al., 2012).

dlPFC as the ideal 'controller'?

As mentioned earlier, lateral PFC regions are widely believed to play a fundamental role processing and integrating a wide range of information in order to implement top-down control on behaviour, thanks to their cellular profile and extensive reciprocal connections with motor and high-order association areas, as well as with performance-monitoring and reward-sensitive areas like the ACC and the basal ganglia. Therefore, one would expect these regions to be particularly engaged when the need for cognitive control is higher, such as in the context of conflict detection as per the CMCC model. Imaging studies have indeed consistently reported lateral PFC activation, and more specifically dlPFC activation, in the same type of conflict tasks that have been shown to activate the ACC, and particularly during adaptation trials (Durston, 2003; Egner & Hirsch, 2005; Kerns, 2006; Kim et al., 2012), suggesting a role for this region as the 'controller' to the ACC's 'monitor'.

There are three broad categories of functions attributed to dlPFC that make it a very plausible candidate for the role of 'controller' in conflict tasks. Temporal functions that involve maintaining and manipulating information across time (which are essential to support adaptation from one trial to another), representational functions that involve the acquisition and encoding of goals, rules and instructions (which are essential to guide behaviour according to internal goals) and more direct control functions that involve the suppression and inhibition of irrelevant information or responses (which are essential to actively ensure that contextually advantageous responses win the competition with contextually disadvantageous responses).

Temporal functions of dlPFC: working memory

Perhaps the most consistent finding about dlPFC function is its involvement in processes that include the maintenance, or processing, of information across time delays. As illustrated earlier, cells in dlPFC have been found to be actively firing across delays (Bunge et al., 2003; Fuster & Alexander, 1971; Goldman-Rakic, 1995; Mansouri, et al., 2007; Miller, 1996). For example, Bauer and Fuster (1976) found that cooling dlPFC (therefore causing a temporary lesion) caused deficits in tasks involving delayed matching-to-sample or delayed response. In this same region, Miller (1996) later found cells that showed heightened activity after the disappearance of the sample in the same type of task. As conflict-induced adaptation involves adjusting response behaviour from one trial to another, these properties of dlPFC cells are particularly relevant. Indeed,

Mansouri and colleagues (2007) found that dlPFC cells in the macaque maintained information about the level of conflict on the previous trial across the inter-trial interval.

Representational functions of dlPFC: learning and rules

Several studies have shown that dlPFC encodes rules, and is necessary for their implementation, ranging from basic stimulus-response associations (Asaad et al., 2000; Asaad et al., 1998) to more complex abstract rules (Buckley et al., 2009; Christoff & Keramatian, 2007; Mian et al., 2012; Wallis et al., 2001; Wallis & Miller, 2003). Therefore dlPFC plays a fundamental role in driving goal-driven or instructed behaviour. dlPFC is particularly important during the acquisition of a new rule, strategy or behaviour, with higher dlPFC involvement during learning or practice, than during the implementation of an established motor skill or familiar response strategy, and with severe impairments in rule-acquisition after a lesion (Boettiger & D'Esposito, 2005; Lungu, Binenstock, Pline, Yeaton, & Carey, 2007; Pasupathy & Miller, 2005; Petrides, 2007; Poldrack et al., 2005). Learning and motor-skill acquisition are characterized, especially during the earlier stages, by great uncertainty, where conflicting information and distractors prevail, until the agent has acquired sufficient exposure to the current contextual contingencies. Indeed, dlPFC activity has also been related to the suppression of distractors and of irrelevant responses, which is a feature of cognitive control that is essential for learning.

Control functions of dlPFC: inhibition and suppression

The dlPFC has been implicated in reversal learning, which requires subjects to suppress a previously overpracticed response that is no longer advantageous (Jahanshahi et al., 1998; Kadota et al., 2010; Knight, Staines, Swick, & Chao, 1999; Knoch, Brugger, & Regard, 2005; Mitchell, Rhodes, Pine, & Blair, 2008; Treichler, 1973), and patients with lesions to dlPFC are impaired at suppressing distractors (Chao & Knight, 1995; Godefroy & Rousseaux, 1996; Knight et al., 1999). Some neurophysiological studies have revealed a surface negative potential, with a possible source in monkey dlPFC, that seems to be related to the inhibition of responses to non-reinforced stimuli (Sasaki & Gemba, 1986). Inhibitory control by dlPFC might be mediated by inhibitory GABA neurons in this region (Barr et al., 2009; Wang, 2006).

dlPFC and conflict-monitoring: abundance of imaging evidence, but is dlPFC necessary?

The aforementioned temporal and representational functions ascribed to dlPFC, combined with the inhibitory properties of dlPFC, make a role for this region during conflict tasks (which are heavily dependent on the implementation of instructed behaviour in the face of distractors and competing responses, and particularly for adaptation across trials) particularly likely.

Indeed, dlPFC has been found to be activated in response to conflict across a variety of tasks, such as the Stroop (Fan, 2003; Kerns et al., 2004; Kim et al., 2012; MacDonald, 2000; Milham et al., 2001; Vincent van Veen & Carter, 2005), Simon (Kerns, 2006; Wittfoth et al., 2009) and Flanker (Durston, 2003; Fan, 2003; Luks et al., 2007; Wittfoth et al., 2009) tasks, which have all been shown to activate the ACC. However some subtle, but important differences in the conflict-related activation of these two regions has been observed. While ACC is more active on the first conflict trial in a sequence (i.e. when there is a change in conflict levels from low to high, indicating conflict detection) than on subsequent conflict trials (see Durston, 2003; Kerns, 2006; Kerns et al., 2004), dlPFC shows the opposite pattern of activation, with higher engagement on the second conflict trial in a pair (Egner & Hirsch, 2005a). Given the temporal dynamics of this activation, with ACC preceding dlPFC activation, the fact that both ACC and dlPFC activation was found to correlate with good behavioural adaptation (Durston, 2003; Egner & Hirsch, 2005; Kerns, 2006; Kim et al., 2012) and that ACC activation was found to reliably predict dlPFC activation on subsequent trials (Kerns, 2006), the imaging literature very strongly implicates dlPFC as the 'controller' that responds to ACC signalling in order to implement cognitive control and drive behavioural adaptation, as suggested by the CMCC model.

While there is abundant imaging evidence regarding the role of dlPFC in conflict-induced behavioural adaptation, most of the neuropsychological literature of conflict

has focused on the ACC. Therefore, evidence as to whether or not dlPFC is necessary for conflict-induced behavioural adaptation is comparatively scarce.

As some of the tasks used in the conflict-monitoring literature, such as the Stroop tasks and the WCST, are part of the standard battery of tests used to asses the consequences of prefrontal damage on cognition, there is evidence of the effects of dlPFC damage on the 'incongruent' (i.e. conflict) versions of these tasks. For example, damage to the dlPFC has been associated with impairments in the Stroop task, when patients are required to name the colour (see Gläscher et al., 2012; Stuss, Floden, Alexander, Levine, & Katz, 2001). However, there is little to no evidence about the effects of dlPFC damage on conflict-induced adaptation in humans, as this is a rather specific process that most assessments of prefrontal damage are not particularly concerned about. Using a conflict analogue of the WCST, Mansouri and colleagues (2007) confirmed the crucial role of dlPFC for adaptation in the monkey: targeted lesions to the macaque dlPFC canceled out the adaptation effects observed in control animals.

Conflict-induced adaptation is a process that occurs across the span of two or more trials, and the imaging literature has shed some light on the temporal dynamics of ACC and dlPFC activation during this process. Neuropsychological studies can only determine whether or not an area is necessary for the overall process, not the timing at which it is crucial. Furthermore, lesions in human patients are often widespread to

regions other than the dlPFC, making localization of function more difficult. This is an issue that can be investigated using Transcranial Magnetic Stimulation (TMS), which can be used to induce 'virtual lesions' in a target area at specific points in time during a task. Due to the biophysical limitations of TMS, selectively targeting ACC is a problematic task, due to its deep location within the cortex (see Chapter 2 for more details). But dlPFC, being lateral and more superficial, is a perfect candidate structure for this type of methodology, and has indeed been targeted very commonly in the literature in order to investigate a wide variety of phenomena, including conflict tasks (Ko, Monchi, Ptito, Petrides, & Strafella, 2008; Vanderhasselt et al., 2007; Vanderhasselt, De Raedt, Baeken, Leyman, & D'haenen, 2006; Wagner, Rihs, Mosimann, Fisch, & Schlaepfer, 2006). Despite this, it has very rarely been applied specifically to the investigation of conflict-induced adaptation. To the best of my knowledge, only Sturmer and colleagues (Stürmer, Redlich, Irlbacher, & Brandt, 2007) have done so, and they found that TMS to the dlPFC canceled out the adaptation effects in the Simon task only when this was restricted to a time-window of 500-300 ms prior to the presentation of stimuli, indicating that dlPFC is unlikely to aid adaptation by function of maintaining information in working-memory across trials, but might do so via other mechanisms, such as reactivating recent task representations or enhancement of task relevant information. There is therefore a lot of scope to investigate the role of dlPFC in conflict-induced behavioural adaptation using TMS.

FPC: THE EVALUATOR

FPC anatomy and connectivity

The FPC is approximately synonymous with Brodmann area 10, and, as the name suggests, occupies the frontal pole (Figure 6). This region extends into the most anterior portion of the frontal gyri and has the most anterior coronal plane in which the three frontal gyri are present as its posterior border. It constitutes the largest area of the prefrontal cortex, making up 1.2% of the total brain volume in humans (Semendeferi, Armstrong, Schleicher, Zilles, & Van Hoesen, 2001). Several subdivisions have been proposed within the FPC, the most common being between lateral and medial FPC, based on cytoarchitecture, connectivity and attributed functions (Bludau et al., 2013). Gilbert and colleagues (Gilbert, Spengler, Simons, Frith, & Burgess, 2005; Gilbert et al., 2006), in a metanalysis of over a hundred imaging studies, found that lateral FPC is more active during working memory and episodic memory tasks, and its activity is associated with slow response times, whereas medial FPC is more active during tasks requiring mentalizing (i.e. processing internal emotional and mental states), and its activity is associated with faster response times.

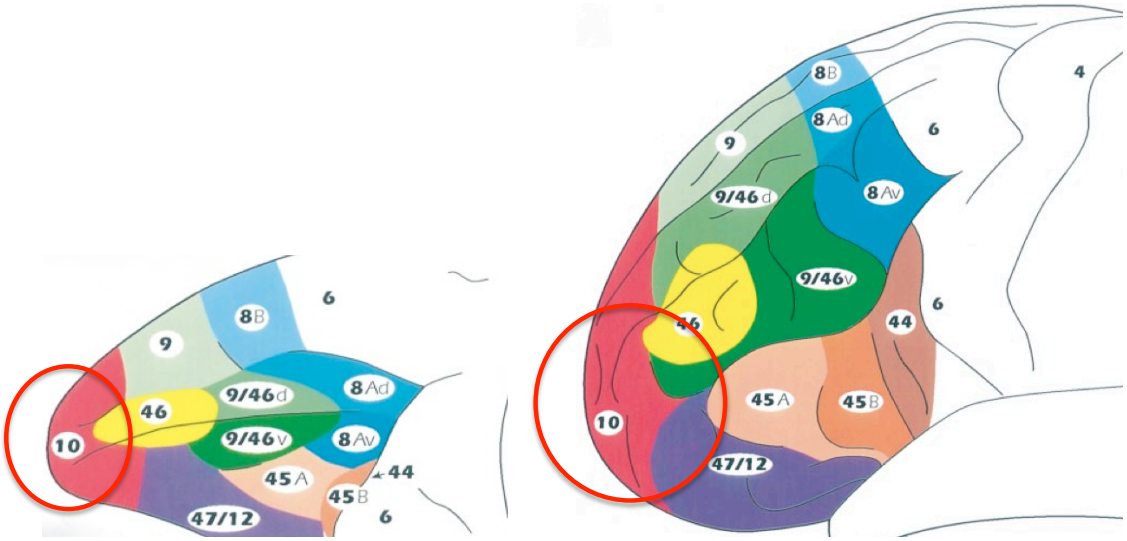


Figure 6. *The Frontopolar cortex, in the red circle, in the monkey (left) and human (right) brain, corresponds approximately to Brodmann area 10. Figure adapted from Petrides & Pandya (1999, pp. 1015).*

FPC is characterized a homogeneous population of pyramidal and granular cells (Petrides & Pandya, 1994). Compared to other cortical areas, neurons in the FPC have high number of dendritic spines (69% greater than primary somatosensory cortex) and dentrites are longer (31% longer than primary somatosensory cortex). However, the density of cell bodies is lower (Jacobs et al., 2001). This makes the FPC very suitable for establishing higher numbers of connections than other areas.

Like dlPFC and the ACC, FPC has very dense local connectivity with most other higher-order prefrontal areas. In particular, it presents a very high number of terminations of local fibres originating in area 32 of the ACC and area 9 of the dlPFC, as well as from OFC and some prefrontal motor areas (Burman, Reser, Yu, & Rosa, 2011). The FPC has bidirectional connections with the ACC and the retrosplenial cortex, via the

efferent connections of the dorsomedial limbic stream that runs through the cingulate fasciculus to target the anterior and posterior cingulate cortices (Petrides & Pandya, 2007) and via the afferent projections from ACC spindle neurons (Allman et al., 2002). It also receives inputs from more posterior association areas via the uncinate fasciculus and external capsule, such as the auditory and multimodal association areas in superior temporal gyrus and sulcus and somatosensory parietal areas, as well as paralimbic structures. More specifically, lateral FPC is rich in connections with limbic and paralimbic cortical areas, whereas medial FPC receives many afferents from somatosensory and premotor areas. However it lacks connections with more primary sensory areas (Petrides & Pandya, 2007; Yeterian et al., 2012; Ramnani & Owen, 2004). Subcortically, it receives projections mainly from the claustrum (which mostly seem to target medial FPC) and the mediodorsal thalamic nucleus, with also some sparse projections from the hypothalamus and the amygdala (Burman, Reser, Richardson, et al., 2011).

Overall, there are fairly consistent similarities between human and monkey FPC (Petrides & Pandya, 1994), making comparison more reliable than the previously discussed prefrontal regions. Nonetheless, differences in the organization of the FPC between the human and monkey brain have been noted. The human FPC is proportionally larger than the monkey FPC and is characterized by a lower cellular density, allowing more space for connections (Semendeferi et al., 2001). In terms of correspondence, in both the human and monkey brain, area 10 can be divided into a medial (10m) and orbital (10o) portion, with a further rostra-caudal subdivision within

10m (10r and 10c)(Ongür, Ferry, & Price, 2003). Human 10o is greatly expanded compared to monkey 10o, but the two subdivision appear more similar between the two species than the medial portions of FPC (Semendeferi et al., 2001).

FPC as the evaluator?

So far, we looked at how ACC and dlPFC might constitute a monitor-controller network to ensure the implementation of cognitive control during conflict in order to avoid interference, from competing responses and/or distractors, and select the most contextually advantageous response. However, "contextually advantageous" is a relative term: some responses may be more likely than others to consistently result in a reward, because the environmental contingencies that will lead to a positive outcome following that response may be more stable or commonplace than others. Consequently, some responses will be advantageous in most situations and others will be advantageous only during specific situations. Furthermore, a response that is generally advantageous might not be so under all circumstances and permanent changes in environmental contingencies might turn a previously advantageous response into a disadvantageous one and viceversa. Conflict tasks, in a way, provide a perfect example of this type of scenarios: a behaviour that normally leads to a positive outcome (e.g. in most context, reading a word will be more useful than noticing the colour it is printed in; or responding to the location of an item before noticing its identity might be beneficial, for example if we are about to be hit by a car), no longer does, and therefore

other options need to be evaluated and explored. The cognitive system needs to be able to support this kind of flexibility. If ACC and dlPFC are involved in monitoring contextual demands and acquiring/applying contextually appropriate responses, respectively, one question concerns whether there is a part of the PFC that performs an explorative and/or evaluative function to initiate the acquisition and/or application of a different response. The FPC could fulfil this role.

FPC, exploration and option values

Imaging studies have shown that the FPC is particularly active during the early stages of learning (Koechlin, Danek, Burnod, & Grafman, 2002; Koechlin & Hyafil, 2007; Ramnani & Passingham, 2001; Strange, Henson, Friston, & Dolan, 2001). Learning, as previously illustrated, is also a process that highly engages the dlPFC. FPC is also particularly active during situations of high uncertainty or conflict (Badre & Wagner, 2004; Koechlin, Corrado, Pietrini, & Grafman, 2000; Milham et al., 2001; van Veen et al., 2001; Yoshida & Ishii, 2006), which, as we saw, is a context that has been shown to also activate the ACC. Uncertainty and learning are both contexts that require alternatives to be explored and evaluated in order to select appropriate behavioural strategies, given the current environmental contingencies. In a meta-analysis involving 165 patients with focal lesions to PFC, Glascher and colleagues (Gläscher et al., 2012) identified a network involving the ACC and the dlPFC that is associated with conflict-monitoring and cognitive control, whereas a network including the orbitofrontal cortex and the FPC was

associated with value-based decision-making. Patients with FPC lesions have indeed be found to be particularly impaired in settings that requires them to deal with complex, uncertain and volatile contexts (Burgess, Veitch, de Lacy Costello, & Shallice, 2000; Dreher, Koechlin, Tierney, & Grafman, 2008). Accordingly, recent research has hypothesized a role for this region in the exploration of previously unchosen options and in encoding the value of unchosen options (Boorman, Behrens, Woolrich, & Rushworth, 2009; Boorman, Behrens, & Rushworth, 2011; Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006; Koechlin & Hyafil, 2007). This seems to indicate that FPC could be a good candidate region to fulfil the role of 'evaluator' to the ACC's 'monitor' and the dlPFC 'controller', to meet the demands of behavioural control and flexibility that are imposed on the cognitive system.

Competing theories: Hierarchical control, abstraction and mentalizing

On the other hand, a large number of other functions have also been linked to FPC, mostly due to the frequently reported activation of this region in a wide variety of complex cognitive tasks. Ramnani and Owen (2004) carried out an extensive evaluation of the existing frameworks of FPC function, and suggested that the most likely role for FPC, out of the ones proposed, concerned decision-making in contexts that require the integration of two or more cognitive operations (see Badre & D'Esposito, 2007; Badre, 2008; Christoff et al., 2001; Kroger et al., 2002; Reynolds, McDermott, & Braver, 2006).

However, this was before the role of FPC in the exploration and evaluation of alternatives had been proposed.

Furthermore, as it may be apparent by the studies cited so far, the vast majority of the evidence about FPC function comes from functional neuroimaging. While FPC may be involved in a large number of tasks, it may not be necessary for optimal performance in all of them. Also, the involvement of FPC across such a variety of tasks is likely to be an indicator of an underlying component shared by all tasks, rather than an indicator that FPC is carrying out a wide variety of processes. Such component should therefore be more sensitive to FPC damage than other components. As FPC is still the least well understood area of the PFC, there is scope to explore this region further by applying methodologies other than imaging, such as neuropsychological investigations. As FPC appears to be active during fairly complex tasks, it is likely that performance in these contexts is also reliant on several other structures. This makes reliance on human neuropsychology difficult, given the widespread lesions affecting several different areas that are usually found in patients. Nonetheless, given the relative similarities between the human and the monkey FPC, it should be possible to selectively target this region in the monkey brain in order to address whether FPC is necessary to support particular aspects of cognition.

CONCLUSIONS: A PREFRONTAL NETWORK FOR COGNITIVE CONTROL AND FLEXIBILITY

Given the cytoarchitectonic, connectivity and functional profiles of ACC, dlPFC and FPC, these three subdivisions of the frontal lobe seem to underpin a major part of what could be a prefrontal modulatory network to support cognitive control and behavioural flexibility in a changing environment. However, as discussed so far, research about each of these regions still presents several outstanding questions.

Although progress has been made to clarify the role of ACC as a structure that may alert the cognitive system for the need to implement cognitive control that results in behavioural adaptation, and we can confidently conclude that ACC is indeed somehow involved in this process, some questions still remain. Is human ACC, differently from the monkey, necessary for detecting conflict and drive adaptation? And if so, is this a mechanism that is sensitive only to specific types of conflict (i.e. perceptual vs response)? We will try to address this question in two studies: one investigating the effects of ACC lesions on performance in a conflict analogue of the Wisconsin Card Sorting Task that has been previously administered to macaques with lesions to the ACC. Second, investigating the effects of Deep Brain Stimulation on performance of the stimulus-based WCST analogue and the response-based Simon task, as well as recording intracranial Local Field Potentials from human ACC during performance of these two tasks.

There is also a lot of scope for investigating the role of human dlPFC in conflict-induced behavioural adaptation, both in the patient population and applying TMS. We will try to address this question in two studies: one investigating the effects of dlPFC damage on performance in a conflict analogue of the Wisconsin Card Sorting Task. And, the second, studying the effects of TMS on healthy participants during the same task, in order to investigate the consequences of more targeted, within-subjects, 'virtual lesions' specifically timed to coincide with adaptation periods.

Finally, we will address the issue of FPC function in a chapter investigating the effects of targeted lesions to the macaque FPC when learning to combine, and subsequently apply, two abstract rules. As the learning phase involves acquiring one rule after another has already been encoded, it represent a good scenario to test whether FPC is necessary in order to discover what is the new advantageous behaviour when the environmental contingencies suggest that the previous is no longer advantageous. At the same time, the paradigm requires the animal to perform the integration of two behaviours at a later time, which tests the hypothesis that FPC is necessary for cognitive and relational integration.

SCOPE OF THE THESIS

The present thesis will comprise six chapters, which aim to address a number of issues that have been raised in the course of the review of the literature.

CHAPTER 1 - Human neuropsychological investigations of the conflict WCST analogue

In the first chapter, we will investigate whether human ACC and dlPFC are necessary for conflict-monitoring and conflict-induced behavioural adaptation in a task where conflict is elicited between two instructed responses, using a conflict-analogue of the Wisconsin Card Sorting Test (Grant & Berg, 1948) .This is a task that has been widely used in the neuropsychological literature to investigate the effects of prefrontal damage on cognitive control (Milner, 1963; Stuss et al., 1982) and has been shown to heavily recruit both the ACC and the dlFPC in imaging studies (Buchsbaum, Greer, Chang, & Berman, 2005; Lie, Specht, Marshall, & Fink, 2006; Monchi, Petrides, Petre, Worsley, & Dagher, 2001; Specht, Lie, Shah, & Fink, 2009). Findings from targeted lesions to the macaque's ACC and dlPFC have shown that while the lesions to the latter only affect conflict-induced behavioural adaptation, lesions to the former affects neither conflict-related processing on the current trial nor adaptation on the subsequent trial (Mansouri et al., 2007). We will apply the same task to a cohort of patients with lesions to the ACC and dlPFC in order to provide a direct comparison between the animal and human literature. This chapter will help addressing the issue of mislocation between the human

and monkey ACC (Cole et al., 2009). Furthermore, this study represents the first time a conflict task where conflict is elicited between two competing instructed responses (as opposed to an habit and an instructed response) is administered to a human clinical population.

CHAPTER 2 – Effects of Transcranial Magnetic Stimulation on conflict-induced behavioural adaptation in the conflict WCST analogue

In the second chapter, we will investigate the effects of Transcranial Magnetic Stimulation of dlFPC on measures of conflict-induced behavioural adaptation, using the same WCST analogue that was administered to our clinical population. This method helps overcoming the limitations of neuropsychological investigations, by ensuring a higher degree of control over the location of the stimulation (i.e. of the 'virtual lesion') and providing a direct within-subject comparison to investigate whether stimulation affects a specific type of adaptation trials over another (i.e. the first or the second conflict trial in a pair of trials).

CHAPTER 3 – A comparative investigation into the possible underlying dynamics of conflict-induced adaptation effects in the conflict WCST analogue and classic conflict tasks

In the third chapter, we will describe a series of behavioural investigations into the dynamics of adaptation effects, in order to determine whether adaptation effects

might be driven by two dissociable components of cognitive control (caution and attention/task-relevant processing), which might be differentially contribute to adaptation depending on whether the tasks generate conflict between a habit and an instructed response (as, for example, in the case of the Simon and Flanker tasks) or between two instructed responses (as in the case of the WCST analogue). This is an aspect of conflict-induced adaptation that is often overlooked in conflict research, but could represent an important consideration when interpreting neuroimaging, electrophysiological and neuropsychological data obtained from these tasks. Furthermore, it can help shed some light on how the different routes to response (direct/habitual vs. indirect/instructed) may be differently modulated depending on task demands.

CHAPTER 4 – Quantification of conflict-induced adaptation components in the Linear Ballistic Accumulator model of choice response time

In the fourth chapter, we will test whether the two components (caution and attention/task-relevant processing) can be dissociated by the parameters of an evidence accumulation model of choice response time: the Linear Ballistic Accumulator (Brown & Heathcote, 2008). The quantification of these components as parameter values helps providing a more explanatory account for the modulation of the direct/habitual and indirect/instructed routes to response, as well as a series of parameters that could be applied in future research to investigate the correlation between different components

of adaptation and their underlying neurophysiological dynamics. This chapter represents one of the very first applications of this type of model to the investigation of conflict-monitoring and adaptation.

CHAPTER 5 – Effects of Deep Brain Stimulation on conflict-monitoring and conflict-related activity in Local Field Potentials recorded from the ACC

In the fifth chapter, we will look at a study involving the administration of the conflict WCST analogue and the Simon task to patients who have been fitted with Deep Brain Stimulation (DBS) implants in their ACC. This allows us to test the effect of stimulation on behavioural performance, as well as providing us with the rare opportunity to record Local Field Potential (LFP) data directly from human ACC. The methodology used to gather evidence in support of conflict processing in humans (mostly neuroimaging) often differs from that used in monkeys (mostly intracranial single-cell or LFP recordings). Intracranial recordings in humans are extremely rare (Cohen, Ridderinkhof, Haupt, Elger, & Fell, 2008; Davis & Taylor, 2005; Sheth et al., 2012). We hope to address a series of questions within this experimental chapter. First, does Deep Brain Stimulation of the ACC affect conflict and adaptation effects, and does it do so differently for a task that involves stimulus-based conflict between two instructed responses (WCST) as opposed to response-based conflict between a habit and an instructed response (Simon)? Secondly, can we observe conflict-related activity in Local Field Potentials recorded from human ACC, and, if so, does it differ between the

two tasks? Third, if any conflict-related activity is found, is it modulated by changes in conflict levels in a way that may be driving adaptation?

CHAPTER 6 – Effects of Frontopolar Cortex lesions on abstract rule acquisition and integration in the macaque monkey

In the final chapter, we investigate whether the FPC could play a fundamental role in behavioural flexibility by evaluating and exploring alternative options. If conflict resolution and adaptation requires the ability to consistently ignore distractors and task-irrelevant information and suppress competing responses, there are also settings where the exploration and evaluation of previously irrelevant information might be beneficial, for example in situations where it might represent a source of new, or greater, rewards or where it can guide the establishment of new behavioural strategies to substitute a behaviour that is no longer rewarded. We tested whether targeted lesions to the macaque FPC affect the acquisition of a response strategy that requires the integration of two abstract rules, its application or both. This allows us to test whether FPC is crucial for driving behaviour based on the integration of abstract, higher-order, information and/or the exploration, and evaluation, of alternative options in order to acquire a new response strategy.

CHAPTER 1

Human neuropsychological investigations of the conflict WCST analogue

Introduction

In the last decade, the conflict monitoring and cognitive control (CMCC) model developed by Botvinick and colleagues (Botvinick, Braver, & Barch, 2001; Botvinick, 2007; Carter & van Veen, 2007) has become the most influential framework for the interpretation of behavioural, imaging and neurophysiological data obtained from tasks involving interference between relevant and irrelevant dimensions of stimulus-response associations (i.e. conflict tasks).

In conflict tasks, subjects are required to respond to the relevant dimension of a stimulus (for example the identity of the stimulus), while ignoring an irrelevant dimension (for example, the spatial location of the stimulus). The overlap between the relevant and irrelevant dimensions is manipulated so as to produce either facilitation in the response (for example if the colour and the spatial location both cue the same response), or interference (for example if the colour and the spatial location cue different responses), creating congruent/low-conflict and incongruent/high-conflict trials, respectively. High-conflict trials usually result in performance costs, with slower response times and a higher error rate than low-conflict trials. The performance cost is

often referred to as congruency effect or conflict effect (Hedge & Marsh, 1975; Kornblum, Hasbroucq, & Osman, 1990; Lu & Proctor, 1995; Simon & Small, 1969; Simon, 1990; Simon, Sly, & Vilapakkam, 1981).

In order to perform the correct response during high-conflict trials, extra resources must be allocated to the relevant dimension. According to the CMCC model, the presence of conflict is detected by conflict-monitoring units, which then, in turn, bias cognitive control units that allocate the resources necessary to implement conflict resolution. Once these resources have been engage (i.e. following high-conflict trials), the effects of conflict on performance are reduced, resulting in faster speed of response and higher accuracy, a phenomenon known as sequential-effects or conflict-induced behavioural adaptation. The Anterior Cingulate Cortex (ACC) and the dorsolateral Prefrontal Cortex (dlPFC) have often been implicated in the detection of conflict and the implementation of cognitive control respectively. As such, they are considered critical components of the model.

Supporting evidence for the role of ACC and dlPFC in conflict monitoring and cognitive control comes particularly from imaging studies across a wide variety of tasks. ACC activation has been correlated with the level of conflict in a trial (Barch, Braver, Sabb, & Noll, 2000; Braver, Barch, Gray, Molfese, & Snyder, 2001; Chen, Wei, & Zhou, 2006; Durston, 2003; Egner & Hirsch, 2005; Fan, 2003; Hazeltine, 2003; Kim, Chung, & Kim, 2012; Kim, Kroger, & Kim, 2011; MacDonald, 2000; Peterson, Kane, & Alexander,

2002; Sohn, Albert, Jung, Carter, & Anderson, 2007; van Veen & Carter, 2005; Weissman, Giesbrecht, Song, Mangun, & Woldorff, 2003). Additionally, in a subset of the same studies, dlPFC activation correlated with the degree of cognitive control implemented to resolve conflict on subsequent trials, predicting good adaptation (Durston, 2003; Egner & Hirsch, 2005; Kerns, 2006; Kim et al., 2012).

Kerns and colleagues (2004, 2006) provide some of the most convincing evidence in support of the CMCC model, as the patterns of activation they describe very closely resemble the ones that would be predicted by the model, and they do so across two different tasks. In two separate studies, they found conflict-related activation in the ACC during the Simon and Stroop tasks. ACC activation was higher on high-conflict trials than low-conflict trials. High ACC activity predicted good adaptation on subsequent trials. ACC activity was lower on high-conflict trials preceded by another high-conflict trial (HH) compared to high-conflict trials preceded by a low-conflict trial (LH), indicating that ACC is more active when conflict is first detected and less active when a higher level of cognitive control has already been already engaged, such as in the case of HH trials. dlPFC, on the other hand, has been shown to be more active during good adaptation trials characterized by faster responses on HH trials compared to LH trials (Egner & Hirsch, 2005; Kerns, 2006), indicating that this region is specifically active when cognitive control is proactively engaged in order to reduce the effects of conflict on performance. Furthermore, dlPFC activation was found to significantly correlate with ACC activation on the previous trial (Kerns, 2006).

Some studies have also argued for a role of these areas in sustained performance. Durston and colleagues (2003) found that activation on high-conflict trials during a Flanker task increased in dlPFC and ACC as a function of the number of preceding low-conflict trials and decreased as a function of the number of preceding high-conflict trials. The offset in the timing of activation in the two regions (ACC during the early stage after stimulus presentation and dlPFC at a later stage), were interpreted as support for the role of ACC in detecting conflict and dlPFC in implementing control following ACC signalling. Horga and colleagues (2011), using a Simon task, showed that areas in the dorsomedial prefrontal cortex were positively correlated with the number of preceding high-conflict trials, whilst the pre-supplementary motor area showed a reversal in activation after more than 2 consecutive high-conflict trials.

The imaging literature therefore provides strong support for the CMCC model. However, while imaging studies can suggest whether a region is involved in a particular process, they cannot conclusively determine whether that region is necessary to support that process. As we outline below, evidence from lesion studies and from neurophysiology does not provide as strong a support in favour of the CMCC model as imaging does.

Neurophysiological studies have provided inconclusive evidence regarding the presence of conflict-sensitive neurons in the ACC. Ito and colleagues (Ito, Stuphorn,

Brown, & Schall; 2003) found no conflict-sensitive neurons in area 24 of the macaque ACC while the animals carried out a saccade-counterdemanding task. Similar results were obtained by Nakamura, Roesch, & Olson (2005), who found no conflict-sensitive neurons in the macaque ACC using the Simon task. Emeric and colleagues (2008) recorded local field potentials in the monkey ACC and did not find any correlation between conflict or adaptation and single-cell activity in a similar task. However, recordings from human ACC have indicated that some degree of conflict modulation is present in the neural population activity. In a count Stroop task, Davis and Taylor (2005) found that some cells are sensitive to the conflict levels in the task, although not exclusively to conflict. Modulation due to the amount of interference present in the current and previous trial was also found in single ACC neurons (Sheth et al., 2012). This might point to potential functional differences between monkey ACC and human ACC, as suggested by Cole et al. (2009), who argue that the absence of an homologue for human area 32' in the monkey brain might explain the discordant results from animal neurophysiology and human imaging research.

However, lesion studies in human patients also provide evidence that is inconsistent with the CMCC view. Some studies on patients with damage to ACC do indicate impairments in conflict tasks compared to controls (Cohen, Kaplan, Moser, Jenkins, & Wilkinson, 1999; di Pellegrino, Ciaramelli, & Làdavas, 2007; Sheth et al., 2012; Swick & Turken, 2002), while other studies show the opposite pattern of results, with no impairment in patients compared to controls (Fellows & Farah, 2005; Stuss, Floden,

Alexander, Levine, & Katz, 2001; Swick & Jovanovic, 2002; Vendrell, Junqué, Pujol, & Jurado, 1995). Most recently, Sheth et al. (2012) investigated the effect of dorsal ACC cingulotomy on a multiple interference task, and found that while cingulotomy did not affect performance during high-conflict trials, it did affect the patients' ability to adapt on subsequent trials. Fellows & Farah (2005) investigated the effects of lesion to dorsal ACC in four patients carrying out the Stroop task, and found no impairments in response speed or accuracy. Patients were overall slower at the task compared to controls, but there were no significant differences in performance on high conflict trials.

The evidence regarding the role of ACC in conflict monitoring and its modulatory activity on dlPFC to drive adaptation therefore appears not conclusively defined. Despite imaging evidence showing high ACC activation for current conflict trials, there is an indication that ACC may be playing a role in driving adaptation, rather than responding to conflict during the current trial.

Recently, evidence about the presence of separate networks to support resolution of different types of conflict has begun to emerge. Conflict can be elicited at different levels of cognitive processing, such as at the stimulus/perceptual level or at the response level (Egner, Delano, & Hirsch, 2007; Egner, 2008; Jia et al., 2007; Kim et al., 2012; Milham et al., 2001; Verbruggen, Notebaert, Liefooghe, & Vandierendonck, 2006; Wendt, Heldmann, Münte, & Kluwe, 2007). Tasks such as the Simon or Flanker tasks are generally thought to elicit response-based conflict, whereas tasks like the Stroop task

are thought to elicit stimulus-based conflict. Furthermore, conflict can be elicited when task-relevant and task-irrelevant information is mapped onto two separate dimensions (e.g., identity and location, in the case of the Simon task), or when it is mapped onto the same dimension (e.g., identity, in the case of the Stroop task). These have all been shown to activate the ACC during high-conflict trials, albeit some evidence suggests that different subdivisions of ACC are involved in processing response versus stimulus-based conflict (see van Veen & Carter, 2005; Kim et al., 2011; 2012). Studies have shown that conflict between semantic or conceptual representations can also recruit the ACC (Badre & Wagner, 2004; van Veen & Carter, 2005; Weissman et al., 2003).

Most of the aforementioned lesion studies used the Stroop task to investigate the effects of lesion on conflict monitoring. Another task that can be used to elicit stimulus-based conflict is the Wisconsin Card Sorting Test (WCST).

Mansouri and colleagues (Mansouri, Buckley, & Tanaka, 2007) investigated the role of the macaque ACC and dlPFC in a conflict analogue of the WCST. In this version of the WCST (see Figure 9 and Methods section for a detailed description of the task), the number dimension of the task is omitted and trials in which one of the possible choices is identical to the sample item are introduced, to provide a low-conflict condition. Animals were presented with a sample-item and had to match the sample to one of three test-items by applying one of two abstract rules (i.e. 'match by shape' or 'match by colour'), which alternated throughout the task. On low-conflict trials, one of the test-

items shared both the colour and the shape dimension with the sample. On high-conflict trials, one of the choice-items shared the same shape with the sample, creating competition between the two dimensions, and another shared the same colour (with a third item sharing neither dimension). Contrary to the CMCC's predictions, they found that the behaviour of animals with lesions to the ACC was unchanged both on high-conflict trials overall, and also on adaptation trials. In other words, they were still significantly faster on low-conflict trials than high-conflict trials (therefore showing a classic conflict effect), and their response time on high-conflict trials improved if the preceding trial was also a high-conflict trial (i.e. on HH trials, as compared to LH trials, therefore showing an adaptation effect). Animals with lesions to the dlPFC, on the other hand, were impaired in measures of adaptation, with the adaptation effect usually observed after high-conflict trials being abolished by the lesion. This is consistent with the idea that the dlPFC plays a crucial role in adaptation, but inconsistent with a fundamental role for ACC in conflict-monitoring.

In the current study, we sought to investigate the role of ACC and dlPFC in conflict-monitoring and adaptation, by analyzing the performance of neuropsychological patients, using the same WCST analogue used by Mansouri and colleagues (2007). By using the same experimental paradigm (both in terms of task and of effector systems) that has been used to investigate the effects of lesions to the analogue areas of the monkey brain, we can avoid some of the confounds caused by the use of different methodologies in human (e.g. neuroimaging) versus non-human (i.e. single-cell

recordings) primate studies, which are suggested to be partly responsible for the discrepancy between human neuroimaging literature and findings from animal work (Cole, et al., 2010; Schall & Emeric, 2010). This allowed us to provide a direct comparison with Mansouri and colleagues' study in order to investigate functional similarities or differences between the human and the monkey brain.

Furthermore, the WCST analogue presents an important difference when compared to other conflict tasks. That is, conflict is being elicited between two instructed behaviours ('match by colour' or 'match by shape'), rather than between a habitual/automatic response (e.g. 'read the word') that needs to be suppressed in favour of an instructed response (e.g. 'name the colour'). This will allow us to investigate whether ACC is crucial for optimal performance in situations where conflict is generated between two instructed responses, rather than when competition is between a habit and an instructed response.

Following on from Mansouri and colleagues' findings (2007), we hypothesized that if ACC is a crucial region to support conflict-monitoring and to bias conflict resolution, we should observe slower responses on high conflict trials, with a larger conflict effect in ACC patients than patients with lesions to other areas or controls. If ACC detects conflict in order to drive adaptation on subsequent trials, we should observe a smaller adaptation effect in ACC patients than in patients with lesions to other areas or controls. Our second hypothesis is that if dlPFC is a crucial region to support conflict-induced

behavioural adaptation in humans, as it has been shown to be in monkeys (Mansouri et al., 2007), we should observe a smaller adaptation effect in dlPFC patients than in controls or in patients with lesions to other areas.

Methods

Participants

27 participants (15 male, mean age 60.55 years) took part in the study. Of these, 9 participants were included the lesion groups. They were identified through the Medical Research Council Cognition and Brain Unit (CBU) volunteer database at the University of Cambridge and through the volunteer panel of the Psychology Department at the University of Birmingham. Three patients had damage to the ACC, three patients had damage to the dlPFC and three patients had damage to the OFC. The average period between the onset of the damage and testing was 131 months (range: 53-168 months). 18 control participants, matched on age, gender and National Adult Reading Test (NART) scores, were recruited through the CBU database, as well as from advertisement in the Oxford area. Control participants were not taking any psychoactive medication and were free of current or past neurological or psychiatric conditions, as determined by their history. All subjects had normal or corrected-to-normal vision. All subjects were native English speakers and provided written consent prior to their participation in the study.

Lesions

Lesions were traced from MR or CT images and mapped onto the standard Montreal Neurological Institute brain using MRIcro software. A detailed summary of the extent and location of the lesions is provided in the Appendix (Appendix 2a). The inclusion criteria for the lesion groups required that over 10% of the total lesion volume had to be localized in the Brodmann area of interest (see Appendix 2b).

ACC damage encompassed Brodmann areas 32 and 24. 1 patient presented damage to left ACC, 1 patients presented damage to right ACC and 1 patient presented bilateral damage to the ACC (see Figure 7 - for more detailed views of individual lesions, see Appendix 1a).

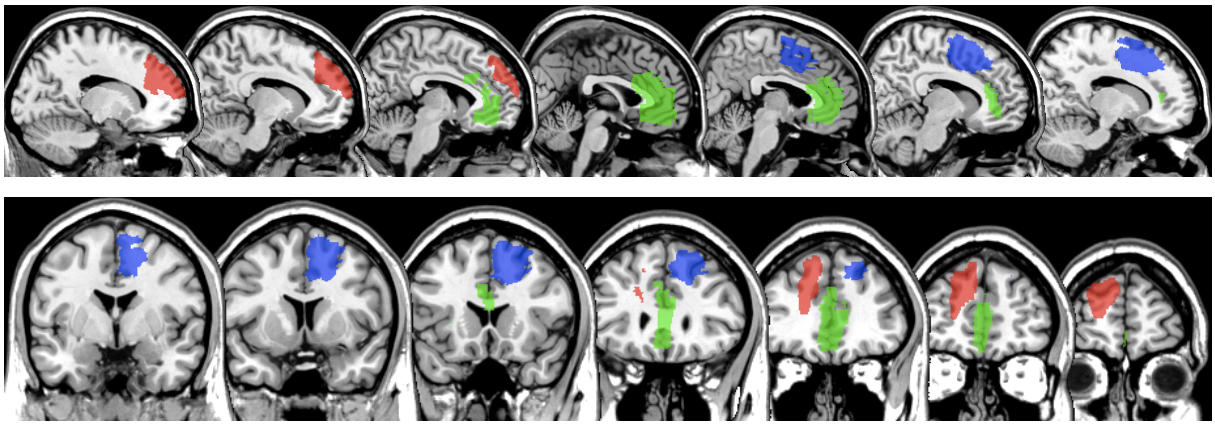


Figure 7. *Location and extent of ACC lesions in the three ACC patients included in this study, presented in a sagittal (top, from left to right hemisphere) and coronal (bottom, from posterior to anterior) view of a standard MNI brain. Each color denotes a different patient. Patient 1 = blue, Patient 2 = red, Patient 3 = green.*

dlPFC damage encompassed for the most part Brodmann area 46, and, to a lesser extent, area 9. Two patients presented damage to left dlPFC and one patient presented damage to right dlPFC (see Figure 8 - for more detailed views of individual lesions, see Appendix 1b).

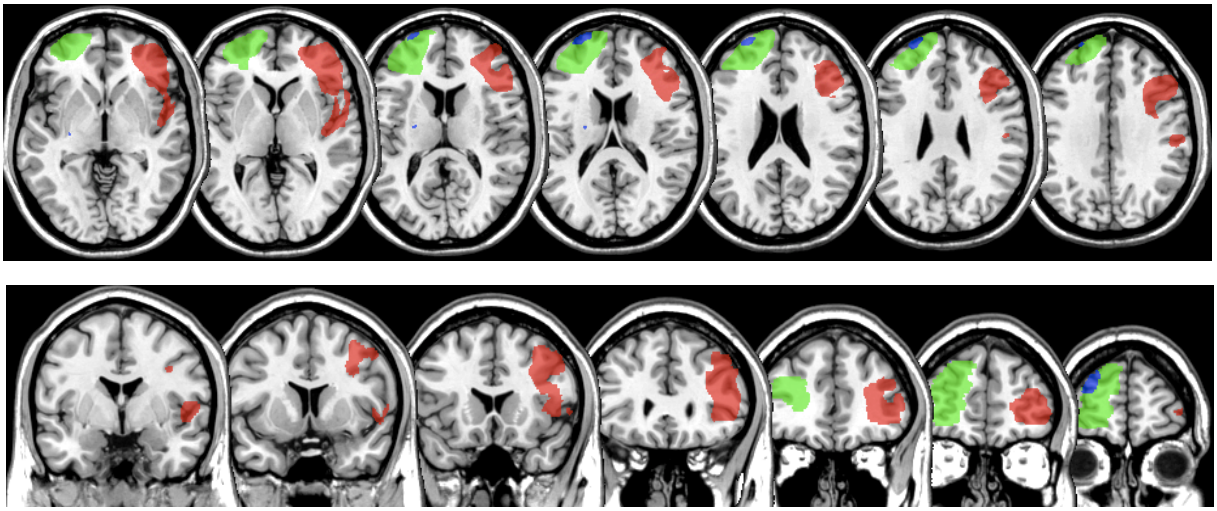


Figure 8. *Location and extent of dlPFC lesions in the three dlPFC patients included in this study in horizontal (top, ventro-dorsal) and coronal (bottom, posterior to anterior) views on a standard MNI brain. Each color denotes a different patient. Patient 4 = red, Patient 5 = blue, Patient 6 = green.*

OFC damage encompassed Brodmann area 11. 2 patients presented damage to right OFC and 1 patient presented damage to left OFC (see Appendix 1c).

The aetiologies included meningioma, aneurysm and encephalitis. A summary of the participants' background information, including aetiology of the damage, is included in Table 1.

Patient ID	Age	Gender	NART IQ	Area(s)	Volume (mm^3)	% of total volume	Aetiology
1	69	M	121.57	24+32R	6320	21	Meningioma
2	65	M	93.09	32L	4128	20	Meningioma
3	58	M	109.11	24+32	7315	19	Encephalitis
4	62	F	123.35	46R	5320	11	Aneurysm
5	66	M	117.12	46L	464	34	Aneurysm
6	54	M	120.68	46L	6760	19	Meningioma
7	57	F	124.24	11R	16520	22	Aneurysm
8	65	F	110	11R	19440	19	Aneurysm
9	46	F	119.79	11L	13088	22	Aneurysm

Table 1. *Patients' demographic information, including aetiology, total lesion volume and percentage of the total lesion volume localized to the specific areas of interest.*

Apparatus, stimuli and procedure

The task used in the study was a computerized analogue of the Wisconsin Card Sorting Test, which was originally developed by Mansouri and colleagues (2007). The task was programmed using Turbo Pascal (Borland), run in DOS on a desktop PC and presented on a 20.1" colour touchscreen (TFT LCD TS200H GNR), which was used to collect responses. Participant sat at a distance of 40cm from the screen and they were instructed to respond using the index finger of their dominant hand to touch the items on the screen.

The stimulus set consisted of six possible shapes (triangle, circle, square, hexagon, ellipse or cross), 2.4 degrees of visual angle in width 2.4 degrees in height, in six possible colours (red, green, blue, cyan, magenta or yellow), for a total of 36 possible combinations, and they were presented against a black background. The sample item

was always presented in the centre of the screen, and the test items were presented 2.6 degrees to the right, left and bottom of the sample item.

A typical trial was as follows. At the start of the trial, a random sample item was presented in the centre of the screen. Participants were instructed to touch the sample item when they were ready to start the trial. Once the sample item was touched, three test items appeared on the screen. Participants were instructed to carry out a matching-to-sample task, where the rule for matching could be either "match by shape" (i.e. pick the test item that shared the same shape as the sample item) or "match by colour" (i.e. pick the test item that shared the same colour as the sample item). All items remained on the screen until a response was made or until 10 seconds had elapsed. Correct trials were identified by a high-pitch sound, indicating positive feedback. Incorrect trials were identified by a low-pitched sound accompanied by the presentation of a large, gray circle, indicating negative feedback. After 2 seconds from response, another sample item appeared on the screen, indicating the start of another trial.

Conflict levels were manipulated by changing the degree of feature overlap between the sample and test items. In low conflict trials, one of the test items was identical to the sample item (i.e. shared both the same colour and the same shape), and the other two test items shared neither shape nor colour with the sample item. In high conflict trials, one of the test items shared the same colour as the sample item, whilst another test item shared the same shape, and a third item shared neither colour nor

shape with the sample item. High and low conflict trials were presented in randomized order. Examples of a high and a low conflict trial are presented in Figure 9.

Participants were informed that one rule would be 'correct' for several trials and then the other would be 'correct' for several trials with the correct rules switching unpredictably during the task such that they would have to periodically reassess which rule they needed to apply in order to perform the correct response. The rule switch occurred only once an accuracy criterion of 85% on the current rule had been reached.

Participants carried out the task for 12-15 minutes, for a total of approximately 150 trials. Reaction times and errors were recorded for analysis.

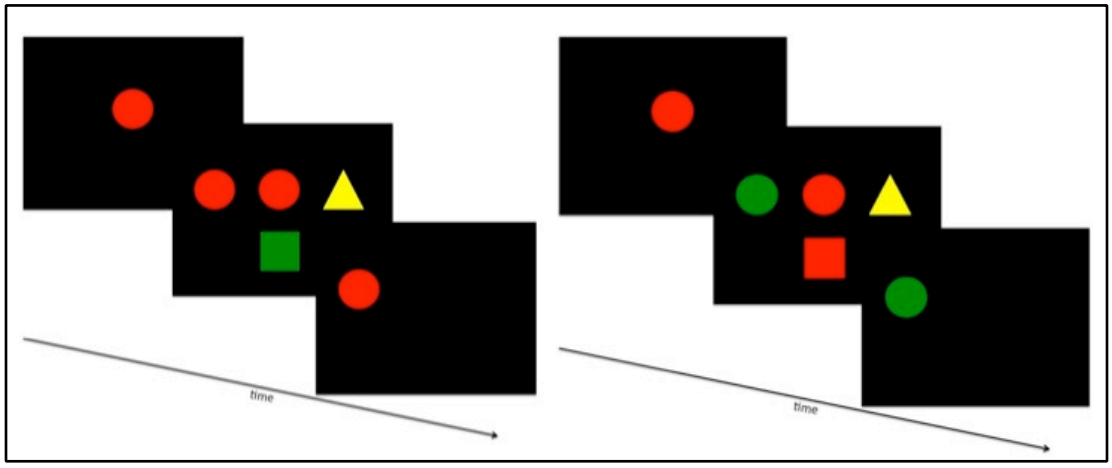


Figure 9. *Examples of a low-conflict trial (left) and a high-conflict trial (right), where the correct rule is 'match by shape', in the WCST analogue.*

Results

The Speed of Target Selection (STS) was computed by taking the reciprocal of the reaction time (RT) data in order to minimize the potential impact of occasional high response time values (without the need to remove these potential outliers from the analysis) on the RT distribution, as well as to use a unit of measurement consistent with Mansouri and colleagues' (2007) primate study, to which we wished to make direct comparisons. Therefore, it is useful to remember throughout that high STS values correspond to fast responses, and low STS values correspond to slow responses. STS analysis was carried out on correct trials only (and hence pairs of consecutive correct trials in the adaptation analyses). As the proportion of correct responses was very high, with the vast majority of errors restricted to switch trials (average number of errors that were non-switch trials was 4.75 for patients and 5.15 for controls) our analyses were carried out on STS only (and not accuracy, or number of switches).

In order to test the presence of conflict and adaptation effects within our samples, a 2x2 Repeated Measures ANOVA was carried out with Current Conflict (High and Low) and Previous Conflict (High and Low) as within-subject factors.

As expected, there was a significant main effect of Current Conflict ($F_{(1,26)} = 58.33$, $p < .001$), with Low-conflict trials showing larger STS values (mean = 1.11 , sd = .20) than High-conflict trials (mean = .97 , sd = .23). There was a significant main effect of

Previous Conflict ($F_{(1,26)}$ = 6.22, p = .019), with trials following High-conflict trials showing larger STS values (mean = 1.06 , sd = .21) than trials following Low-conflict trials (mean = 1.03, sd = .21) and importantly, a significant interaction between Current Conflict and Previous Conflict ($F_{(1,26)}$ = 13.95, p = .001). Together the above analyses confirm that both conflict and adaptation effects were present in human participants just as they were with non-human primates performing the same task (Mansouri et al., 2007).

Bonferroni-corrected paired-samples t-tests were carried out to investigate the significant interaction between Current and Previous Conflict. There was a significant difference between HH and LH trials ($t_{(26)}$ = 4.36, p < .001), but no significant difference between HL and LL trials ($t_{(26)}$ < 1). STS values for HH trials were larger (mean = 1.007, sd = .22) than for LH trials (mean = .94, sd = .23). This indicates that the adaptation effect is due to better adjustment on High-conflict trials following High conflict trials (HH), rather than differences in response speed on Low-conflict trials.

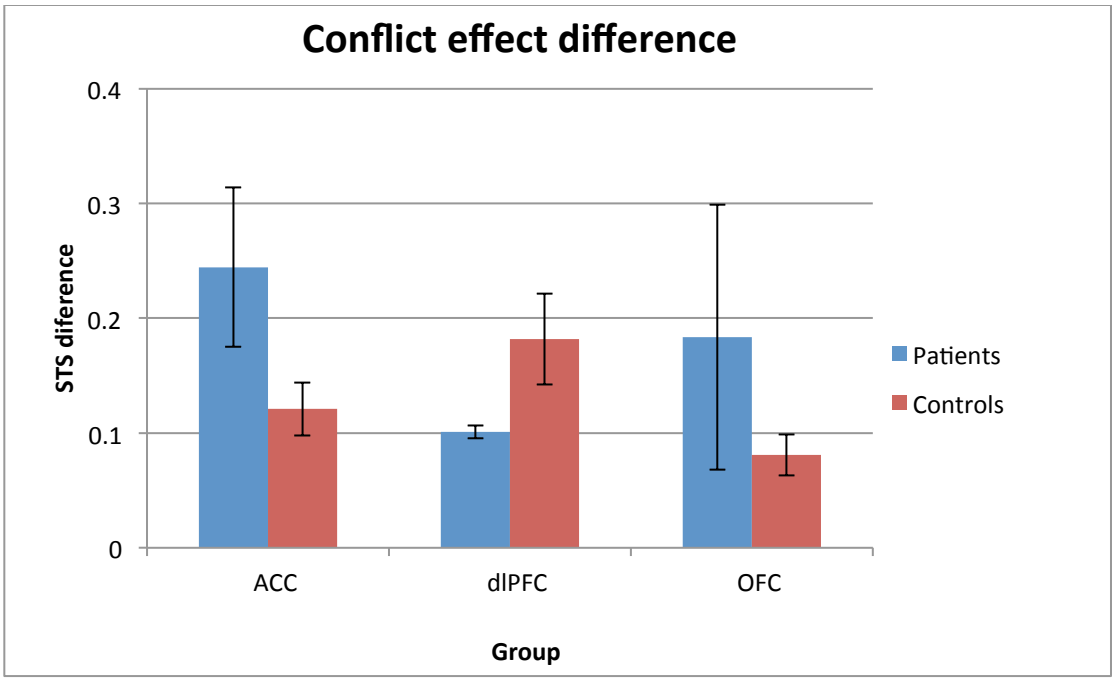


Figure 10. *Mean Speed of Target Selection difference between low- and high-conflict trials (i.e. conflict effect) for each patient group plotted against their respective matched control group. The presence of a conflict effect indicates faster responses on low-conflict compared to high-conflict trials.*

A series of non-parametric t-tests was then carried out to directly compare patient groups with their matched controls on their measures of conflict and adaptation effects. Wilcoxon Signed Ranks independent samples t-tests revealed no significant differences in the size of conflict effects between any of the patients groups and their corresponding matched controls. More specifically, the ACC patients did not significantly differ from their control in the size of the conflict effect ($z = 1.29$, $p = .197$), neither did the dlFPC patients compared to their controls ($z = -1.55$, $p = .121$), nor the OFC patients ($z < 1$). A conflict effect was present in all groups, with higher STS values on low-conflict than on high-conflict trials (Figure 10).

As the non-human primate study (Mansouri et al., 2007) indicated a clear difference in adaptation between the dlPFC and their control group, we predicted that there may be a similar difference (i.e. in the same direction, hence a 1-tailed t-test) in the adaptation effect (measured as the difference between mean HH and mean LH STS values) of dlPFC patients compared to their matched control group. A Wilcoxon Signed Ranks independent-samples t-test indicated a significant difference in the adaptation effect between the dlPFC group and their control group ($z = -1.81$, $p = .035$, 1-tailed test), with the control group showing significantly larger STS difference between HH and LH (mean = .0819 , sd = .063) than the dlPFC group (mean = -.0013 , sd = .026). There was no significant difference in the size of the adaptation effect between the ACC group and their controls ($z = -1.29$, $p = .197$), or between the OFC group and their controls ($z < 1$). These results are illustrated in Figure 11 and replicate the effects previously observed in non-human primates with dlPFC lesions on the same task (Mansouri et al., 2007).

In order to investigate whether the absence of an adaptation effect in the dlPFC group was due to effects of lesion on HH or LH trials, a 2-tailed Wilcoxon Signed Ranks independent-samples t-test revealed a significant difference in the STS between dlPFC and their control group for LH trials ($z = -2.32$; $p = .020$), with higher STS values for dlPFC patients (mean = .99, sd = .02) than the control group (mean = .80, sd = .08). There was a numerical difference in the speed of HH trials, with higher STS values for dlPFC patients (mean = .99, sd = .04) than controls (mean = .88, sd = .08), but it did not reach

significance (z = -1.81, p = .071). This indicates that the absence of an adaptation effect in patients with lesions to the dlPFC is due to faster responses on LH trials, which become as fast as HH trials.

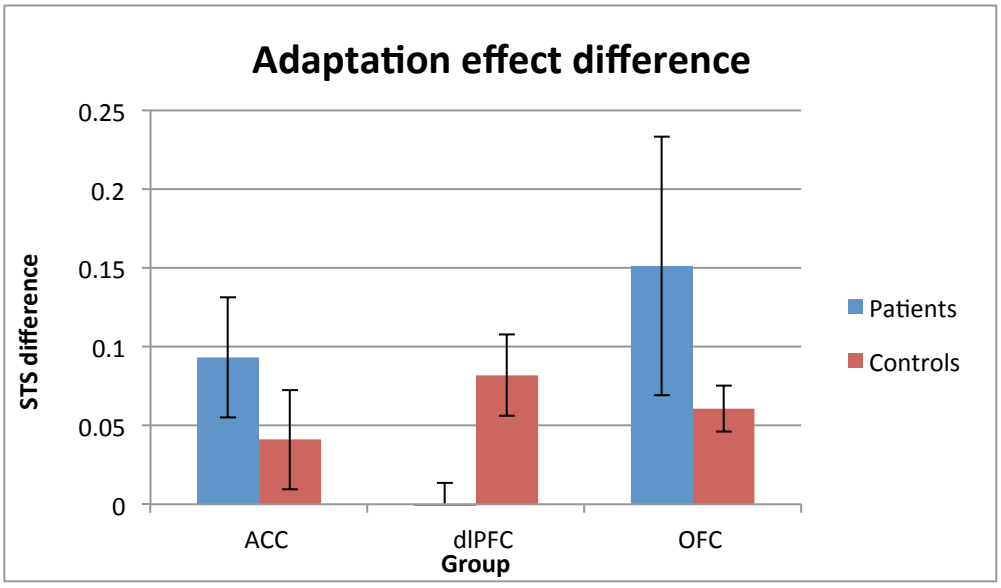


Figure 11. *Mean Speed of Target Selection difference between HH and LH trials (i.e. adaptation effect) for each patient group plotted against their respective matched control group. The presence of an adaptation effect indicates faster responses on HH trials compared to LH trials.*

Discussion

In this study, we aimed to investigate the role of the ACC and dlPFC in conflict-monitoring and conflict-induced behavioural adaptation using a conflict version of the WCST analogue.

We first investigated whether ACC lesions affected speed of response on high-conflict trials. Consistently with Mansouri and colleagues' (2007) findings in non-human

primates, our findings indicate that patients whose lesions include the ACC are not impaired in their responses to high-conflict trials, as compared to their matched controls. This confirms that the human ACC is, like non-human primate ACC, not generally necessary to respond to conflict on the current trial in all tasks designed to assess this.

The absence of impairment in the ACC group on high conflict trials is perhaps unsurprising. Most of the lesion studies that investigated the role of the ACC in responding to high-conflict trials and subsequent adaptation found either no impairment (Vendrell et al., 1995; Stuss et al., 2001; Swick & Jovanovic, 2002; Fellows & Farah, 2005) or impairment restricted to adaptation (diPellegrino et al., 2007; Sheth et al., 2012). diPellegrino and colleagues (2007) found that the conflict effect in a Simon task was significantly larger in the patient groups than in the control group. However, there was no significant difference amongst patient groups, regardless of the specific region affected by the lesion. They instead found that adaptation was affected by ACC damage. These findings were recently replicated in a study on ACC cingulotomy (Sheth et al., 2012). This might indicate that while the ACC might be sensitive to the level of conflict in the task, as demonstrated by imaging, it is not involved in conflict resolution on the current trial and/or that the areas that are involved in conflict resolution on the current trial are not dependent on ACC alone for their performance (see Carter et al., 2000). However, our ACC patients were also not impaired on measures of conflict-induced behavioural adaptation. While this, again, confirms the findings from the non-

human primate study (Mansouri et al., 2007), it is at odds with the literature showing impairments in adaptation following ACC damage (Swick & Turken, 2002; DiPellegrino et al., 2007; Sheth et al., 2012).

One possible explanation for this finding might be related to the specific conflict task we used. Tasks commonly used to investigate conflict-monitoring (e.g. Simon, Flanker and Stroop), include low-conflict trials that can exploit a habitual and/or overpracticed response (e.g. word-reading, or responding to the location of a stimulus) in order to aid performance. In the WCST analogue, just as in other conflict tasks, low-conflict trials can also be solved without the need to retrieve rules or specific stimulus-response associations. However, unlike in classic conflict tasks, there is no habitual response that can aid performance on these trials in the WCST. Conflict in the WCST is elicited between two instructed responses, rather than a habitual response and instructed responses. Studies that used neutral trials (i.e. trials that do not involve an element of facilitation due to the exploitation of a habit), found neutral trials to be generally slower than low-conflict trials (e.g. Milham et al., 2001; Whur & Ansorge, 2005; Galashan et al., 2008; Chen & Melara, 2009), and reported that adaptation after neutral trial, while present, is smaller than after high-conflict trials (Whur & Ansorge, 2005). Also, higher ACC activation is found for congruent and incongruent trials than for neutral trials (Milham & Banich, 2005). This indicates that ACC damage might lead to impairments in adaptation only in tasks that involve an element of facilitation due to the possibility of exploiting habitual responses (and therefore in contexts that require

suppression of these habitual responses on high-conflict trials). This could be due to the fact that, if absence of facilitation means that low-conflict trials are relatively slower in the WCST, the drop in speed of response from a low-conflict to a high-conflict trial (i.e. during LH trials) might not be as pronounced as in other tasks, and, consequently, the difference in speed of response between HH and LH might not be as large. Therefore, ACC damage might be more evident when the difference between HH and LH is larger. Another possibility is that that adaptation effects in tasks involving facilitation might be due to adjustments in performance during low-conflict trials, rather than adjustments in performance during high-conflict trials, and that the role of ACC is to modulate the former, rather than the latter. If this were the case, since adaptation in our task was exclusively restricted to an enhancement in performance on high-conflict trials, ACC damage would not be expected to have any effects on adaptation. We will further explore the task-specific dynamics of adaptation in Chapters 3 and 4.

The similarity in the results obtained from the current study and from the non-human primate study by Mansouri and colleagues (2007) helps inform the debate regarding similarities or differences between human and monkey ACC anatomy, both in terms of cytoarchitectonic discrepancies between the two species (Cole et al., 2009) and in terms of methodological differences limiting the reliability of comparisons between studies across species (Cole et al., 2010; Schall & Emeric, 2010). Although human lesion studies, as opposed to animal neuropsychological investigations, are often limited by the impossibility to control for the precise localization of the lesion, comparing the

performance of a neuropsychological population to that of animals with lesions to the analogue areas in such a direct manner, using the same task and the same effector systems (touchscreens) overcomes many of the limitations that are normally encountered when trying to reconcile findings from different methodologies across species. Our findings of indicate that, at least as far as the conflict WCST analogue is concerned, anatomical differences between the human and monkey brains are unlikely to account for the contradictory evidence regarding the role of ACC that has been obtained from imaging as opposed to neurophysiology and lesion studies.

One area that did affect the patients' ability to adjust their behaviour on high-conflict trials, as a function of previous trial type, was the dlPFC. In patients with dlPFC lesions, there was little to no difference between high conflict trials preceded by high-conflict trials (HH) and high-conflict trials preceded by low-conflict trials (LH). This is consistent with the non-human primate findings (Mansouri et al., 2007), and also with the body of literature implicating the dlPFC as an important structure for the implementation of cognitive control (Egner & Hirsch, 2005; Kerns et al., 2004; Kerns, 2006; Durston et al., 2003; Kim et al., 2012).

An interesting finding regarding the dlPFC lesion group was that the reduction in the adaptation effect seemed to be due to LH trials being faster in the lesion group compared to their controls, as opposed to HH trials being slower. If we assume the role of dlPFC is to implement sustained enhanced cognitive control, as suggested by the

increase in dlPFC activation on good adaptation trials shown by imaging (Kerns, 2006; Egner & Hirsch, 2005), we would expect patients with dlPFC lesions to be slower on HH trials, rather than faster on LH trials. HH trials do seem affected by dlPFC lesions (although non-significantly), but through also an increase, rather than decrease, in speed, compared to the control group. The results from the effect of dlPFC lesions on adaptation therefore indicate that dlPFC might be sensitive to the change in conflict level, from low- to high-conflict, rather than involved in sustaining cognitive control across trials in order to proactively prevent the detrimental effects of conflict on performance. Using the same task as in our lesion study, Mansouri et al. (2007) recorded single cell activity in the macaque dlPFC, and found that cells encoded information about the conflict level of a trial during the inter-trial period, suggesting that dlPFC might aid adaptation by maintaining information about recent conflict history online. They found that conflict-sensitive cells in dlPFC were sensitive to LH trials as well as HH trials. This is consistent with the idea that dlPFC might be encoding information about changes in conflict levels during the early stages after a change from low- to high-conflict, and it might maintain this information only for a brief period (as its activation decreases with increasing numbers of high-conflict trials, see Durston et al., 2004) in order to initially trigger adaptation mechanisms.

It is difficult to draw any specific conclusion about the finding of faster responses on LH trials from our neuropsychological population because the absence of pre-lesion and post-lesion data, and the small sample size, makes it unclear whether this might be

due to the effects of the dlPFC damage, or rather to other confounds. Similarly, there were no pre-op and post-op data to compare in Mansouri and colleagues' study (2007). We will further investigate this issue in the following chapter, looking at the effects of Transcranial Magnetic Stimulation on adaptation in the WCST.

To conclude, lesions to the ACC do not affect performance on high-conflict trials in a task eliciting stimulus-based conflict between two instructed responses, and not involving suppression of a habitual response. Lesions to the ACC also do not affect adaptation on the subsequent trial as a function of the level of conflict on the current trial. Lesions to dlPFC, on the other hand, do affect adaptation, by reducing the adjustment effect in patients compared to controls. These findings suggest that while the CMCC model might explain some aspects of conflict-monitoring and adaptation (e.g. dlPFC does play an important role in adaptation), it might overestimate the importance of ACC in modulating behaviour in response to conflict and might not adequately describe the similar, yet distinct, cognitive processes that may underlie adaptation. Additionally, the precise mechanism by which dlPFC modulates conflict-driven adaptation seems to be dependent on changes in conflict levels, rather than sustained recruitment of cognitive control across high-conflict trials.

CHAPTER 2

Effects of Transcranial Magnetic Stimulation on conflict-induced behavioural adaptation in the conflict WCST analogue

Introduction

The results of our lesion study in Chapter 1 indicate that dlPFC is a crucial structure to mediate conflict-induced behavioral adaptation in the WCST analogue, with lesions to this region abolishing the adaptation effect usually observed after high conflict trials. One of the limitations of lesion studies in humans is that, in the vast majority of clinical cases, the lesions are not localized exclusively to the region of interest and might involve, sometimes large, lesions to other brain areas. Whilst overlap across lesions in different patients helps to overcome part of this limitation, the level of control over this independent variable is relatively low. Low sample sizes also generally limit the power of the design. Furthermore, there is often no access to pre- and post-op data in order to assess the effects of lesions within-subjects and to control for the effects of plasticity and compensatory processes in the period following the lesion, which may vary from patient to patient. Control measures need to be collected from a separate group or participants.

Hence one valuable, complementary methodology is Transcranial Magnetic Stimulation (TMS). TMS allows the experimenter to interfere with neural activity in the brain, either on selected trials or whole blocks, in a way that has often been described as a temporary 'virtual lesion' (Pascual-Leone, Walsh, & Rothwell, 2000; Walsh & Cowey, 2000; Walsh & Rushworth, 1999). This is accomplished by generating a magnetic field, which, in turn, induces a perpendicular electric field that causes artificially synchronized activity in the target population of neurons, thus interfering with their function (Pascual-Leone et al., 2000). The effects of the magnetic field on the target population are heavily dependent on the intensity of stimulation and the distance between the scalp and the region of interest. Activation is stronger nearer to the source of the magnetic field (i.e. the scalp) and weakens as the distance between the coil and the target region increases (see Stokes et al., 2005; 2007; Trillenberg et al., 2012). Therefore, the intensity of the stimulation needs to be carefully adjusted in order to obtain the maximum possible interference effect, while, at the same time, ensuring the highest degree of control over the localization of the stimulation. Higher stimulator output is needed as the distance between the scalp and the target region increases, but high intensity results in less focal activation. On the other hand, too low stimulator output can result in understimulation, which could mask significant experimental effects, particularly at higher coil-cortex distances. Thus, TMS cannot be used to reliably target deeper, or subcortical structures (Walsh & Cowey, 2000).

Furthermore, the spatial resolution of TMS is dependent on the size of the magnetic field that is generated and the coil used. Generally, when using an intensity that is strong enough to elicit a significant 'virtual lesion' without overstimulating the region, the spatial resolution of TMS is powerful enough to be able to differentially affect areas that are only 1 to 0.5 cm apart (Brasil-Neto et al., 1992), and is particularly precise with figure-of-eight-coils (Hallet, 2007; Walsh & Rushworth, 1999), but the effects can spread very rapidly across neighboring cortical areas, as well as interconnected areas (Ilmoniemi et al., 1997), decreasing spatial resolution across time.

Overall, while the 'virtual lesion' caused by TMS is not as precise and controlled as the kind of selective lesions to entire specific cytoarchitectonic areas that are possible with neurosurgery, it nonetheless provides a higher degree of control than neuropsychological investigations over localization, timing of the 'virtual lesion', control measures and individual variability. This is particularly useful to further investigate the results of our lesion study, where we observed that abolition of the adaptation effect was seemingly due to LH trials being faster in the patient population compared to their controls. As there was a numerical tendency for HH trials to also be faster in the patient population than their controls, and in the absence of pre-op and post-op data, it is unclear whether the significantly faster LH trials are an effect of the lesion to dlPFC or of other confounds (e.g. individual differences) making the dlPFC group faster overall. Therefore, by applying TMS to the dlPFC, we can expand on our patient lesion study by

measuring within-subject response times and assess what specific changes in response speed are due to the virtual lesion.

Several studies have used TMS to investigate the mechanisms underlying performance in conflict tasks. However, while areas such as the medial PFC (mPFC) - including Pre-Supplementary and Supplementary Motor Areas and the ACC – (Hayward, Goodwin, & Harmer, 2004; Jin, Olk, & Hilgetag, 2010; Mars et al., 2009; Soutschek et al., 2013; Taylor, Nobre, & Rushworth, 2007), Pre-Motor and Motor cortices (Mars et al., 2009; Praamstra et al., 1999; Stürmer et al., 2000) and Posterior Parietal Cortex (Jin et al., 2010; Stürmer et al., 2007) have been commonly targeted, very few studies have looked at the effects of TMS on the dlPFC during conflict tasks, especially with regards to adaptation.

While TMS to mPFC has been found to generally interfere with optimal conflict processing, for example by increasing RTs to response-based conflict trials (Soutscheck et al., 2013) and errors on high-conflict trials (Taylor et al., 2007) during Flanker tasks, some also reported the abolishment of the incongruency effect during the Stroop task (Hayward et al., 2004). Given the prominent ACC role in conflict-monitoring suggested by the conflict-monitoring account, it is particularly interesting to investigate the effects of TMS on the ACC, over other mPFC structures. However, one issue with the application of TMS to the mPFC is that it is impossible to target the ACC proper in isolation. ACC is a deeper cortical structure, whose stimulation can be achieved only by

also stimulating more superficial mPFC areas such as SMA and pre-SMA. Therefore, one cannot conclude that the effects of stimulation to the ACC during conflict tasks are specific to interference with activity within this region alone (as in the case of Hayward and colleagues; 2004).

dlPFC, on the other hand, is a superficial cortical region that is more suitable for TMS investigations. TMS to the dlPFC has been found to have no effect on conflict-related performance on the current trial during a Stroop task (Vanderhasselt et al., 2007; Vanderhasselt, et al., 2006). Similarly, Wagner and colleagues (2006) found no effect of left dlPFC TMS on either the incongruent condition of the Stroop task, nor the standard WCST. Ko and colleagues (2008) found that TMS of the right dlPFC impaired performance on the WCST only when applied during the feedback period, but not if applied during matching itself.

However, involvement of dlPFC in adaptation is strongly implicated by imaging studies (Durston, 2003; Egner & Hirsch, 2005; Kerns, 2006; Kerns et al., 2004; Kim, Chung, & Kim, 2012) and the necessity of dlPFC has been confirmed by non-human primate lesion studies (Mansouri et al. 2007). Despite good evidence of dlPFC involvement, the form this involvement takes is still unclear, and several accounts have been proposed: maintenance of conflict history or attentional demands in working memory during the inter-trial period (MacDonald, 2000; Mansouri et al., 2007a), reactivation of recent mental representations or rule retrieval (Badre, 2008; Raye,

Johnson, Mitchell, Reeder, & Greene, 2002; Roth, Johnson, Raye, & Constable, 2009), engagement of attentional resources and cognitive control (Botvinick, et al., 2001; Kerns, 2006), or enhancement of task relevant information (Depisapia & Braver, 2006; Egner & Hirsch, 2005).

To the best of my knowledge, Sturmer and colleagues (2007) carried out the only study that looked at effects of TMS of the dlPFC on adaptation. They found that the reduction in the Simon effect (measured as the difference in speed between low- and high-conflict trials), usually observed after high-conflict trials was abolished by 20Hz repetitive TMS (rTMS), when this was applied to the left dlPFC. The effect was restricted to a time-window 500-300ms before stimulus presentation, and not to earlier time-windows, indicating that dlPFC is unlikely to be involved in maintenance of information in working-memory across trials. However, Sturmer and colleagues did not specify whether the adaptation effect is abolished through the effects of TMS on specific trial sequences (e.g. through increasing speed LH trials, or decreasing speed on HH trials). One limitation of their study also concerns the timing of the stimulation. They noted that, whilst a conflict effect was present after both high- and low-conflict trials, the effect was nonetheless smaller after high-conflict trials than after low-conflict trials, indicating some degree of adaptation. They suggested that this might be due to dlPFC processing being largest in the time directly preceding the new stimulus set, and that stimulation at 500-300ms before stimulus presentation might be too early to interfere

with dlPFC activity during this critical time-window, and cause complete abolition of the adaptation effect.

In the current experiment, we followed up on our lesion study by investigating whether TMS to the dlPFC affects adaptation in our WCST analogue and, if so, whether the adaptation effect is abolished via increase in speed on LH trials or decrease in speed in HH trials (or both). Using a paradigm similar to Sturmer and colleagues' (2007), we applied on-line repetitive rTMS to the left dlPFC during selected trials, at a 10Hz frequency. Given the evidence that increases in rTMS frequency leads to longer-lasting effects, as compared to single-pulse TMS (Hallet, 2007), we used lower frequency rTMS than Sturmer in order to limit the effects of TMS to a shorter, and more precise, time-period, while still exploiting the stronger effects of rTMS. As Sturmer's study indicated that rTMS affects performance when administered in time periods closer to stimulus onset, rather than earlier during the inter-trial period (thereby implicating an effect on the retrieval or re-activation of task-relevant information as opposed to maintenance of information across the inter-trial period), and that such effect could be larger when stimulation is administered closer to the start of the new trial, we applied TMS at the precise moment of presentation of the test items, and therefore at the moment conflict was triggered, requiring rule-retrieval.

Since the equivalent of low-conflict trials in the WCST analogue have been shown not to be affected by TMS (Ko et al., 2008), and in order to maximize the number

of pairs of trials that could be reliably considered free of TMS effects carrying over from previous trials, only high-conflict trials were stimulated. We targeted left dlPFC as this is an area that has been found to be to be activated during adaptation periods in conflict tasks (e.g. Kim et al., 2012; Kerns et al., 2006; MacDonald, 2000) and, unlike right dlPFC, has been shown to negatively affect adaptation when stimulated (Sturmer et al., 2007). Furthermore, 2 out of 3 of our patients presented unilateral lesions to left dlPFC. We hypothesized that TMS to the left dlPFC would abolish the adaptation effect normally observed on non-TMS trials, thus replicating the findings from our lesion study. In order to investigate whether there might be any subdivisional specificity within dlPFC in aiding adaptation, we targeted two different sites in this region: a more lateral part of dlPFC, which approximately corresponded to areas 46 or 9/46 (which was the same area targeted in Sturmer's study) as well as an area corresponding approximately to midline dorsal area 9.

The different accounts of dlPFC involvement mentioned earlier also allow for different predictions with regards to which type of trial would be affected by TMS. We predict that, if the dlPFC is involved in maintaining information about recent conflict-history in working-memory, LH and HH trials should both be affected by TMS, as response speed in both should be dependent on conflict-history. On the other hand, if dlPFC is involved in refreshing tasks instructions, we predict that only LH trials should be affected, becoming even slower, as these are the trials signaling the need to retrieve task instructions. However, we would predict no changes in the adaptation effect, as

performance on HH trials can still be enhanced once the task instructions have been refreshed in working memory. If the role of the dlPFC is to engage cognitive control in order to reduce the detrimental effects of conflict on subsequent trials, we would expect HH trials, the trials that should most benefit from this type of proactive engagement of cognitive control, to be affected by stimulation, with response speed dropping to LH levels, therefore canceling out the adaptation effect.

Methods

Participants

33 participants (15 male, mean age 24.18 years) took part in the experiment. Participants were students and staff at the University of Oxford, recruited through advertisement, and received monetary compensation for their participation in the study. All participants had no history of current or previous neurological or psychiatric condition and were not taking any psychoactive medication, as established by a screening questionnaire. All participants were fluent English speakers and had normal or corrected-to-normal vision. They all provided written consent prior to their participation in the study. Participants were randomly assigned to one of the two experimental groups, one group receiving stimulation to left area 9/46 (12 participants) and one group receiving stimulation to midline dorsal area 9 (10 participants), or to a control group (11 participants), who received stimulation at the vertex (see below for details).

Repetitive Transcranial Magnetic Stimulation

Repetitive Transcranial Magnetic Stimulation was carried out using a biphasic Super Rapid Magstim stimulator (Magstim, Dyfed, UK) and a 70mm figure-of-eight coil. First, the Resting Motor Threshold (RMT) was measured individually in each participant. In order to measure the RMT, stimulation was applied to the left primary motor cortex (M1), targeting the site that elicited the largest twitch in the index finger of the participant at the lowest stimulator output (site search started from a spot localized 5cm laterally and 1cm rostral of the vertex). The RMT was defined as the intensity needed to produce Motor Evoked Potentials (MEPs) of at least 50 μV in the First Dorsal Interosseus muscle (FDI) of the right hand in at least 5 out of 10 trials. MEPs were recorded using Ag–AgCl electrodes in a belly-tendon montage. The signal was acquired through a CED 1902 amplifier, a CED 1401 analog-to-digital converter, and the Signal software (Cambridge Electronic Design, Cambridge, UK).

The stimulation sites selected for the experimental groups were left dlPFC (corresponding approximately to BA 9/46, from here on referred to as BA9/46) and midline dlPFC (approximately dorsal BA 9, from here on referred to as BA9d). The site of stimulation in the BA9/46 group was localized using the Beam F3 Location System (Beam, Borckardt, Reeves, & George, 2009). This system allows the measurement of the location of the F3 electrode position in the 10-20 EEG coordinate system (which has been shown to be a reliable reference point for the localization of dlPFC in the absence

of structural brain scans - see, for example Stürmer et al., 2007), by taking into account individual variability in skull sizes. BA9/46 was localized 1cm caudally of F3, as according to Stürmer and colleagues (2007). The BA9d site was localized 8cm rostral to the vertex and 7cm superior to F3, along the midline; in order to ensure stimulation at a dlPFC site that would be as reliably as possible out of lateral BA9/46. We chose to target this more superior location within dlPFC as this is an area whose analogous region in the macaque monkey has been shown not to affect performance on the WCST analogue, when lesioned (Buckley et al., 2009), unlike BA 46. The control site, the vertex, was localized at a site corresponding to the electrode CZ location in the 10-20 EEG coordinate system, measured as half the distance between inion and nasion and intersecting with half the distance between the two aural canals. Stimulation sites are illustrated in Figure 12.

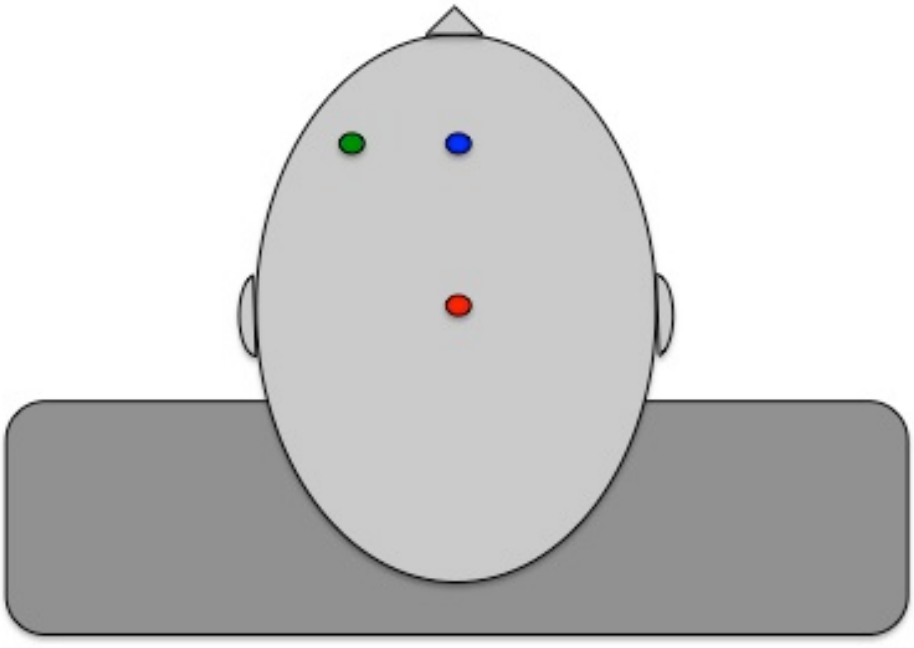


Figure 12. *Stimulation sites used in the experiment. BA9/46 (green), BA9d (blue) and vertex (red).*

Stimulation was performed at 90% of the RMT stimulator output. As a 2.8% reduction in stimulator output for every mm closer to the skull has been recommended (Stokes et al., 2005; 2007; 2013), this intensity was within an appropriate range considering the average difference in scalp-cortical surface distance between M1 and the stimulation sites (i.e. approximately 3-4mm for the Inferior and Middle Frontal Gyri). The stimulation coil was held at a 45° angle off the midline and rTMS pulse-trains consisted of 3 pulses administered 100ms apart (10Hz), on selected high conflict trials.

Apparatus, stimuli and procedure

The task used in the study was a computerized analogue of the Wisconsin Card Sorting Test, which was originally developed by Mansouri and colleagues (2007). The task set-up and procedure was identical to the one used in the neuropsychological study illustrated in the previous chapter, with some slight modifications to adapt the procedure to the requirements of TMS (see below). The task was programmed using Turbo Pascal (Borland), run in DOS on a desktop PC and presented on a 20.1" colour touchscreen (TFT LCD TS200H GNR), which was used to collect responses. Participant sat at a distance of 40cm from the screen and they were instructed to respond using the index finger of their dominant hand to touch the items on the screen.

Participants carried out a total of 4 blocks of 150 trials each. rTMS pulse-trains were applied to half the total number of H trials. On TMS trials, pulse-trains were triggered at the time the sample item was touched by the participant, initiating the trial.

Testing sessions were pre-programmed so that participants would not be able to anticipate stimulation, and in such a way that an equal number of LH and HH trials were available for stimulation in each session. As rTMS effects can outlast the period of stimulation itself (Hallett, 2007) and we were interested in the effect of stimulation on adaptation, and therefore the interaction between the nature of the previous trial (L or H) and stimulation on the current trial, we ensured that possible additive effects of a TMS trial on subsequent trial were controlled for. In order to allow for the effects of stimulation to dissipate, pulse-trains were never administered less than 3 trials apart. Trials were considered "clear" trials only if the previous two trials were both non-TMS trials. Similarly, TMS was administered only on trials that were preceded by two clear trials. A block contained, on average, 25 TMS trials and 25 clear trials. Reaction times and errors were recorded for analysis.

Results

As in the previous study, the speed of target selection (STS) was computed by taking the reciprocal of the response time (RT) data. STS analysis was carried out on correct trials only. The overall number of errors was too low to carry out any meaningful analysis on error data (average number of errors that were non-switch trials per condition was 1.58), therefore, analysis was carried out on STS data only. A small number of participants showed a negative adaptation effect on non-TMS trials. As we

were interested in investigating the effects of TMS on a standard (i.e. positive) adaptation effect, these participants were excluded from the analysis. In total, 9 participants were included in the BA9/46 group, 7 participants were included BA9d group and 10 participants were included in the control group.

A 2x2x3 mixed ANOVA was carried out on the STS data, with Adaptation (LH or HH) and Stimulation (TMS or non-TMS) as the within-subject factors and Group (BA9/46, BA9d and Vertex) as the between-subject factor. The ANOVA indicated a main effect of adaptation ($F_{(1,22)}$= 24.22, $p < .001$), with lower STS values for LH trials (mean = 1.31, sd = 0.18) than HH trials (mean = 1.35, sd = 0.18). This confirms the presence of an adaptation effect, with faster responses to HH trials than to LH trials. There ws a significant main effect of Stimulation ($F_{(1,22)}$= 6.46, p = .019)., with higher STS values for TMS trials (mean = 1.35, sd = 0.19) than non-TMS trials (mean = 1.31, sd = 0.18). There was no significant main effect of Group ($F_{(1,22)}$ = 1.52, p =.240) and no significant interaction between Group and Stimulation ($F_{(1,22)}$ = 1.39, p =.270). There was a significant interaction between Adaptation and Stimulation ($F_{(1,22)}$= 21.85, $p < .001$), suggesting that TMS modulates speed of response differently for different types of adaptation trials (HH vs LH). There was also a significant three-way interaction between Adaptation, Stimulation and Group ($F_{(1,22)}$ = 3.53, p =.047), indicating that the modulation of adaptation by TMS differs between the three groups. In order to investigate the significant three-way interaction, we carried out three separate 2x2 repeated-measures ANOVA for the three groups.

In the BA9/46 group, there was a significant main effect of Adaptation ($F_{(1,8)}$ = 7.84, p =.023), with lower STS values for LH trials (mean = 1.33, sd = 0.19) than HH trials (mean = 1.37, sd = 0.16), confirming the presence of an adaptation effect for this group. The main effect of TMS was not significant ($F_{(1,8)} < 1$). There was however a significant interaction between Adaptation and Stimulation ($F_{(1,8)}$ = 14.56, p =.005), indicating a significantly smaller adaptation effect on TMS trials (mean = -.01, sd = .05) than on non-TMS trials (mean = .10, sd = .07). We investigated this interaction further by carrying out paired-samples T-tests on the speed of response to HH and LH trials, depending on Stimulation, in order to determine whether TMS affected one type of trial more than the other. There was no significant difference between non-TMS LH and TMS LH trials ($t_{(8)}$ = -1.87, p > .05), and no significant difference between non-TMS HH and TMS HH trials ($t_{(8)}$ = 1.51, p > .05). TMS in this group, therefore, affects the overall adaptation effect, with no selective effect on either type of trial (see Figure 13).

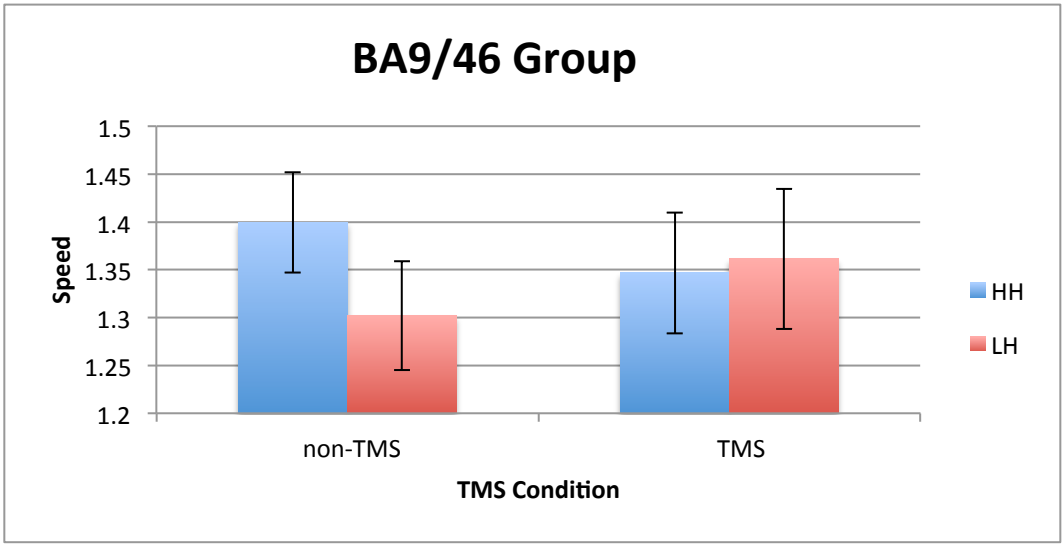


Figure 13. *Effect of TMS on speed of response to HH and LH trials in the BA9/46 group. Higher STS values for HH than LH trials indicate an adaptation effect. This is only present on non-TMS trials (leftmost columns), and is abolished by stimulation (rightmost columns).*

The BA9d group showed a strong trend towards a significant main effect of Adaptation ($F_{(1,6)} = 5.36$, $p = .060$), with higher STS values for HH trials (mean = 1.41, sd = .16) than LH trials (mean = 1.38, sd = .17). There was also a significant main effect of Stimulation ($F_{(1,6)} = 13.23$, $p = .011$), with higher STS values on TMS trials (mean = 1.42, sd = .16) than non-TMS trials (mean = 1.37, sd = .17), indicating that stimulation sped up responses overall. The interaction between Adaptation and Stimulation was also significant ($F_{(1,6)} = 8.47$, $p = .027$), indicating a modulatory effect of TMS on adaptation effects, with a larger adaptation effect on non-TMS trials (mean = .07, sd = .02) than on TMS trials (mean = -.001, sd = .07). Paired-samples t-tests revealed a significant difference in speed ($t_{(6)} = -6.09$, $p = .001$), of response between TMS LH trials (mean = 1.42, sd = .18) and non-TMS LH trials (mean = 1.33, sd = .17), but not between TMS and non-TMS HH trials ($t_{(6)} < 1$). This indicates that the adaptation effect in this group is cancelled out by an increase in speed of response on LH trials following stimulation (see Figure 14).

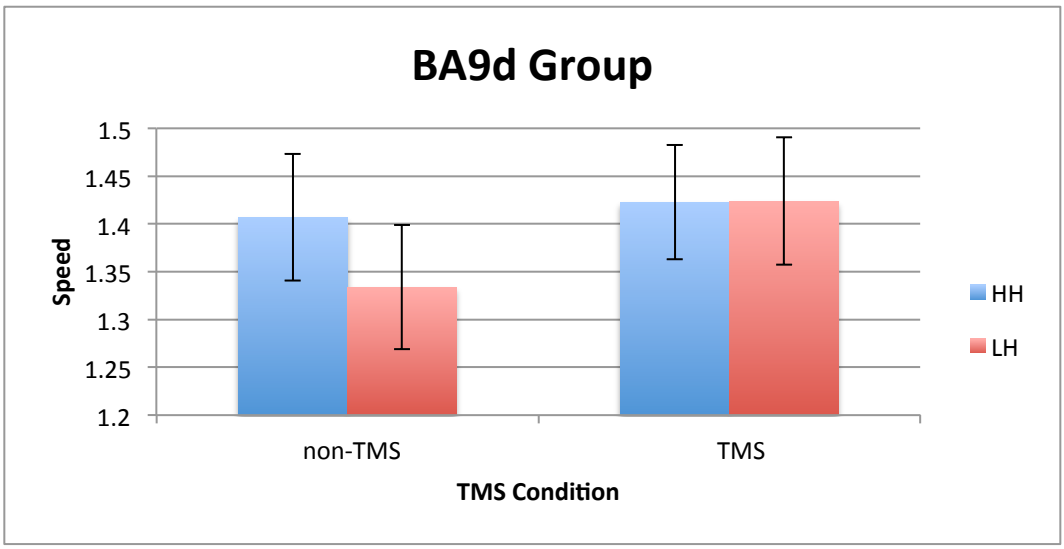


Figure 14. *Effect of TMS on speed of response to HH and LH trials in the BA9d group. Higher STS values for HH than LH trials indicate an adaptation effect. This is only present on non-TMS trials (leftmost columns), and is abolished by stimulation (rightmost columns).*

Lastly, the control group (Figure 15) showed a significant main effect of Adaptation ($F_{(1,8)}$ = 13.43, p =.006), with higher STS values for HH (mean = 1.26, sd = .18) than LH trials (mean = 1.22 , sd = .20), indicating a typical adaptation effect. There was also a significant main effect of Stimulation ($F_{(1,8)}$ = 6.23, p =.037), with higher STS values for TMS (mean = 1.27, sd = .20) than non-TMS trials (mean = 1.22, sd = .18), indicating that stimulation sped up responses overall. Unlike the other two groups, the control group showed no significant interaction between Stimulation and Adaptation ($F_{(1,8)}$ = 1.07, p =.331).

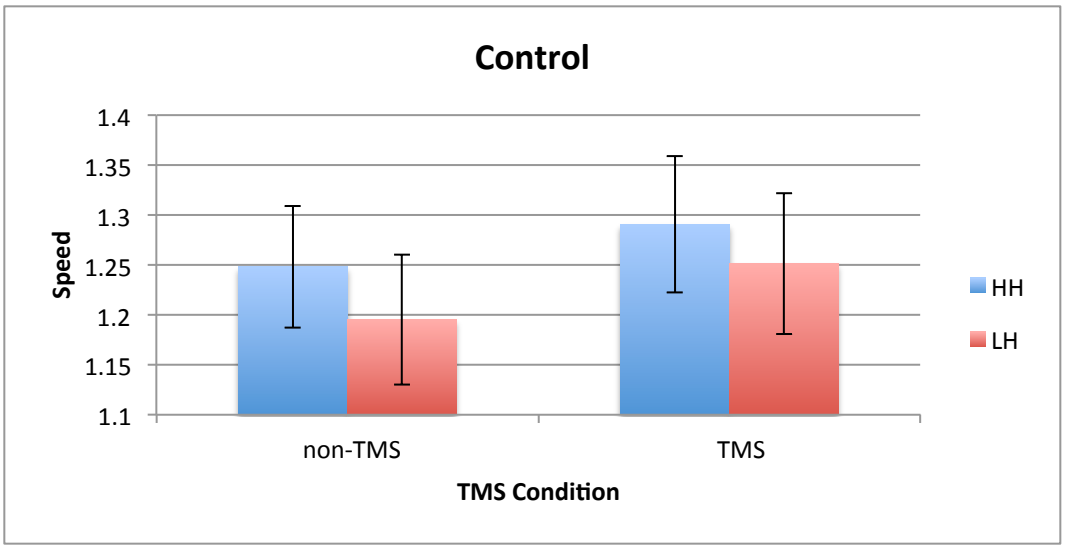


Figure 15. *Effect of TMS on speed of response to HH and LH trials in the Control group. Higher STS values for HH than LH for HH than LH trials indicate an adaptation effect. In this group, the adaptation effect is not affected by TMS.*

Discussion

In this study, we investigated the role of the dlPFC in conflict-induced behavioural adaptation using TMS. We were interested in whether applying TMS on high conflict trials of the WCST analogue would affect performance to the same extent as lesions and whether the effect would be restricted to specific pairs of trials (LH as opposed to HH) or be observable as a reduction of the overall adaptation effect only (i.e. the difference between LH and HH trials).

The results from our TMS study replicated the ones obtained in our lesion study and, in turn, the results from Mansouri and colleagues' primate study (2007). TMS of the dlPFC cancelled out the adaptation effect, reducing the response speed difference that

is normally observed between LH and HH trials. There appeared to be some differentiation in the effects of TMS to different sites within dlPFC. While TMS to the left BA9/46 group cancelled out the adaptation effect by modulating both types of trials (LH trials were numerically, but not significantly, faster, and HH trials were numerically, but not significantly, slower, when TMS was applied), TMS to BA9d appeared to cancel out the adaptation effect specifically through an increase in speed on LH trials. This increase of speed in LH trials was also observed in our patient data (where the absence of pre- and post-op data limited our interpretation) and could be due to dlPFC lesions in the patient cohort expanding beyond the more ventral part of dlFPC, and specifically area 9/46, into more dorsal dlPFC areas such as BA9d. This is inconsistent with the non-human primate lesion studies, which show that performance on the WCST analogue is not affected by lesions to areas of dlPFC from 3mm above the principal sulcus right up to the midline inclusive (Buckley et al., 2009). However, the BA9d group in our TMS study only showed a trend towards a significant adaptation effect and so at present we do not feel there is sufficient data to describe this as a robust species difference.

Our study is consistent with previous literature showing that conflict-induced behavioural adaptation is negatively affected by TMS to structures thought to be involved in conflict-related performance (e.g., Soutschek et al., 2013; 2010; Jin et al., 2010; Mars et al., 2009; Taylor, Nobre, & Rushworth, 2007; Hayward, Goodwin, & Harmer, 2004; Stürmer, et al., 2000; Praamstra, Kleine, & Schnitzler, 1999) and with TMS and imaging literature showing that dlPFC in particular is involved in adaptation

(Durston, 2003; Egner & Hirsch, 2005; Kerns, 2006; Kerns et al., 2004; Kim, Chung, & Kim, 2012). Although studies such as Wagner and colleagues' (2006) and Ko and colleagues' (2008) found no effect of left dlPFC stimulation on the WCST, our results are not necessarily inconsistent with these findings, as we specifically looked at whether TMS affected the way the nature of previous trials modulated performance on the current trial. We did not find any effect of TMS on the overall performance in the experimental groups, consistently with these studies.

Stimulation of the dlPFC at different time points should have a different effect on performance depending on the nature of dlPFC involvement. If dlPFC were crucial to maintain information online in between trials, we would expect stimulation during the inter-trial period to affect performance. If dlPFC were crucial to re-activate recent mental representations (such as rules, stimulus-response associations, or memory traces), we would expect performance to be affected by stimulation during the stimulus-presentation period, but not during the inter-trial period. If dlPFC were crucial to enhance task-relevant information or paths, we would expect performance to be affected by stimulation at the time of, and shortly after, stimulus-presentation. As previously mentioned, Sturmer and colleagues (2007) found that stimulation affected adaptation when this was administered close to stimulus presentation, but not at earlier stages during the inter-trial interval, suggesting that dlPFC does not modulate adaptation by maintaining information online across trials.

The timing of stimulation used in our experiment cannot inform on the specific type of involvement of dlPFC in adaptation during the WCST, as stimulating at the time of stimuli-onset could be interfering with maintenance of information from a previous trial, as well as with retrieval or reactivation of information at the start of the trial, or with decision-making processes pertaining to the post-stimulus period (as 10Hz stimulation is likely to affect more than just the stimulation period). Due to time constraints, we were not able to test the effects of stimulation on an earlier time-period, in the way Sturmer and colleagues (2007) did, in order to investigate whether the effect is restricted to the stimulus-presentation. This is something we would be interested in exploring in further research. However, we can attempt to infer about the specific role of dlPFC in adaptation by looking at the different effects of TMS on specific pairs of trials.

We saw that the adaptation effect was abolished by TMS, and that this was at least partly due to an increase in speed on LH trials, as well as a numerical decrease in speed of response on HH trials. The absence of an adaptation effect, and a slower speed of response on HH trials, is consistent with the hypothesis that dlPFC might be involved in engaging cognitive resources in order to minimize the effects of conflict on subsequent trials, after this has recently been experienced, either through the enhancement of task-relevant information, sustained attention or rule-maintenance (Botvinick, Braver, & Barch, 2001; De Pisapia, Sandrini, Braver, & Cattaneo, 2012; Egner & Hirsch, 2005; Kerns, 2006). However, faster responses on LH trials are not consistent

with a role of dlPFC in implementing cognitive control. Being the first trial after a change in the level of conflict, from a trial where retrieval or maintenance of the rule, while helpful, is not necessary to perform a correct response (i.e. two identical items are present on the screen) to a trial where retrieval of the current relevant rule is necessary, LH trials represent a context where cognitive control is a necessary requirement for accurate and fast responses. Similarly, the improvement in performance on LH trials is also inconsistent with a role of dlPFC in the reactivation or retrieval of recent representations (Badre & Wagner, 2004; Raye et al., 2002), as interference with these mechanisms should result in slower responses.

However, abolition of the adaptation effect that is dependent both on modulation of HH and on modulation of LH trials may still be consistent with a role of dlPFC in maintaining conflict-related information. Maintenance of conflict information might not refer specifically to the maintenance of strictly high-conflict information. Cells in dlPFC have been found to respond to both low- and high-conflict trials of the WCST analogue, as well as being sensitive to both LH and HH trials during the inter-trial period (Mansouri et al., 2007). Furthermore, the dlPFC has been found to be more active on high-conflict trials as the number of preceding low-conflict trials increase and to then decrease as the number of consecutive high-conflict trials increase (Durston, 2003), indicating that it does not simply encode the presence of sustained high conflict. dlPFC may therefore be maintaining recent conflict-history or contextual information in working memory, and be sensitive to changes in conflict-levels which require some type

of adjustment on the part of the cognitive system - in the form of enhancement of task-relevant information, increased cognitive-control, strategy shifts, etc.

Regardless, if the role of dlPFC were specifically about maintaining conflict history online and reacting to changes in conflict levels, it is unclear why damage or TMS to dlPFC would lead to faster responses on LH trials. We can attempt to reconcile these findings with the maintenance view by considering them under the dual-route to response framework (Kornblum et al., 1990; Stoffels, 1996; Sturmer et al., 2002; Wühr & Ansorge, 2005). These authors suggest that behaviour can be supported by two routes to response: a direct route, supporting habitual, reflexive and overpracticed responses, and an indirect route, supporting instructed, rule-guided behavior, either through the creation of links between arbitrary stimuli and responses or the application of more abstract rules. These two routes can be modulated, by suppression or activation, depending on the demands of the current trial. For example, in the Simon task, responding to the location of the stimulus is a reflexive type of behaviour encoded in the direct route that aids performance on low-conflict trials, but that needs to be suppressed on high-conflict trials in favor of an arbitrary stimulus-response association that is encoded in the indirect route.

A question remains as to whether the cognitive system has a “default” mode of suppression or activation of one route or the other, and evidence seems to suggest that rather than a hardwired “default” route, both routes are likely to be concurrently

modulated depending on task demands (Sturmer et al., 2002). The default route might therefore be task-dependent, i.e. be the route that is more beneficial to performance for the overall task, and therefore more often exploited. As low-conflict trials on the WCST cannot benefit from exploiting habitual responses, the default route might be the indirect route supporting rule-guided behaviour. If the so-called "default" route to response for the WCST is the indirect route, and the role of dlPFC is to maintain a memory trace for recent conflict-history, when this memory trace is lost following TMS, the cognitive system might reset to the default, in this case indirect, route to response and this may lead to more efficient responses on LH trials. This issue will be further explored in Chapters 3 and 4.

Overall, while the findings from our study are in agreement with findings showing that TMS to the dlPFC appears to cancel adaptation effects, we find that the abolition of adaptation effects could be due, more specifically, to a combination of slower responses on some types of trials (such as HH trials) and faster responses on others (such as LH trials). Many of the studies that argue for impairments in adaptation, following lesions or TMS, measure adaptation as the difference in the size of the conflict effect (i.e. the difference between response times to low- and high-conflict trials), depending on whether the preceding trial was high- or low-conflict (di Pellegrino et al., 2007; Praamstra et al., 1999; Sheth et al., 2012; Stürmer et al., 2007; Diane Swick & Turken, 2002). A significant adaptation effect therefore corresponds to a significantly smaller, or null, conflict effect after high-conflict trials, compared to the conflict effect observed

after low-conflict trials. The presence of a conflict effect after high-conflict trials that is as large as the one observed after low-conflict trials is seen as an index of impairment in adaption. However, in our WCST studies, we measure adaptation as the difference in speed between high-conflict trials only, as a function of preceding trial type (i.e. low-conflict or high-conflict).

Studies measuring adaptation as a reduction in conflict effects following high-conflict trials are effectively looking at a difference between two differences (i.e. the difference between LL and LH compared to the difference between HL and HH) and often do not specify whether this reduction is due to an increase in speed for HH trials or a decrease in speed for HL trials (or both), making it impossible to determine whether there are specific aspects of adaptation that are particularly affected by lesion or TMS (such as in the case of Sturmer et al., 2007) and whether this might be task dependent. In the following chapters, we will look more closely at direct comparisons between the WCST and versions of the most commonly used conflict tasks (i.e. Simon and Flanker tasks), to investigate whether adaptation effects may be driven by different mechanisms depending on the task, and whether this could explain different effects of TMS or lesions on adaptation effects.

To summarize, our TMS study replicated the findings of our lesion study, by showing that interfering with dlPFC activity abolishes the adaptation effect. The abolition of the adaptation effect seems to be due to different modulation of the speed

of response on HH and LH trials. Given the finding of faster responses on LH trials after TMS, the role of dlPFC in enhancing attentional and cognitive processing, or in reactivating recent representations or rules, in order to drive adaptation, is questioned. These effects might be better explained in terms of interference with a role for dlPFC in maintaining recent conflict-history or contextual information in working memory, which could be used to modulate two context-dependent routes to behaviour, in response to changes in conflict-levels.

CHAPTER 3

A comparative investigation into the possible underlying dynamics of conflict-induced adaptation effects in the conflict WCST analogue and classic conflict tasks

Introduction

Conflict effects are generally defined as the difference in speed of response between high- and low-conflict trials, specifically due to faster speed of response on the latter compared to the former. The larger the difference in speed, the larger the magnitude of conflict effects. Reductions in the magnitude of these conflict effects following high-conflict trials are a fairly established phenomenon in cognitive psychology (e.g. Akçay & Hazeltine, 2008; Botvinick, Braver, & Barch, 2001; Burle, Allain, Vidal, & Hasbroucq, 2005; Chen & Melara, 2009; Egner & Hirsch, 2005; Gratton, Coles, & Donchin, 1992; Hommel, Proctor, & Vu, 2004; Kerns et al., 2004; 2006; Kunde, 2003; Mansouri, Tanaka, & Buckley, 2009; Nieuwenhuis & Stins, 2006; Notebaert & Verguts, 2007; Purmann, Badde, & Wendt, 2009; Stürmer & Leuthold, 2003; Stürmer, Siggelkow, Dengler, & Leuthold, 2000; Ullsperger, Bylsma, & Botvinick, 2005; Wendt, Kluwe, & Peters, 2006; Wühr & Ansorge, 2005) and are generally referred to as sequential effects or adaptation effects. Adaptation effects are a key topic of investigation for conflict-monitoring and cognitive control research, as they are thought to represent a direct measure of the engagement of cognitive control in response to conflict, and of how

cognitive control can be implemented to minimize the effects of conflict on behaviour. Despite the substantial amount of research devoted to investigating these effects, the precise mechanisms that are responsible for adaptation are still unclear, as will be discussed below.

Adaptation effects were first observed by Gratton and colleagues (Gratton et al., 1992) in the Eriksen Flanker task (Eriksen & Schultz, 1979), and were subsequently replicated across a variety of conflict tasks such as the Simon task (Burle, Possamaï, Vidal, Bonnet, & Hasbroucq, 2002; Chen & Melara, 2009; Hommel et al., 2004; Kerns et al., 2004; Stürmer & Leuthold, 2003; Stürmer et al., 2000; Wendt et al., 2006; Wühr, 2005), other variants of the Flanker task (Durston, 2003; Fernandez-Duque & Knight, 2008; Forster, Carter, Cohen, & Cho, 2011; Nieuwenhuis & Stins, 2006; Soutschek, Taylor, Müller, & Schubert, 2013; Ullsperger et al., 2005; Verbruggen, Notebaert, Liefooghe, & Vandierendonck, 2006), the Stroop task (Egner & Hirsch, 2005; Egner & Hirsch, 2005; Kerns et al., 2004; Kim, Chung, & Kim, 2012; Larson, Kaufman, & Perlstein, 2009) and hybrid tasks incorporating elements from each of these tasks (Sheth et al., 2012; Verbruggen et al., 2006).

Several theoretical accounts have been advanced to explain the phenomenon of adaptation effects. While no consensus has been reached about which particular account provides the most exhaustive explanation of the range of results obtained across the literature, each framework has been thoroughly scrutinized and weaknesses,

as well as merits, of each have been identified. Broadly speaking, we can identify four primary potential accounts: priming effects (Mayr et al., 2003), feature-integration mechanisms (Chen & Melara, 2009; Hommel et al., 2004), modulation of different routes to response (Stoffels, 1996; Sturmer et al., 2002) and conflict-induced enhancement of attention/cognitive control (Botvinick et al., 2001; Botvinick, 2007; Carter & van Veen, 2007). While each of these theories is able to successfully account for parts of the findings, no single account has been so far able to explain all aspects of adaptation.

As seen in previous chapters, the conflict-monitoring account posits that high-conflict trials trigger an increase in cognitive control, as they require the allocation of more attentional resources to task-relevant information and the suppression of competing responses and influence of distractors. Once these resources have been engaged, the costs on speed of response and performance decrease, resulting in reductions in the magnitude of conflict effects (Botvinick et al., 2001; Botvinick, 2007; Carter & Veen, 2007; Yeung, Botvinick, & Cohen, 2004). This is a broad, comprehensive framework that can be applied to a wide variety of tasks and that has found considerable support in the imaging literature (Barch, Braver, Sabb, & Noll, 2000; Carter et al., 2000; Hazeltine, 2003; Horga et al., 2011; Kerns, 2006; Kerns et al., 2004; Kim et al., 2012; Luks, Simpson, Dale, & Hough, 2007; van Veen & Carter, 2005). However, the CMCC model does not clearly specify what are the exact mechanisms cognitive control implements in order to reduce conflict effects. As described in the following section,

several more specific, as well as alternative, processes that may contribute to adaptation effects have been suggested in the literature.

Mayr and colleagues (2003) were the first to propose a role for repetition priming, instead of cognitive control, in adaptation. Repetition priming can be defined as an improvement in performance (either in response-time, accuracy, or both) when a participant is required to respond to two or more identical consecutive trials. Priming can also be negative, when a stimulus or feature that was irrelevant on a previous trial needs to be attended to on the current trial, resulting in alternation costs. Mayr and colleagues pointed out that, in the case of many conflict tasks, sequences of high-conflict trials that show adaptation effects are often complete stimulus-response repetitions. Similarly, others (Nieuwenhuis & Stins, 2006; Wendt, Heldmann, Münte, & Kluwe, 2007) noted that most, if not all, high-conflict trials following low-conflict trials are alternations that could incur negative priming costs. However, extensive testing has shown that adaptation effects are still observable when repetitions are removed or controlled for (Akçay & Hazeltine, 2008; Egner & Hirsch, 2005a; Fernandez-Duque & Knight, 2008; Kerns, 2006; Kerns et al., 2004; Nieuwenhuis & Stins, 2006; Notebaert & Verguts, 2007; Ullsperger et al., 2005), or when repetitions rarely, if ever, occur, such as in the case of the WCST analogue that we used in Chapters 1 and 2 (see also Mansouri et al., 2007). Futhermore, a simple solution to this problem is to use tasks with a wider stimulus-set, in order to limit the numbers of repetitions or costly alternations (Akçay & Hazeltine, 2008; Egner, 2007). Therefore, while priming can constitute a confound in

measurements of adaptation, other mechanisms are also likely to contribute behavioural adjustments.

The feature-integration account (Chen & Melara, 2009; Hommel et al., 2004) is closely linked to the priming account, but posits that repetition benefits or alternation costs are due to the formation of short-term memory links between stimuli and responses (Tagliabue et al., 2000). These links require effortful 'unbinding' on trials where the stimulus repeats but the response alternates or viceversa. This type of partial repetitions is very common on HL and LH trials in two-choice tasks, resulting in response time costs (i.e. slower responses). These slower partial repetitions cause HL trials to become as slow as HH trials, therefore abolishing the conflict effect. This framework finds some support in EEG studies showing that the N2 component, which is believed to reflect pre-response conflict on correct trials (Forster et al., 2011; Nieuwenhuis et al., 2003; Wendt et al., 2007) and to originate in the ACC (Nieuwenhuis et al., 2003; Van Veen & Carter, 2002; West, Jakubek, Wymbs, Perry, & Moore, 2005), is sensitive to partial mismatches in stimulus or response features (Wang et al., 2003). However, as in the case of the priming account, the feature-integration account cannot explain why adaptation effects are also present in tasks such as our WCST analogue, where no short-term memory link formation between a stimulus and a response is required in order to perform the correct response, or in tasks where the stimulus set is large enough to limit the number of partial repetitions/alternation to a minimum. Furthermore, some authors

have shown how even when feature integration is accounted for, conflict-induced adaptation still accounts for variance in the data (Notebaert & Verguts, 2007).

In order to explain what other mechanisms might be involved in adaptation, some authors (Kornblum et al., 1990; Stoffels, 1996; Sturmer et al., 2002) propose a parallel processing system comprising a dual-route to response, with a direct/unconditional route supporting habitual responses, and an indirect/conditional route, supporting instructed responses. The former supports performance in low conflict/congruent trials, whereas the latter supports performance in high conflict/incongruent trials. According to this view, adaptation effects can be explained in terms of suppression of the direct route during high-conflict trials and how this might affect both the current and the subsequent trial. On the first high-conflict trial in a sequence, after a low-conflict trial (LH), the direct route is open and needs to be suppressed, resulting in slower responses. On the subsequent trial, the direct route will still be suppressed, and therefore there should be little to no costs in response speed due to the need to suppress an inappropriate response tendency. This will ultimately result in a smaller conflict effect.

While this is an elegant framework and can find correspondence in underlying neural structures (for example in cortico-striatal pathways supporting actions versus habits and response suppression), it also has some limitations. First, it remains unclear whether the cognitive system has a 'default' mode, where the direct route is always suppressed, and only released when it is deemed advantageous, or the direct route is

always active and only suppressed when the context indicates it is not advantageous. One way to address this question is to investigate conflict effects following neutral trials, as these are trials that should not elicit conflict between the two routes. If a conflict effect is observed after a neutral trial, then the direct route is likely to be active by default: neutral trials do not require the suppression of the direct route, therefore normal conflict costs should occur on the first high-conflict trial only if the habitual response needs to be suppressed; i.e. if the direct route is active, as it is the case on LH trials. On the other hand, if there is no conflict effect following a neutral trial, the direct route is likely to be suppressed by default (and, consequently, the indirect route will be the default route to response), and conflict effects are therefore only observed after this route is released from suppression, i.e. on low-conflict trials. However, findings from studies that have used neutral trials are contradictory, as conflict effects are observed after neutral trials in some tasks but not in others (Chen & Melara, 2009; Wühr & Ansorge, 2005). Furthermore, conflict effects and adaptation effects are also observed in tasks such as our WCST, which does not rely on the direct route for any type of trial, indicating that adaptation effects are not entirely due to the modulation of the direct route, but might also be driven by biasing and enhancement of indirect route activity. This suggests that there could be different underlying dynamics that are responsible for the adaptation effects observed in many conflict tasks.

In summary, we can see how each account can explain some aspects of adaptation, but not others. Priming definitely plays a role in adaptation, but there are

also other mechanisms at play. The dual-route to response framework can explain adaptation effects in tasks involving competition between a habitual response and an instructed response, but cannot fully explain how they might manifest in tasks that do not rely on the direct route, such as the WCST, or tasks where competition is created between neutral trials, without relying on pre-existing habits (i.e. Chen & Melara, 2009). The feature-integration account can explain response-time modulations and conflict-frequency modulations, but similarly to priming, cannot explain adaptation effects in tasks that do not involve, or control for, unbinding processes.

The conflict-monitoring model posits that, when experiencing conflict on a high-conflict trial, the cognitive system signals the need to increase the allocation of attentional and cognitive resources to task-relevant information and instructions. This, in turn, drives adaptation on subsequent trials. As previously mentioned, most studies investigating adaptation effects measure adaptation as the difference in the size of the conflict effect following high-conflict trials and the size of the conflict effect following low-conflict trials. This can be summarized by the following equation:

$$(1) \qquad \text{Adaptation} = (HL\text{-}HH) - (LL\text{-}LH)$$

As we can infer from equation (1), adaptation effects, measured as the difference between factor 1 (HL – HH) and 2 (LL – LH), can therefore be due to either a decrease in speed for L trials following H trials (HL) or an increase in speed for high-

conflict trials following high-conflict trials (HH). It is therefore possible that the reduction in speed in HL trials may contribute more to adaptation than the increase in speed in HH trials. Viceversa, adaptation could be due entirely to increases in HH trials speed. The end result, in terms of reduction of conflict effects after high-conflict trials, is the same regardless of whether adjustments on HL or HH trials contribute more.

Of the large amount of literature exploring adaptation effects, some studies do report both increases in HH speed and decreases in HL speed (e.g. Egner, Delano, & Hirsch, 2007; Egner & Hirsch, 2005; Kerns, 2006; Ullsperger et al., 2005), some only find reductions in HL speed (e.g. Horga et al., 2011; Stoffels, 1996), some only find increases in HH speed (e.g. Kerns et al., 2004; Mansouri, Buckley, & Tanaka, 2007), but many only report reductions in conflict effects, without specifying what types of trials are affected (e.g. Chen & Melara, 2009; Kunde, 2003; Purmann et al., 2009; Sturmer et al., 2002; Stürmer, Redlich, Irlbacher, & Brandt, 2007; Verbruggen et al., 2006; Wühr & Ansorge, 2005). However, we argue it is important to report adaptation effects in terms of the specific trials that are being affected, rather than as a difference between two differences, as they might reveal different underlying adaptation mechanisms, as we will see further on in this chapter.

Given the very common findings of adaptation being due, either partially or totally, to slower responses on HL trials, the reductions in conflict effects reported in classic conflict tasks that create competition between a habit and an instructed

response could be partially, or mostly, due to costs on the facilitatory mechanisms that can speed up responses on low-conflict trials (such as the exploitation of a habit). These costs could even exceed the benefits of the proactive engagement of attentional resources on high-conflict trials. This is consistent with the notion of direct route (i.e. the route supporting habitual responses) suppression on a high-conflict trial carrying over on the subsequent trial, but, rather than reducing the costs of conflict on the subsequent high-conflict (HH) trial, it reduces the benefits of exploiting the direct route on a low-conflict (HL) trial.

Indeed, Kerns (2004) observed that, in the Stroop task, HL trials might be slower because focusing on the colour causes a reduction on the facilitative influence of the word that is normally observed on LL trials. This is in line with Jacoby and colleagues' Process Dissociation (PD) procedure (Jacoby et al., 2003), which stipulates that color-naming (i.e. the indirect route) and word-reading (i.e. the direct route) processes both contribute to performance. However, while the probability of naming the color correctly on congruent items is determined by the extent to which both color-naming and word-reading processes influence response, the probability of correctly naming the color on incongruent trials is determined only by the extent to which color-naming processes influence response. Lu & Proctor (1995) have similarly shown that the irrelevant location of an item in the Simon or spatial-Stroop task influences performance when there is strong prior association of the irrelevant dimension with the response and when the temporal overlap between the response primed by the irrelevant dimension and the

response primed by the relevant dimension is larger: the faster the RT in the congruent mapping, the greater the interference in the incongruent condition (Fitts & Deniger, 1954). Therefore, the stronger the influence of the direct route, the larger the facilitation on low-conflict trials, and the higher the cost on HL trials may be when the direct route is being suppressed.

Adaptation effects are therefore likely modulated by two mechanisms: suppression of habitual responses (when a context allows for the exploitation of the direct route that supports pre-existing habitual/reflexive behaviour on low-conflict trials) and enhanced rule-guided behaviour (possibly due to increased biasing of task-relevant information and the possibility of establishing new patterns for rule-guided behaviour in the indirect route). Adaptation might manifest itself in different forms depending on whether the highest level of performance can be achieved by relying on one route or the other.

We carried out several small behavioural experiments using different versions of some commonly used conflict tasks (Simon and Flanker) as well as a novel task (Arrows) and the WCST analogue, in order to compare the specific dynamics of adaptation effects between tasks that can exploit habitual responses to facilitate behaviour (i.e. the Simon tasks, the Arrows task and, to a lesser extent, the Flanker task, as detailed below), and a task that does not (i.e. the WCST). We predicted that all tasks would show a reduction in conflict effects following H trials. As per equation (1), reductions in the size of the

conflict effect should be observed regardless of whether the precise mechanism driving adaptation is reduction of speed on HL trials or increase of speed in HH trials. However, only the WCST should show adaptation restricted exclusively to faster HH trials. As this is a task where low-conflict trials cannot be facilitated by exploiting a habitual response, there should be no suppression of the direct route carrying over from previous high-conflict trials to slow down responses on HL trials. However, tasks like the commonly used Simon task, which can exploit the reflexive response of saccading or responding to the location of the stimulus to facilitate performance on low-conflict trials, might result in conflict reduction due to a reduction of facilitation on HL trials.

The Arrows task was an adaptation of the classic arrow-Flanker task, where participants had to respond by touching the side of the screen indicated by arrows presented on-screen. However, instead of all arrows being presented in a row, they were scattered across the screen, and participants had to respond to the direction the majority was pointing towards. This task, therefore, included facilitation on low-conflict trials, where all arrows pointed in the same direction and participants could exploit habitual behaviour of responding to the direction arrows point towards. High-conflict trials, on the other hand, were characterized by the necessity to scan the visual field for task-relevant information, rather than simply ignoring distractors, or reinforcing rule-guided behaviour. This would limit the usefulness of the indirect route (as no specific rule-guided behaviour can be implemented and biased as a shortcut to response), while, at the same time, require suppression of the direct route that can facilitate performance

on low-conflict trials. We therefore predicted that conflict-reduction would be entirely dependent on slower HL trials due to suppression of the direct route.

We also investigated how varying the frequency of high-conflict trials affected the WCST analogue. In most conflict tasks, increasing the proportion of high-conflict trials results in no conflict effect, (indicating a modulation similar to adaptation mechanisms, where no conflict effect is observed after high-conflict trials) or even a reversal of the conflict effect (Hommel, 1994; Logan, 1979; Nieuwenhuis et al., 2003; Sturmer et al., 2002). Investigating the effects of varying conflict-frequency in the WCST allowed us to elucidate whether the elimination or reversal of conflict effects by increased conflict frequency is due to more sustained engagement of the indirect route, when in the presence of sustained conflict (leading to faster responses on high-conflict trials), or, rather, it is mostly due to suppression of the direct route to response (leading to slower responses on low-conflict trials). If we observe no or reversed conflict effect in the WCST, we can conclude that enhanced biasing of the indirect route is sufficient to account for this type of adaptation, as this is not a task that is modulated by direct route suppression. If not, it would indicate that direct route suppression might play a bigger role in adaptation than enhanced rule-guided behaviour.

Methods

EXPERIMENT 1

All tasks were programmed using Turbo Pascal (Borland), run in DOS on a desktop PC and presented on a 20.1" colour touchscreen (TFT LCD TS200H GNR), which was used to collect responses in all tasks except for the task in Experiment 1c. Participants were students at the University of Oxford and took part in the study in exchange for course credits. All participants had normal or corrected-to-normal vision, were fluent English speakers and provided written consent prior to their participation in the study.

Experiment 1a – Simon task

Participants

20 participants (7 male, mean age 21.75) participated in this experiment.

Apparatus, stimuli and procedure

Participant sat at a distance of 40cm from the screen and they were instructed to respond using the index finger of their dominant hand to touch the items on the screen, and to be as fast and as accurate as possible.

The stimulus set consisted of coloured circles (4.3 degrees of visual angle in diameter) in two possible colours (blue or green), presented against a black background. White dots (0.35 degrees in diameter) indicated the touch-sensitive zones of the screen that could be selected for a response.

A typical trial was as follows. At the start of the trial, a white fixation cross (2 degrees in width and 2 degrees in height) was presented in the centre of the screen. Participants were instructed to touch the fixation cross when they were ready to start the trial. Once the fixation cross was touched, it disappeared, and a randomly selected coloured circle appeared on the screen, positioned either 20 degrees to the right of centre or to the left of centre. Depending on the colour of the circle, participants were instructed to touch either the right or the left side of the screen, marked by a white dot. For example, they could be instructed to touch the right side of the screen whenever a red circle was presented and the left side of the screen whenever a green circle was presented. The contingencies were randomly varied across participants. The stimulus remained on the screen until a response was made or until 10 seconds had elapsed. Correct trials were identified by a high-pitch sound, and the correct choice remaining on the screen for one second, indicating positive feedback. Incorrect trials were identified by a low-pitched sound accompanied by the presentation of a large, gray circle, indicating negative feedback. After 2 seconds from response, the fixation cross was presented again on the screen, indicating the start of another trial.

Conflict levels were manipulated by changing the degree of overlap between stimulus location and response location. In low-conflict trials, the stimulus was presented in the same (congruent) location as the location the participants were instructed to select given the colour of the stimulus. In high-conflict trials, the stimulus was presented in the opposite (incongruent) location. High- and low-conflict trials were presented in randomized order. Example of a trial is presented in Figure 16.

Participants carried out 2 blocks of 150 trials each. Reaction times and errors were recorded for analysis.

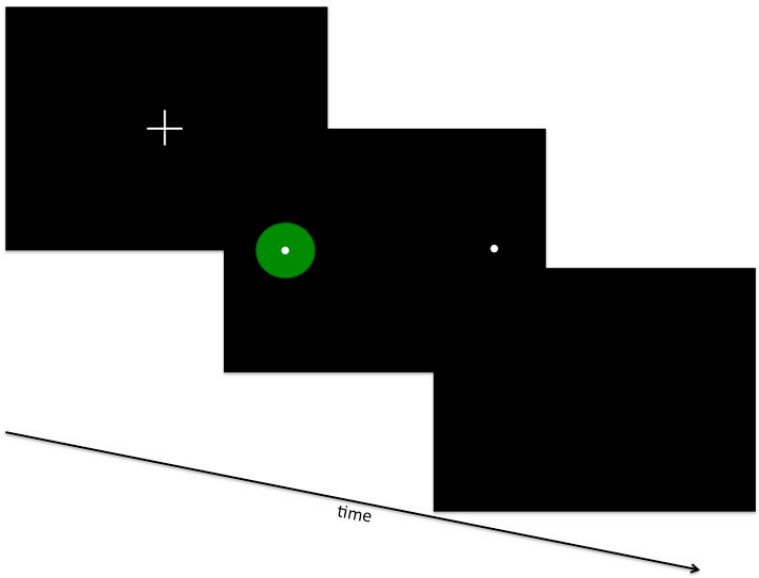


Figure 16. *Example of a typical trial in the Simon task. The small white dots indicate touch-sensitive regions on the screen that can be selected for response.*

Experiment 1b – Simon task with 3 cues

Participants

20 participants (13 male, mean age 22.15) participated in this experiment.

Apparatus, stimuli and procedure

The task used in this experiment was essentially identical to the Simon task described above. The only difference involved the addition of one extra test item. The third item added to the task consisted in a red circle of the same size and shape previously used for the other two items. During low conflict trials, the item was presented at a centered location 14 degrees of visual angle in distance from the to the top of the centre of the screen. During high conflict trials, it was presented in one of the other two incongruent locations. The other two test items were also presented at the new top-centre location on 50% of their respective incongruent trials. An example of a trial is presented in Figure 17.

Participants carried out 2 blocks of 150 trials each. Reaction times and errors were recorded for analysis.

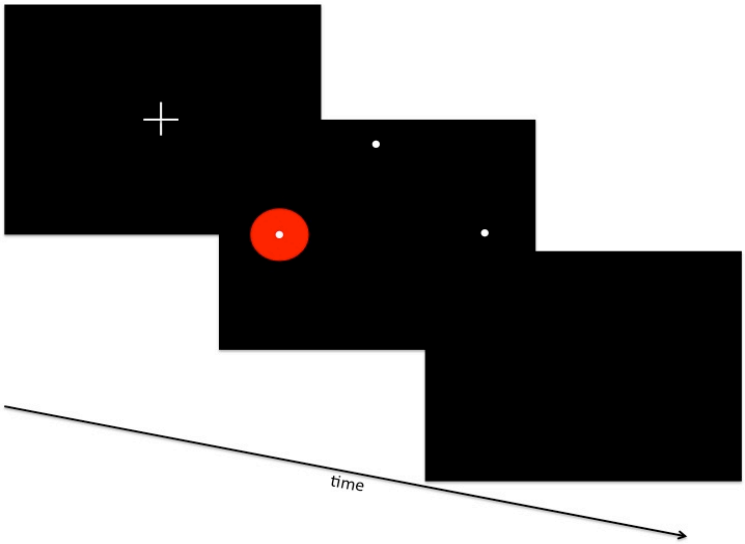


Figure 17. *Example of a typical trial in the 3-item version of the Simon task. The small white dots indicate touch-sensitive regions on the screen that can be selected for response.*

Experiment 1c – Letter Flanker

Participants

17 participants (7 male, mean age 19.21) participated in this experiment.

Apparatus, stimuli and procedure

A 3-button box was used to collect responses. Participant sat at a distance of 40cm from the screen and they were instructed to respond using their index, middle and ring fingers of their dominant hand to press the left, central and right button respectively. They were instructed to respond as fast and as accurately as possible.

The stimulus set consisted of 2 pairs of letters 0.7 degrees of visual angle in size (lower-case b/d and p/q), presented against a black background.

A typical trial was as follows. At the start of the trial, a white fixation cross was presented in the centre of the screen. Participants were instructed to press the central button when they were ready to start the trial. The fixation cross then disappeared, and a row of 5 letters appeared in the centre of the screen. The distance between the letters and the top/bottom and left/right ends of screen was 20 degrees of the visual angle. Position was varied by ± 1 one degree from one trial to another in order to reduce the possible confound of participants maintaining their focus on the same point of the screen in order to attend to the relevant stimulus. Participants were instructed to identify the central letter by pressing the right (for b and p) or left (for d and q) buttons, whilst ignoring the identity of the flanker distractors. The contingencies between response buttons and stimulus identity were alternated randomly across participants. The stimuli remained on the screen until a response was made or until 10 seconds had elapsed. Correct trials were identified by a high-pitch sound, and the correct choice remaining on the screen for one second, indicating positive feedback. Incorrect trials were identified by a low-pitched sound accompanied by the presentation of a large, gray circle, indicating negative feedback. After 2 seconds from response, the fixation cross was presented again on the screen, indicating the start of another trial.

Conflict levels were manipulated by changing the identity of the flanker distractors with respect to the central, target letter. In low-conflict trials, the distractors and the target letter shared the same identity. In high-conflict trials, the distractors and

target letter were two different letters from the same stimulus set pair. High- and low-conflict trials were presented in randomized order. Examples of a high- and a low-conflict trial are presented in Figure 18.

Participants carried out 2 blocks of 150 trials each. Reaction times and errors were recorded for analysis.

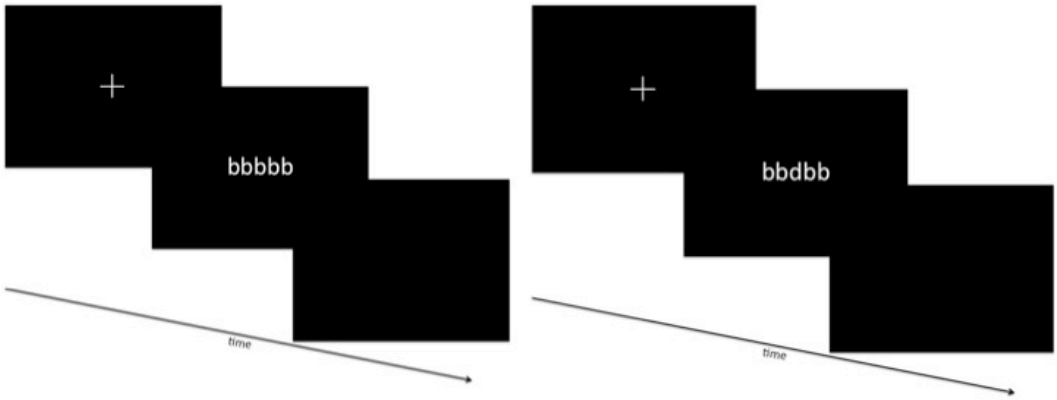


Figure 18. *Example of a low-conflict (left) and high-conflict (right) trial in the Letter Flanker task.*

Experiment 1d – Arrows task

Participants

13 participants (5 male, mean age 20.67) participated in this experiment.

Apparatus, stimuli and procedure

Participants sat at a distance of 40cm from the screen and they were instructed to respond using the index finger of their dominant hand to touch one of four locations

on the screen identified by white dots, as during experiments 1a and 1b. They were instructed to respond as fast and as accurately as possible.

The stimulus set consisted of a set of small white arrows, 0.7 degrees of visual angle in length, presented against a black background.

A typical trial was as follows. At the start of the trial, a white fixation cross (2 degrees in width and 2 degrees in height) was presented in the centre of the screen. Participants were instructed to touch the fixation-cross when they were ready to start the trial. The fixation-cross then disappeared, 7 arrows appeared scattered across the centre of the screen. Participants were instructed to respond by touching the side of the screen indicated by the majority of the arrows. The stimuli remained on the screen until a response was made or until 10 seconds had elapsed. Correct trials were identified by a high-pitch sound, indicating positive feedback. Incorrect trials were identified by a low-pitched sound accompanied by the presentation of a large, grey circle, indicating negative feedback. After 2 seconds from response, the fixation cross was presented again on the screen, indicating the start of another trial.

Conflict levels were manipulated by changing the percentage of arrows pointing in one direction or the opposite direction. In low-conflict trials, all arrows pointed in the same direction. In high-conflict trials, either two or three of the arrows pointed in one direction, while the remaining arrows pointed in the opposite direction. High- and low-

conflict trials were presented in randomized order. Examples of a high- and a low-conflict trial are presented in Figure 19.

Participants carried out 2 blocks of 150 trials each. Reaction times and errors were recorded for analysis.

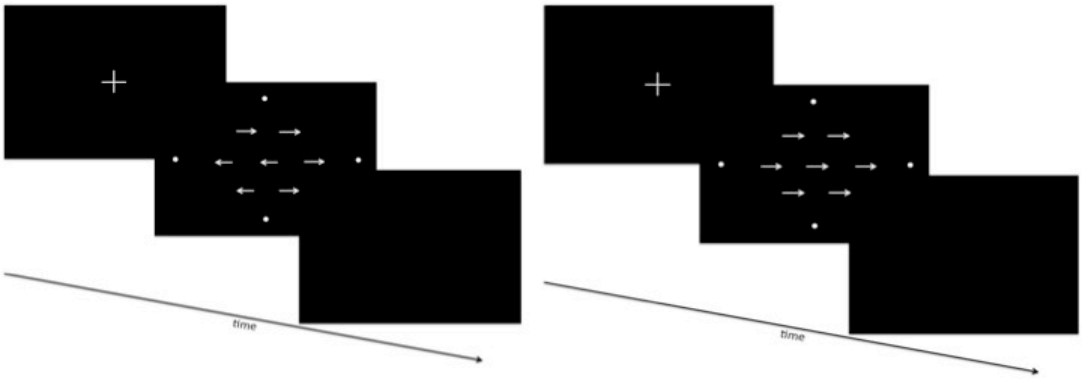


Figure 19. *Example of a low-conflict (right) and high-conflict (left) trial in the Arrows task.*

Experiment 1e – WCST analogue

Participants

22 participants (12 females, mean age 21.89) participated in this experiment.

Apparatus, stimuli and procedure

The task used in this experiment was identical to the WCST analogue described in Chapter 1. Participants carried out 2 blocks of 150 trials each. Reaction times and errors were recorded for analysis.

EXPERIMENT 2

Participants

31 participants (13 male, mean age 21.48) participated in this experiment. The participants were randomly assigned to one of two conditions, in a between-subject design. One group (16 participants) performed a version of the task in which 70% of the trials were high conflict trials, whilst the second group performed a version in which 30% of the trials were high conflict trials.

Apparatus, stimuli and procedure

The task used in this experiment was identical to the WCST analogue described in Chapter 1 and used in Experiment 1e. The only modification to the task involved varying the percentage of high conflict trials to create two conflict-frequency conditions. Participants carried out 2 blocks of 150 trials each. Reaction times and errors were recorded for analysis.

Results

EXPERIMENT 1

Experiment 1a – Simon task

A 2x2 repeated measures ANOVA was carried out on the Speed of Target Selection (STS measured as the reciprocal of response time – this measure was used in all tasks), with Current Conflict (High or Low) and Previous Conflict (High or Low) as factors. Results are illustrated in Figure 20. Exact stimulus-reponse repetitions were removed from analysis. There was a significant main effect of Current Conflict ($F_{(1,19)}$ = 36.33, $p < .001$), with higher STS values for Low-conflict trials (mean = 1.43, sd = .32) than High-conflict trials (mean = 1.29, sd = .31) and a main effect of Previous Conflict ($F_{(1,19)}$ = 11.91, $p = .003$), with higher STS values for trials following Low-conflict trials (mean = 1.39, sd = .31), than High-conflict trials (mean = 1.35, sd = .31). There was also a significant interaction between Current and Previous Conflict ($F_{(1,19)}$ = 13.66, $p = .002$), with a larger conflict effect following an L trial (mean = .19, sd = .13) than following an H trial (mean = .05, sd = .11).

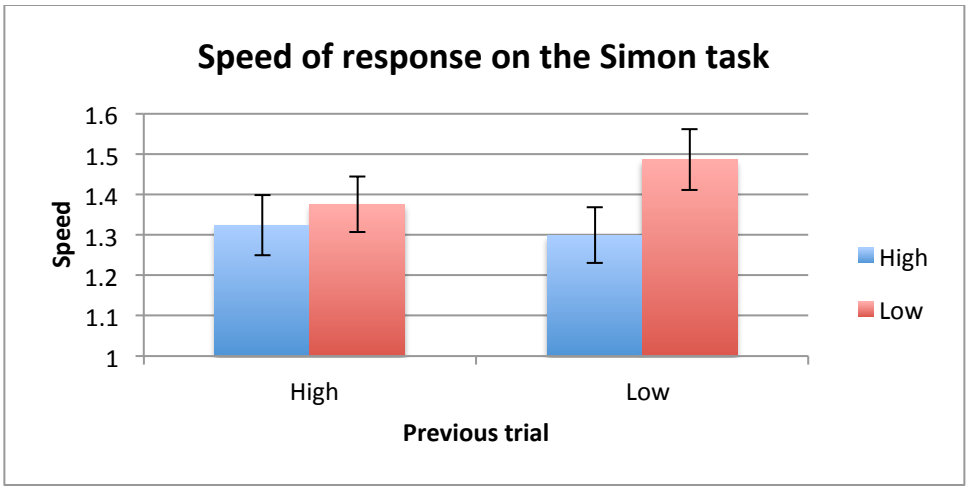


Figure 20. *Speed of response to high- and low-conflict trials, as a function of previous trial type (High or Low), in the Simon task. Faster responses to Low-conflict than high-conflict trials denote a conflict effect.*

Bonferroni-corrected paired samples t-tests revealed a significant difference between LL and HL trials ($t_{(19)}$ = 4.18, p = .001), with higher STS values when Low-conflict trials were preceded by Low-conflict trials (mean = 1.49, p = .34) than High-conflict trials (mean = 1.38, sd = .31). There was no significant difference between HH and LH trials ($t_{(19)}$ = 1.49, p = .153).

Experiment 1b – Simon task with 3 cues

A 2x2 repeated measures ANOVA was carried out, with Current Conflict (High or Low) and Previous Conflict (High or Low) as factors. Exact stimulus-reponse repetitions were removed from analysis. Results are illustrated in Figure 21. There was a significant main effect of Current Conflict ($F_{(1,19)}$ = 124.50, p < .001), with higher STS values for Low-conflict trials (mean = 1.28, sd = .20) than High-conflict trials (mean = 1.15, sd = .23) and a main effect of Previous Conflict ($F_{(1,19)}$ = 13.60, p = .002), with higher STS values for

trials following Low-conflict trials (mean = 1.19, sd = .22), than High-conflict trials (mean = 1.21, sd = .23). There was also a significant interaction between Current and Previous Conflict ($F_{(1,19)}$ = 54.97, p < .001), with a larger conflict effect following a Low-conflict trial (mean = .22, sd = .08) than following a High-conflict trial (mean = .16, sd = .08).

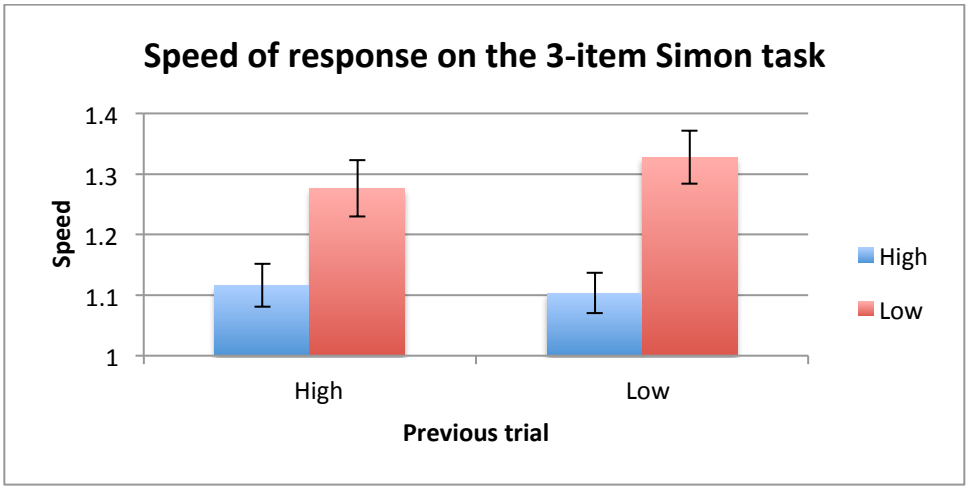


Figure 21. *Speed of response to high- and low-conflict trials, as a function of previous trial type (High or Low), in the 3-item version of the Simon task.*

Bonferroni-corrected paired samples t-tests revealed a significant difference between LL and HL trials ($t_{(19)}$ = 6.45, p < .001), with higher STS values when Low-conflict trials were preceded by Low-conflict trials (mean = 1.33, p = .19) than High-conflict trials (mean = 1.28, sd = .21). There was also a significant difference between HH and LH trials ($t_{(19)}$ = 2.46, p = .024), with higher STS values when High-conflict trials were preceded by High-conflict trials (mean = 1.12, sd = .16) than Low-conflict trials (mean = 1.103, sd = .15).

Experiment 1c – Letter Flanker

A 2x2 repeated measures ANOVA was carried out, with Current Conflict (High or Low) and Previous Conflict (High or Low) as factors. Exact stimulus-reponse repetitions were removed from analysis. Results are illustrated in Figure 22. There was a significant main effect of Current Conflict ($F_{(1,16)}$ = 69.33, p < .001), with higher STS values for Low-conflict trials (mean = 1.33, sd = .27) than High-conflict trials (mean = 1.17, sd = .22) and a main effect of Previous Conflict ($F_{(1,16)}$ = 8.28, p = .011), with higher STS values for trials following Low-conflict trials (mean = 1.27, sd = .26), than High-conflict trials (mean = 1.25, sd = .25). There was also a significant interaction between Current and Previous Conflict ($F_{(1,16)}$ = 6.79, p = .019), with a larger conflict effect after Low-conflict trials (mean = .17, sd = .08) than after High-conflict trials (mean = .14, sd = .08).

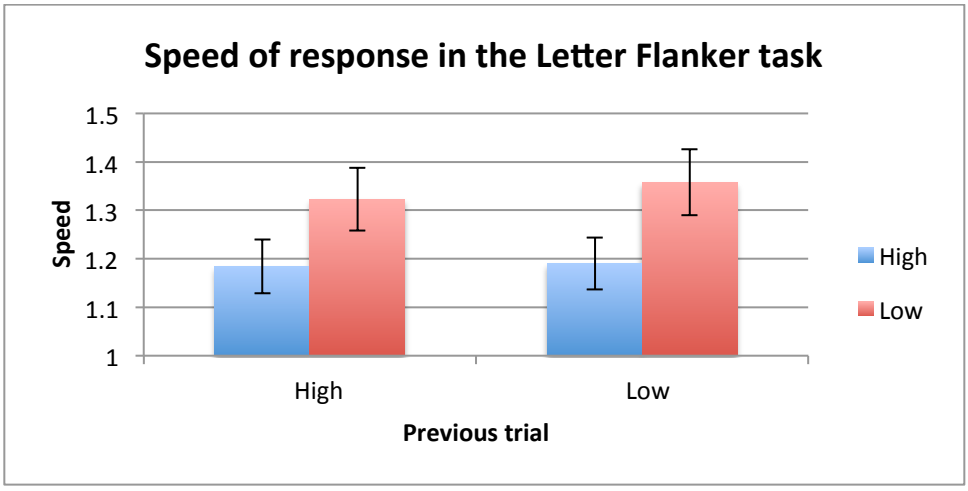


Figure 22. *Speed of response to high- and low-conflict trials, as a function of previous trial type (High or Low), in the Letter Flanker task.*

Bonferroni-corrected paired samples t-tests revealed a significant difference between LL and HL trials ($t_{(16)} = 3.41$, $p = .004$), with higher STS values when Low-conflict trials were preceded by Low-conflict trials (mean = 1.36, p = .28) than High-conflict trials (mean = 1.32, sd = .27). There was no significant difference between HH and LH trials ($t_{(16)} < 1$).

Experiment 1d – Arrows task

A 2x2 repeated measures ANOVA was carried out, with Current Conflict (High or Low) and Previous Conflict (High or Low) as factors. Results are illustrated in Figure 23. There was a significant main effect of Current Conflict ($F_{(1,12)} = 402.75$, $p < .001$), with higher STS values for Low-conflict trials (mean = 1.06, sd = .18) than High-conflict trials (mean = 0.62, sd = .20) and a main effect of Previous Conflict ($F_{(1,12)} = 33.44$, $p < .001$), with higher STS values for trials following Low-conflict trials (mean = 0.86, sd = .19), than Low-conflict trials (mean = 0.82, sd = .19). The interaction between Current and Previous Conflict did not reach significance, but indicated a trend towards a significant interaction ($F_{(1,12)} = 3.71$, $p = .078$). Trials following low-conflict trials showed a numerically larger conflict effect (mean = .45, sd = .09) than after high-conflict trials (mean = .41, sd = .08).

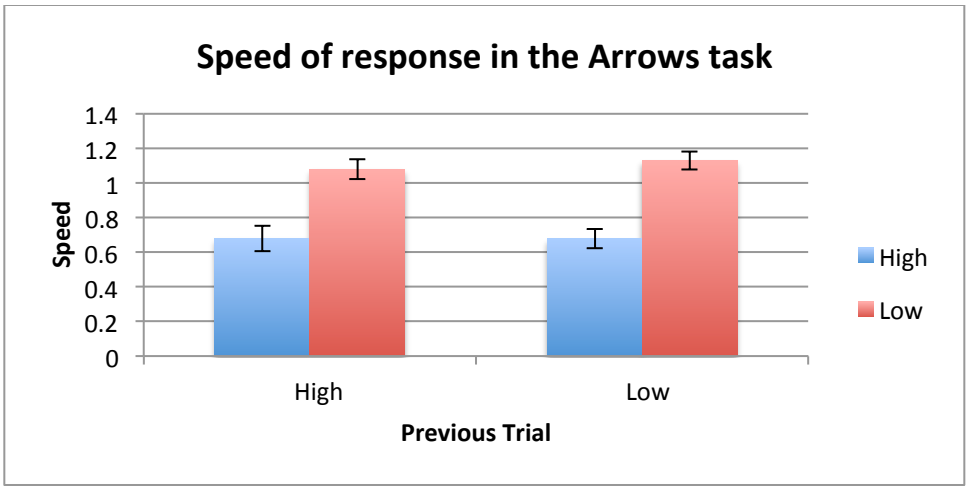


Figure 23. *Speed of response to high- and low-conflict trials, as a function of previous trial type (High or Low), in the Arrows task.*

Bonferroni-corrected paired samples t-tests revealed a significant difference between LL and HL trials ($t_{(12)}$ = 6.04, p < .001), with higher STS values when L trials were preceded by Low-conflict trials (mean = 1.09, p = .18) than High-conflict trials (mean = 1.02, sd = .17). There was no significant difference between HH and LH trials ($t_{(12)}$ = -1.74, p = .108).

Experiment 1e – WCST analogue

A 2x2 repeated measures ANOVA was carried out, with Current Conflict (High or Low) and Previous Conflict (High or Low) as factors. Results are illustrated in Figure 24. There was a significant main effect of Current Conflict ($F_{(1,21)}$ = 157.06, p < .001), with higher STS values for Low-conflict trials (mean = 1.71, sd = .28) than High-conflict trials (mean = 1.48, sd = .26) and a main effect of Previous Conflict ($F_{(1,21)}$ = 5.83, p = .025), with higher STS values for trials following High-conflict trials (mean = 1.61, sd = .26),

than Low-conflict trials (mean = 1.58, sd = .25). There was also a significant interaction between Current and Previous Conflict ($F_{(1,21)}$ = 12.21, p = .002), with a larger conflict effect after Low-conflict trials (mean = .27, sd = .1) than after High-conflict trials (mean = .19, sd = .09).

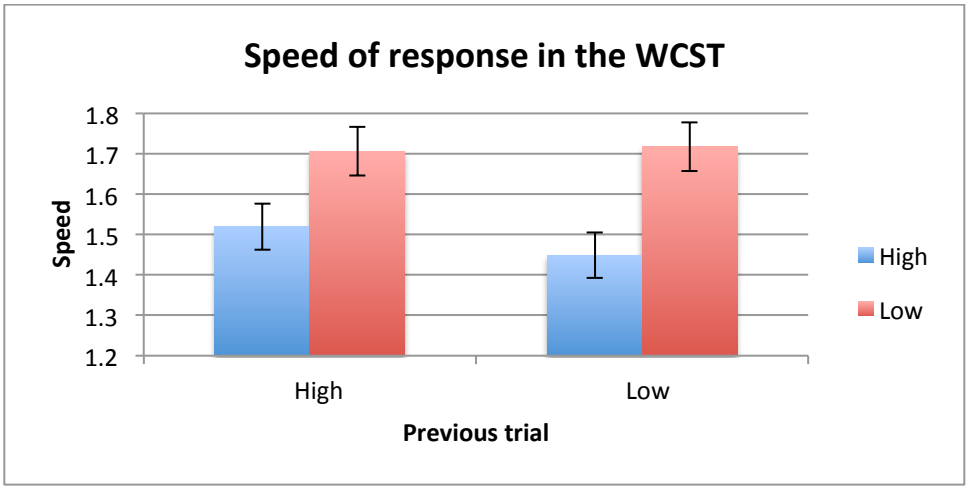


Figure 24. *Speed of response to high- and low-conflict trials, as a function of previous trial type (High or Low), in the conflict version of the WCST analogue.*

Bonferroni-corrected paired samples t-tests revealed a significant difference between HH and LH trials ($t_{(21)}$ = 3.93, p = .001), with higher STS values when High-conflict trials were preceded by High-conflict trials (mean = 1.52, sd = .27) than Low-conflict trials (mean = 1.45, sd = .26). There was no significant difference between HL and LL trials ($t_{(21)} < 1$).

EXPERIMENT 2

A 2x2x2 mixed ANOVA was carried out with Current Conflict (High and Low), Previous Conflict (High and Low) as within-subject factors and Proportion of high-conflict trials (30% and 70%) as the between-subject factor.

There was a significant main effect of Current Conflict ($F_{(1,29)} = 46.20$, $p < .001$)), with larger STS values for Low-conflict trials (mean = 1.38 , sd = .16) than High-conflict trials (mean = 1.31 , sd = .15). There was a significant main effect of Previous Conflict ($F_{(1,29)} = 7.94$, $p = .009$), with trials following High-conflict trials showing larger STS values (mean = 1.45 , sd = .16) than trials following Low-conflict trials (mean = 1.44, sd = .16), as well as a significant interaction between Current Conflict and Previous Conflict ($F_{(1,29)} = 4.23$, $p = .049$). There was no significant interaction between Current Conflict and Proportion ($F_{(1,29)} < 1$): the size of the conflict effect did not appear modulated by frequency of high-conflict trials (see Figure 25). There was no significant main effect of Proportion ($F_{(1,29)} = 1.91$, $p = .178$) but the three-way interaction between Current Conflict, Previous Conflict and Proportion showed a very strong trend towards significance ($F_{(1,29)} = 3.98$, $p = .056$).

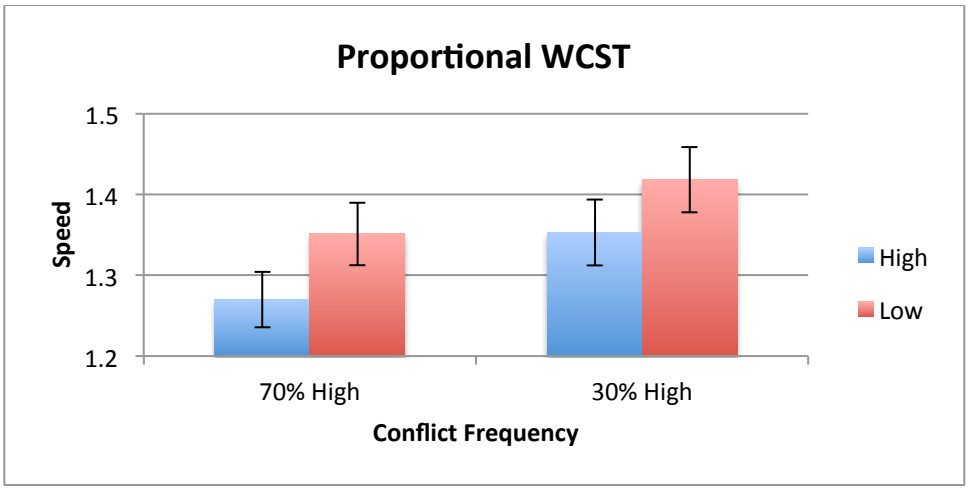


Figure 25. *Speed of response to high- and low-conflict trials, as a function of the proportion of high-conflict trials (70% = 70% high-conflict/30% low-conflict; 30% = 30% high-conflict/70% high-conflict) in the conflict version of the WCST analogue.*

In order to further investigate the trend 3-way interaction, we carried out two separate 2x2 Repeated-Measures ANOVAs, separately for the two groups. The analysis revealed a significant main effect of Current Conflict both in the 70% Group ($F_{(1,15)}$ = 24.24, $p < .001$) and in the 30% Group ($F_{(1,14)}$ = 21.99, $p < .001$), with higher STS values for Low-conflict trials than High-conflict trials. There was a significant main effect of Previous Conflict in the 30% Group only ($F_{(1,14)}$ = 4.69, $p = .048$), with higher STS values for trials following High-conflict trials. The interaction between Current and Previous Conflict was significant only in the 30% Group ($F_{(1,14)}$ = 5.53, $p = .034$). Paired-Sample T-Tests for the 30% Group, revealed that there was a significant difference between HH and LH trials ($t_{(14)}$ = 2.39, $p = .031$), with higher STS values for HH trials (mean = 1.37, sd = 0.15) than LH trials (mean = 1.32, sd = 0.16). There was no significant difference between HL and LL trials ($t_{(14)} < 1$). Results are illustrated in Figure 26.

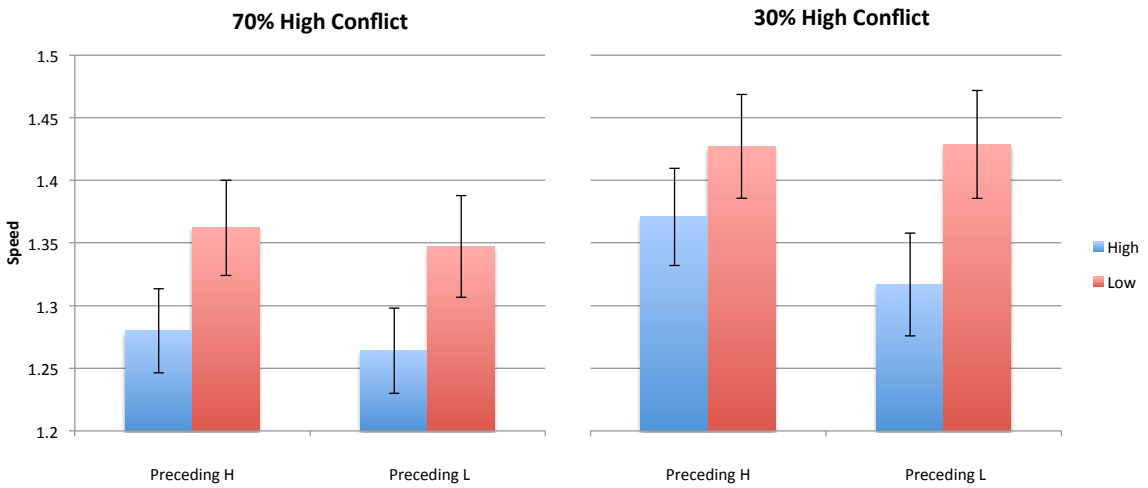


Figure 26. *Proportional version of the conflict-WCST analogue. Speed of response to high- and low-conflict trials, as a function of preceding trial type, presented separately for high frequency of high-conflict trials (70% High-conflict – left) and low-frequency of high-conflict trials (30% High-conflict – right).*

Discussion

The results from Experiment 1 largely confirmed our predictions. All conflict tasks showed a significant reduction in the conflict effect following high-conflict trials (see Table 2). However, these reductions were not due to the same underlying adaptation mechanisms in all tasks. Rather, they emerged as the result of modulation in the speed of response of different trial types in different tasks. In tasks where low-conflict trials could be facilitated by exploiting a habitual response (i.e. Simon, Flanker and Arrows), the reduction was mostly due to slower response-times on HL trials. On the other hand, in the WCST analogue, where no such facilitation is possible on low-conflict trials, adaptation was solely due to increases in HH speed. The 3-item Simon task showed both an increase in HH speed and a decrease in HL speed, indicating that both mechanisms might be at play in this task. This suggests that, while exposure to conflict does induce

adjustments in behaviour on the following trial in all tasks, the specific dynamics of these adjustments may vary depending on whether overall task performance can be facilitated by the exploitation of a direct route to response, or not. If a habit can be exploited, the direct route might be selected as more advantageous (i.e. default), and would be suppressed only when a change in conflict-level indicates it is no longer the most advantageous option. This is consistent with Sturmer and colleagues' dual-route framework (2002), suggesting that the direct route is suppressed when the task context indicates response-priming as a disadvantage. In the case of most conflict tasks, which are designed to elicit conflict between a habitual response and rule-guided behaviour, the direct route might be active by default and, once suppressed, lead to slower responses on HL trials. If there is no habit to exploit within a task, as in the case of the WCST, the indirect route might be selected as the default, and adaptation effects would manifest exclusively as enhanced rule-guided behaviour on sequences of trials that benefit from increased engagement of this route (i.e. HH trials).

Task	SIMON	SIMON 3cues	LETTER FLANKER	ARROWS	WCST
Conflict reduction	yes	yes	yes	trend	yes
HH vs LH	not sig	sig	not sig	not sig	sig
HL vs LL	sig	sig	sig	sig	not sig

Table 2. *Summary of significant and non-significant statistical differences in the speed of response between types of adaptation trials. Conflict reduction is calculated as the difference between the size of the conflict effect after a high-conflict trial and after a low-conflict trial (as per Equation 1).*

This might also help explain why a reverse conflict effect is often observed when conflict-frequency is increased (Hommel, 1994; Logan, 1979; Nieuwenhuis, Yeung, van den Wildenberg, & Ridderinkhof, 2003; Sturmer et al., 2002). Conflict-frequency might act as a contextual cue (similarly to overall task demands, or to trial-to-trial changes in conflict levels) to indicate which route is currently more advantageous. For example, during a Simon task, if the proportion of high-conflict trials is high, exploitation of the direct route is no longer advantageous, but exploitation of the indirect route becomes beneficial. Therefore this may lead to a reduction in facilitation on low-conflict trials (due to direct route suppression) and/or an increase in speed of response on high-conflict trials (which benefit from the increased engagement of the indirect route), canceling out the conflict effect. The reduction of facilitation on low-conflict trials, for tasks that rely heavily on the direct route, may even exceed the enhancement in performance on high-conflict trials, making high-conflict trials faster.

For a task where low-conflict trials cannot be facilitated by exploiting the direct route, the WCST, manipulating conflict frequency had a different effect on response speed. Manipulation of the frequency of high-conflict trials did not affect the overall conflict effect in our WCST analogue, where low-conflict trials were always faster than high-conflict trials. The results from the WCST seem to indicate that sustained engagement of the indirect route, as should be the case in contexts of frequent exposure to conflict, does not speed up high-conflict trials, suggesting that there might be a ceiling for performance on this type of trials that cannot be exceeded simply by

sustained enhancement of task-relevant information or instructions. These results are in line with Jacoby and colleagues' (2001) findings that manipulating conflict-frequency in the Stroop task (which is a task that can rely on habitual responses for facilitation) only influenced word-reading processes, not color-naming. Increasing the number of high-conflict trials had therefore a larger influence on the suppression of the direct route (and, consequently, on facilitation for congruent items), than on the engagement of the indirect route (which supports performance for incongruent items). High-conflict trials in the WCST were affected by conflict-frequency (albeit not significantly) only when we considered modulation related to specific trial-sequences. We found that HH trials were faster than LH trials only in the low-frequency condition. When the frequency of high-conflict trials is high, sustained engagement of the indirect route due to less frequent changes in conflict levels (as the number of low-conflict trials interspersed amongst high-conflict trials, and therefore LH trials, is significantly lower) might reduce the costs of re-engaging the relevant rule on LH trials, effectively making LH trials as fast as HH trials.

The only task not showing a significant reduction in the conflict effect was the Arrows task, although a strong trend toward significance was present. The findings concerning the Arrows task might be due to limitations in the design. Although low-conflict trials presented only one level of facilitation (i.e. all arrows pointed in one of four directions), there were effectively two levels of conflict during high-conflict trials, depending on the configuration of the trial. Some high-conflict trials had 5 of the arrows

pointing in one direction and 2 pointing in the opposite direction, making deduction of the correct direction easier to determine than on other high-conflict trials where 3 of the arrows pointing in the opposite direction. This may influence performance on high-conflict trials overall and, as a consequence, the difference in speed between high- and low-conflict trials observed after high-conflict trials. These trials should have been analysed separately, but were coded as high-conflict, making distinction post-testing impossible. Nonetheless, as predicted, HL trials were significantly slower than LL trials, while no such difference was observed between HH and LH trials. This suggests that when the usefulness of the indirect route is limited, and rule-guided behaviour cannot provide as effective a shortcut to response, adaptation effects are still present but restricted to decreased speed of response on HL trials.

Results therefore indicate that adaptation effects can be due to both reduction in facilitation on low-conflict trials and performance-enhancement on high-conflict trials, and that the extent to which one or the other contributes more to adaptation seems to be task-specific. The task-specificity of this phenomenon might be due to which route is most advantageous to the current context/task overall (i.e. the direct route if the context/task primes the exploitation of a habit, or the indirect route if it does not). This theoretical framework can be seen as complementary to models suggesting the presence of mechanisms that suppress or enhance task-irrelevant or task-relevant information (Depisapia & Braver, 2006; Egner & Hirsch, 2005). Rather than (or alongside) information being suppressed or enhanced, it may be task-advantageous

or disadvantageous routes that are being modulated. This could also accommodate many of the explanations provided by the three prominent accounts of adaptation: conflict-monitoring, routes to response and feature-integration.

Chen and Melara (2009) stated that since adaptation effects can be found for sequences of neutral trials that do not elicit conflict (i.e. trials where competition is created between two arbitrary stimulus response associations, in the case of Chen and Melara's study, these would be Simon-type trials that involve a right or left button-press in response to the identity of items that appear at the top or the bottom of the screen), adaptation can be explained by feature-integration and unbinding costs alone, rather than conflict-monitoring. However, they found that unbinding costs were larger for Simon than for neutral trials (and so were adaptation effects, albeit not significantly), suggesting that unbinding costs are not the same for the two types of trials: unbinding on Simon HL trials is somehow more costly than on neutral HL trials. Reduction in facilitation on HL trials, due to direct route suppression (in addition to unbinding), can explain this phenomenon. If a habit needs to be suppressed, the reduction in facilitation, combined with unbinding costs, will result in higher response speed costs.

Furthermore, the type of neutral trials in Chen and Melara's study, although not eliciting conflict between a habit and an instructed response, are designed to elicit a type of competition between two instructed responses that is not unlike the one in the conflict analogue of the WCST. This type of trials, even if not benefitting from facilitation

advantages, may be exploiting the same indirect route exploited by high-conflict trials, as they essentially constitute rule-guided behaviour. Competition is therefore elicited *within* the indirect route, instead of between routes. When competition is elicited within the same route, between instructed responses (like in the WCST analogue), low-conflict trials are not affected by previous conflict, whereas high-conflict trials are, suggesting that even when the direct route is not exploited at all, adaptation can be present in the absence of unbinding costs or repetitions, in the form of enhanced performance on HH trials. Feature-integration might be one of the processes carried out by the indirect route. Also, whether trial N_{-2} was high- or low-conflict may have an influence on the route suppressed on the current trial, N. Therefore, the finding of adaptation effects for neutral trials is not necessarily incompatible with the idea of different routes to response.

The conflict-monitoring model posits that conflict-induced adaptation is driven by enhanced attention and cognitive control in response to conflict detection. This would generally be expected to improve performance on high-conflict trials. Although we have seen that, in many common conflict tasks, adaptation is driven for the most part by reductions in facilitatory effects on low-conflict trials, rather than increased speed of response on high-conflict trials, this is still consistent with a role for a conflict-monitoring system in adaptation effects. The presence of increased activation in brain structures such as the ACC and dlPFC during LH and HH trials (e.g. Durston, 2003; Kerns, 2004; 2006; Horga et al., 2012; Egner & Hirsch, 2005) and the impairments in adaptation

observed after damage to these structures (di Pellegrino, Ciaramelli, & Làdavas, 2007; Mansouri et al., 2007; Sheth et al., 2012; Swick & Turken, 2002), indicate that conflict-modulated recruitment of control does play a role. It is not presence of conflict on the current trial per-se that triggers adaptation, but, rather, changes in recent conflict-history modulate selection, suppression and biasing of the default route to response in a given context or task. Most, if not all, of the tasks used in the literature, where activation related to conflict-induced adaptation has been found, are likely to exploit the direct route as the default route to response. This route needs to be suppressed on LH trials, when there is a change in conflict levels (from low- to high-conflict) indicating that it is no longer advantageous. If the role of the ACC is to detect changes in context that require the suppression of one route when competing with another, it may explain why ACC is found to be more active on these trials. It may also explain why ACC lesions do not appear to affect responses in tasks where there is no competition between the direct and indirect route, like the WCST.

Consistent with the idea of a role for ACC in suppressing the default/most advantageous route to response, evidence also shows that the N2 component, which is thought to reflect pre-response conflict and originate in the ACC, reverses in a Go/No-Go task with higher proportions of No-Go trials, increasing in amplitude for Go, as opposed to No-go, trials (Nieuwenhuis et al., 2003). Similarly, Wilks and colleagues (2012) have shown that ACC and dlPFC activation increases on low-conflict trials when these become less frequent, and decreases on high-conflict trials as they become more

frequent. This indicates that ACC and dlPFC are not merely signaling the need to suppress specifically the direct route (or to enhance task relevant information) during high-conflict trials, but might instead signal contexts where the current default route is no longer advantageous, and a change must be implemented: in the case of tasks with a high frequency of conflict trials, the indirect route could be more advantageous overall, and selected as a default. In this context, low-conflict trials would represent a situation where the default route is no longer the most advantageous, triggering the need for adjustment, and, as such, ACC and dlPFC might become more active on low-conflict than high-conflict trials.

There is a substantial body of literature suggesting a role for ACC in processing reinforcement history to guide action selection (Kennerley, Walton, Behrens, Buckley, & Rushworth, 2006), assessing advantageous action states (Newman & McGaughy, 2011; Rudebeck et al., 2008; Rushworth, Walton, Kennerley, & Bannerman, 2004) and being sensitive to environmental uncertainty and volatility (Behrens, Woolrich, Walton, & Rushworth, 2007; Rushworth & Behrens, 2008) as well as surprising, novel and unexpected events (Aarts, Roelofs, & van Turennout, 2008; Badre & Wagner, 2004; Brown & Braver, 2008; Bush et al., 2002; Nee, Kastner, & Brown, 2010; Hayden, Heilbronner, Pearson, & Platt, 2011). This would be consistent with the idea that ACC is signalling the need for behavioural adjustments in response to changes in conflict-levels in recent history, rather than signalling the presence of conflict per-se. Indeed, Stoffels (1996) argued that when high- and low-conflict trials are intermixed in a single block,

rather than presented in separate blocks, as it is the case in all conflict tasks in the literature, it is uncertainty that is being manipulated, rather than conflict. Conflict is already present in high-conflict blocks, but when these are mixed with low-conflict trials, it is only the latter that are affected, in the form of a decrease in response speed (Stoffels, 1996). As the task used by Stoffels (a task requiring participants to respond to the visual properties of an arrow while ignoring the direction it points towards) benefits from direct-route exploitation on low-conflict trials, their results are consistent with our findings that such tasks would show adaptation in the form of reduction in facilitation.

Similarly, dlPFC also has been shown to encode recent conflict-history (Mansouri et al., 2007) and to be sensitive to changes in recent-conflict history (Durston, 2003). If the role of the dlPFC is to maintain a memory trace for recent conflict-history and respond to changes in recent conflict-history, when this memory trace cannot be formed or is lost, for example due to lesions or interference with dlPFC activity, the cognitive system might reset to the default route to response. In most conflict tasks designed around competition between a habit and an instructed response, where the default route is the direct route, resetting to the default (i.e. direct) route would result in impairments in adaptation: default route activation would increase the speed of response on HL trials, and, consequently, increase the difference in speed between HH and HL trials, therefore reinstating the conflict effect normally reduced after high-conflict trials.

To summarize, we have confirmed that adaptation effects are due to both reduction in facilitation on low-conflict trials and enhancement in response speed on high-conflict trials. Adaptation effects manifest more prominently as one or the other depending on whether a task, or context, allows for the exploitation of a direct route supporting habitual, overpracticed, responses or creates competition between instructed responses (i.e. within an indirect route supporting rule-guided, instructed, behaviour). The WCST, differently from other conflict tasks, does not elicit conflict between a habit and instructed responses, and is therefore unlikely to rely on the direct route for performance. This results in adaptation effects that are exclusively restricted to faster responses on HH trials. In tasks where the direct route can be exploited, adaptation effects also, or exclusively, include reductions of facilitatory effects on low-conflict trials. Modulation of the routes to response may be triggered by changes in the conflict level in recent history, rather than the presence or the absence of conflict on the current trial, and be dependent on which route is more advantageous for the overall context or task. The ACC and dlPFC are two good candidate regions for the implementation of these mechanisms at the neural level, modulating the activation within the two competing routes to response.

Based on these findings, we can attempt to speculate on the specific mechanisms implemented by cognitive control to drive behavioural adaptation. Cognitive control might implement two distinct, dissociable mechanisms to proactively reduce the conflict-related costs on behaviour: caution (which may result in a reduction

of facilitation, and therefore slower responses, on HL trials) and attention (which may result in better task-relevant processing, speeding up responses on HH trials). Caution may contribute to suppressing the disadvantageous route (often the direct route) and inhibition of task-irrelevant information. Attention, on the other hand, may contribute to biasing the advantageous route (often the indirect route) and to enhance task-relevant information. In the following chapter, we will look at how these two processes could be quantified in the parameters of a modelling framework for reaction-time data, the Linear Ballistic Accumulator. Furthermore, no study, to the best of my knowledge, has looked at how ACC activation might be modulated in contexts requiring the suppression of a habitual response (i.e. competition between-routes), as opposed to competition between two instructed responses (i.e. competition within-route). In the fifth chapter, we attempted to investigate this question, by administering the WCST and Simon tasks to a small cohort of patients fitted with Deep Brain Stimulation implants that allowed us to record localized neural activity from deep within the ACC during execution of these tasks, as well as test the effects of stimulation on behavioural measures of performance.

CHAPTER 4

Quantification of conflict-induced adaptation components in the Linear Ballistic Accumulator model of choice response time

Introduction

Response time (RT) data is the most commonly analysed form of data in experimental psychology and neuroscience. Inferences about the decision-making processes involved in a range of cognitive tasks are often drawn on the basis of RT data as a unified measure representing the speed of response. However, RT would be more adequately described as a measure of latency, of which speed is only one component. Complex effects on mean RT values and distributions, due to the interactions between speed, accuracy and task difficulty, have been observed, even within individual participants (Ratcliff & Rouder, 1998). While two mean RTs might appear numerically identical, they could result from very different underlying dynamics. Similarly, two values might appear statistically different, but this difference may not be due to differences in the same underlying dynamic (Wagenmakers, Krypotos, Criss, & Iverson, 2012). Mathematical models of RT data allow greater depth of analysis of RT values to obtain a more complete picture of the decision-making process and how it is affected by different experimental manipulations (Wagenmakers, van der Maas, & Grasman, 2007).

Several mathematical models have been developed over the last few decades to account for variability in RT distributions (e.g. Brown & Heathcote, 2005; Ratcliff & Rouder, 1998; Ratcliff & Tuerlinckx, 2002; Ratcliff, 1978; Reddi & Carpenter, 2000; Usher & McClelland, 2001; van Maanen et al., 2011; Wagenmakers et al., 2007). Despite their different mathematical approaches, the basic decision-making process in these models is based on the concept of sequential sampling (Stone, 1960). Sequential sampling describes decision-making as the continuous accumulation of evidence, sampled from the environment, in favour of one choice or another, until enough evidence has been accumulated for one of the options to be selected as the response. All models share three main basic elements, or parameters, that interact to determine RT. The first element is the speed at which evidence is sampled from the environment, known as the *drift rate* (*d*). The second element is the amount of response caution implemented by the subject in order to maximize accuracy, calculated as the difference between the *response threshold* (*b*, the total amount of evidence necessary to select a particular response) and the *start-point* (*a*, the point at which the evidence starts to accumulate at the beginning of the decision-making process). The third, and final, element is the *non-decision* time (*T0*), which represents the amount of time due to non-decision aspects of the task (such as the time to process stimulus information, or to physically make the response). The interaction between these three elements determines the RT value, and is sufficient to describe many benchmark phenomena commonly observed in experimental settings. For example, phenomena such as speed-accuracy tradeoffs, response biases and post-error slowing can all be described as modulations in threshold

values (Donkin, Brown, & Heathcote, 2011). An example of the application of the LBA to investigate response biases is illustrated in Figure 27 below.

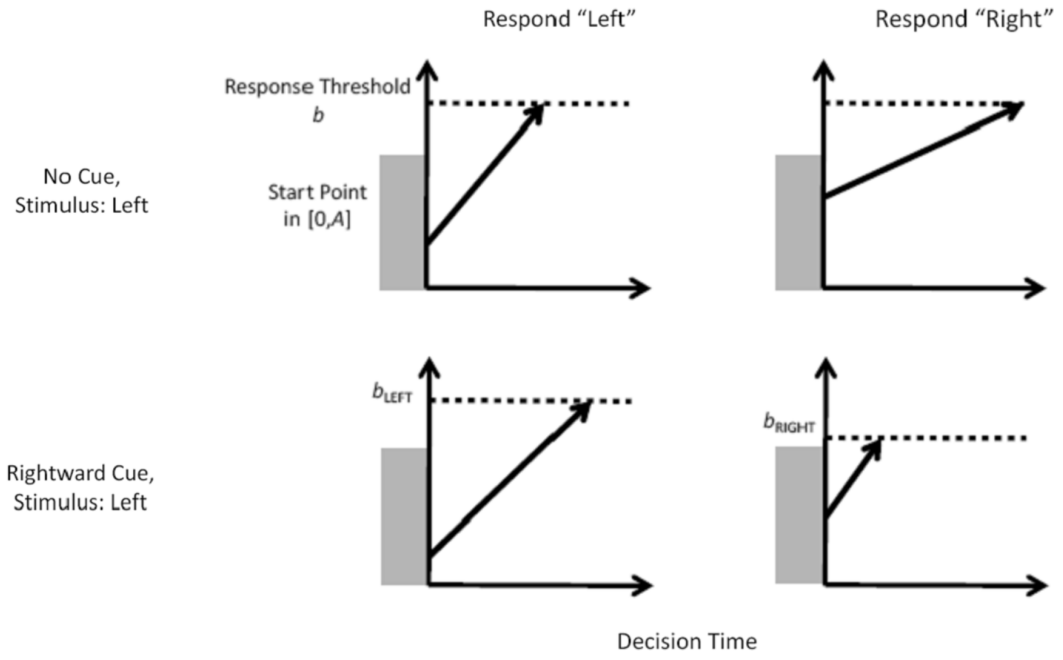


Figure 27. *Example of an LBA application to a cued random-dot motion paradigm. Participants had to determine the direction of motion of a cloud of dots that could be preceded by no cue or by a valid or invalid cue. The top row depicts the accumulators for the correct (left) and incorrect (right) choices during an un-cued trial and the bottom row depicts the accumulators for the correct and incorrect choices during an invalid cued trial. In the presence of a cue, response threshold (dotted line) is lower for the accumulator corresponding to the cued direction (bottom right accumulator). (Figure adapted from Forstmann, Brown, Dutilh, Neumann, & Wagenmakers (2010), pp. 4).*

Differences in response time between conditions, groups or even individual participants might be due to variability in any of the aforementioned parameters that describe RT distributions. These models have been successfully applied to investigate a variety of cognitive processes. One of the first applications of this class of choice-time

models was carried out by Ratcliff (1978), who looked at how evidence in favour of different memory traces accumulates depending on the similarity between a probe and specific memories. Since then, the models have been applied to many other processes, such as decisions based on visual discrimination (Gold & Shadlen, 2001; Reddi, 2001), saccade inhibition (Hanes & Carpenter, 1999), speed-accuracy tradeoffs (Brown & Heathcote, 2005; Forstmann, Anwander, et al., 2010; Reddi & Carpenter, 2000), and even more complex phenomena such as reading (Ratcliff, Gomez & McKoon; 2004). In this instance, we chose to analyse the RT data obtained from some of our conflict tasks using the Linear Ballistic Accumulator (LBA) model developed by Brown and Heathcote (2008) to investigate the contribution of these different decision-making components to conflict-induced behavioural adaptation.

We chose to use the LBA, as opposed to other models, for several reasons. The first reason is the inherent simplicity of the LBA. The LBA is the simplest and most parsimonious model of its kind. As the name suggests, the LBA comprises two or more linear (i.e. where evidence accumulation does not fluctuate over time), ballistic (i.e. non-stochastic, where the trajectory of accumulation is determined once the initial conditions are known) independent accumulators. Most models assume that the accumulation of evidence is stochastic, with fluctuations in the amount of evidence available from moment-to-moment (e.g. Usher & McClelland, 2001; Brown & Heathchote, 2005). The LBA, instead, uses linear accumulators, where evidence accumulates linearly until threshold is reached, with no moment-to-moment variability.

As a consequence, it requires a smaller number of parameters than most other competing models (Brown & Heathcote, 2008). This significantly reduces the computational demands, which is particularly advantageous when modelling data with several responses, conditions or levels. Secondly, it is easily adaptable to tasks with more than two possible response options, thanks to separate, independent accumulators for each possible response, unlike most other models that use a single accumulator with the correct and incorrect choice competing to reach their threshold (e.g. Ratcliff's Diffusion Model, 1978; 1988; 2001).

Despite its simplicity, the LBA has been shown to account for most empirical phenomena as successfully as more complex models (Brown & Heathcote, 2008; Donkin et al.; 2011). At the same time, while even simpler models do exist, these present some shortcomings that can be overcome by the LBA, such as the inadequacy in describing the relative speed of correct and incorrect responses (Wagenmakers et al., 2007), or the difficulty in predicting incorrect responses (Reddi & Carpenter, 2000), due to the lack of separate accumulators for each response. The LBA, therefore, provides a satisfactory middle ground between simplicity and explanatory power.

The LBA has been successfully applied in recent years to a number of cognitive tasks. Forstmann and colleagues have carried out extensive investigations of how the influence of speed/accuracy emphasis on performance, during visual discrimination tasks, may be described in terms of changes in LBA threshold values, and how these

changes may be correlated with neural activity (Forstmann et al., 2008; Forstmann, Anwander, et al., 2010; Ho et al., 2012). Ho and colleagues (Ho, Brown, Abuyo, Ku, & Serences, 2012) investigated the mechanisms underlying attentional enhancement and how these correlate with drift rate values. The single-trial extension of the LBA has been applied to describe correlations between trial-to-trial fluctuations in response caution and hemodynamic changes (van Maanen et al., 2011). Only very recently, it has begun to be applied to the investigation of the mechanisms underlying performance in conflict tasks. King and colleagues (King, Donkin, Korb, & Egner, 2012) recently applied the LBA to investigate conflict-frequency effects in the Flanker task, and found variations in response thresholds for infrequent trial types, suggesting that frequency-modulation influences performance through response caution rather than evidence accumulation.

We saw in previous chapters (see Chapter 3) how adaptation effects may be due to different underlying mechanisms depending on whether the task, or context, promotes the exploitation of a direct route supporting habitual, overpracticed responses, over an indirect route supporting rule-guided behaviour, or viceversa. In tasks where no habitual response can be exploited to facilitate performance on low-conflict trials (e.g. the WCST), adaptation effects appear to be restricted to modulation of speed of response on high-conflict trials, such as increases in speed for HH trials, possibly due to enhanced indirect route biasing. In tasks where a habitual response can be exploited, adaptation effects can be due to both a reduction in facilitation on low-conflict trials (possibly due to suppression of the direct route) and increase in speed of

response on high-conflict trials, as a function of the level of conflict on the previous trial, resulting in slower HL trials and faster HH trials, respectively. The LBA could provide a way to dissociate and quantify these mechanisms. Suppression of the direct route could be reflected in higher response caution on trials where the participant would normally exploit this route for facilitation. Sustained biasing of the indirect route, on the other hand, could result in more efficient processing of task-relevant information, and this might be reflected in faster accumulation of evidence in favour of the correct response.

We therefore applied the LBA model to the data collected from the behavioural studies illustrated in Chapter 3, specifically to the WCST and the Simon tasks. We specifically chose these tasks because they allow for a direct comparison between a task that only involves conflict between two instructed responses, and therefore does not exploit the direct route (the WCST, which shows adaptation due exclusively to enhanced performance on HH trials), a task that specifically involves conflict elicited between a habit and an instructed response, and therefore between the direct and indirect route (the 2-item version of the Simon task, which showed adaptation due exclusively to reduced speed of response on HL trials) and a task that showed adaptation due to both enhanced speed of response on HH trials, and reduced speed on HL trials (the 3-item version of the Simon task).

For tasks that involve competition between a habitual response and an instructed response (i.e. the Simon tasks), we expected to observe higher response

caution values on HL trials than on LL trials, due to suppression of the direct route from the previous trial, which results in slower response times. We did not expect to observe this in the WCST, as this is a task that involves competition between two instructed responses, and does not exploit the direct route to response. We further expected increased speed on HH trials to be characterized by higher speed of evidence accumulation on these trials compared to LH trials, as drift rates are assumed to be dependent on attentional processing and the efficiency of the cognitive system (Donkin et al., 2011). For all tasks, we expected to observe higher response caution high-conflict trials than on low-conflict trials, due to a combination of habit-suppression and effects due to increased task difficulty.

Methods

Behavioural data

We applied the LBA model to the behavioural data obtained from our studies utilizing the 2- and 3-item versions of the Simon task, originally described in Chapter 3, and the WCST analogue originally described in Chapter 1. As we were only interested in modelling adaptation effects, errors due to switch-trials in the WCST task were excluded from the fitting.

LBA model

The LBA comprises one accumulator (see example in Figure 28) for each response option. The amount of time necessary for an accumulator to select a response (the RT) is given by:

$$\frac{b-a}{d}+T0$$

where *d* is the drift rate, *b* is the threshold, *a* is the start point at which evidence begins to accumulate at the beginning of the trial, and *T0* is the non-decision time. The drift rate is sampled from a normal distribution N (μ, σ), whereas the start point is sampled from a uniform distribution U [0,*A*], where *A* is the top of the distribution. *A*, therefore, represents the maximum amount of evidence that can be present in an accumulator before each trial. These two parameters are sampled randomly and independently for each trial. Evidence begins accumulating at the start of the trial from *a*, at a speed *d*, and the decision process ends once one accumulator reaches its threshold *b*, resulting in the corresponding response.

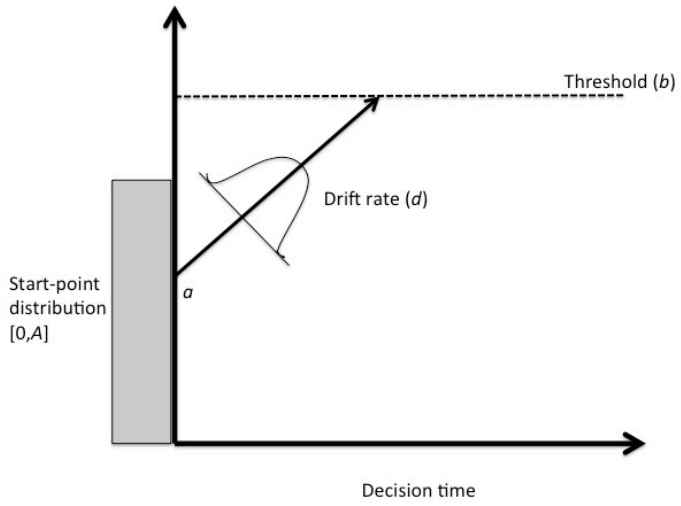


Figure 28. *Example accumulator in the LBA model. The start point of the accumulation process (a) is randomly sampled from the uniform distribution along the Y-axis (gray area), whereas the drift rate (d) is randomly sampled from a normal distribution with mean μ and standard deviation σ. Both a and d are sampled independently from the distribution for each response.*

For H trials in the WCST, there were three possible response options: the correct option, a relevant-incorrect option, which shared a feature with the correct option (colour or shape), and an irrelevant-incorrect option that shared features neither with the correct option nor with the other incorrect option. The LBA generally allows for three accumulators to represent each of these options. However, as none of our WCST participants picked the irrelevant incorrect option during testing, we omitted the third accumulator in our final model fit. This is justified because in the data the probability that this accumulator is selected is zero, and therefore it does not influence the defective probability that one of the other accumulators is selected.

Evidence indicates that, ideally, at least 100 observations per condition should be collected in order to reduce the risk of drift rate overestimation, particularly in high-accuracy conditions (Donkin et al., 2009). Given that our data included less than 100 trials per condition per participant (on the order of approximately 70-80 trials per condition), it was unlikely to provide reliable individual fits. We therefore averaged the experimental data across participants using a Vincentizing procedure (Vincent, 1912), which allows calculation of the group-average response time distribution (thus calculating averages out of hundreds of observations per condition), while preserving the information (location, scale and shape) of the individual response time distribution. The LBA model was then fit to the Vincentized data.

In the most flexible form of the LBA model, all five parameters previously mentioned may be allowed to vary for each accumulator: the mean (*v*) and standard deviation (*s*) of the drift rate (*d*) distribution, the threshold *b*, the top of the start point (*a*) distribution, *A* and the non-decision times *T0*. For our experimental design, this means we could have up to 5 parameters that vary for each response option across different combinations of stimuli and adaptation conditions (HH, LH, HL and LL). This results in hundreds of potential parameters combinations. However, a-priori assumptions can be applied to reduce the number of combinations, and therefore models, that may be fit to the data.

For example, for the WCST, we assumed that, while all parameters could vary across the four levels of the adaptation condition, they would not vary depending on the physical properties of the items in the stimulus set (e.g. whether it was composed of a red square, yellow triangle, etc.), nor depending on the physical properties of the choice item. We made similar assumptions for the Simon tasks, where we also assumed parameters would not vary depending on stimulus properties or positioning. We further assumed that parameters would not vary depending on whether the response was made by touching the top, left, or right of the screen, given that participants always started from a point in the centre of the screen which was equidistant from the response locations. For both tasks, we expected the drift rate parameters (*v* and *s*) and the threshold *b* to be the parameters more likely to reflect adaptation effects (and therefore vary across the adaptation condition).

We applied sequential model building (see Donkin et al., 2011) to find the best fitting model for our data. This involved starting with the simplest version of the model (with the least number of free parameters) and gradually increasing its complexity by including more free parameters in different combinations. We tested all possible combinations where the five parameters varied across the adaptation condition, for a total of 32 models.

Testing these combinations up to the most complex model (the one with all free parameters) allowed us to draw inferences as to which parameters were more likely to

be fixed across adaptation levels, in order to test further combinations of parameters while reducing the computational demands of running all possible permutations. For example, the best fitting models for the WCST all included free drift-rate parameters, whereas *T0* was fixed. Viceversa, many of the best fitting models for the Simon task included fixed drift-parameters. Therefore, we selected the best fitting model out of the 32 fit to the data through sequential building, and tested versions of these models that allowed each parameter to vary in turn only depending on the conflict level on the previous trial or current trial, while the other parameters were fixed or varied depending on their settings in best model. As *A* represents the amount of evidence already present in the accumulator before the start of the trial, we were interested in testing whether *A* may vary only depending on the previous trial, rather than current trial type, and viceversa. We also tested combinations where *T0* was fixed for high- and low-conflict trials, regardless of previous trial type, as non-decision time might be influenced only by the current level of conflict. This resulted in a total of 35 models for each task.

Model fitting was carried out using an adaptation of the R (www.r-project.org) script developed by Donkin and colleagues (2009), and run on Mac OS X. The script searched for parameters that can generate predictions for accuracy and response-time data that best resemble the experimental data. The similarity between predicted and observed data was calculated using Quantile Maximum Products Estimation (QMPE; Heathcote & Brown, 2004; Heathcote, Brown & Mewhort, 2002). The starting

parameter values of the model were estimated using automatic heuristics calculations based on the data and optimized using a SIMPLEX search algorithm (Nelder & Mead, 1965). This algorithm searched for the optimal parameter values that minimized the difference between predicted and observed data, calculated by the QMPE method described earlier. Once the best possible set of parameter values were identified, Bayesian Information Criterion (BIC; Schwartz, 1978) values were calculated for the model. These values are a measure used to identify over-fitting (i.e. the presence of extra parameters that only account for unsystematic variation in the data), as they penalize the model for including extra parameters. We selected the model with the lowest BIC value as the best fit, according to the criteria outlined by Raftery (1994).

Results

We calculated the likelihood of each model being the best fit for the data over other models using the procedure for calculating Schwarz weights (Schwarz, 1978) outlined in Wagenmakers and Farrell (2004). Schwarz weights ($\mathrm{w_i(BIC)}$) were calculated in R according to the following equation:

$$w_i(\mathrm{BIC}) = \frac{\exp\left\{-\frac{1}{2}\Delta_i(\mathrm{BIC})\right\}}{\sum_{k=1}^{K}\exp\left\{-\frac{1}{2}\Delta_k(\mathrm{BIC})\right\}}$$

Where *i* is the model of interest, *k* is the sum of all models, $\Delta(\mathrm{BIC})$ is the BIC difference

between the model and the best candidate model and $\sum w_i(\mathrm{BIC}) = 1$. The Schwarz weight, therefore, gives the probability that model *i* is the best candidate model out of all the tested models. The evidence in favour of one model over the other is given by:

$$\frac{w_i(\mathrm{BIC})}{w_j(\mathrm{BIC})}$$

3-ITEM SIMON TASK

Model selection techniques revealed that the model that best described the experimental data (i.e. the model with the lowest BIC) for the 3-item Simon task was one that allowed the threshold *b, A* and the non-decision time parameter *T0* to vary across adaptation levels, while keeping all drift rate parameters fixed. Based on Schwarz weights comparison, this model is over 10^5 times more likely than the next best model (see Table 3). The next best three models did not majorly differ from the best model in terms of how parameters were allowed to vary.

Model	Free parameters					BIC value
Best fit	*A*	*T0*	*b*	-	-	26093
2nd best fit	A_{curr}	-	-	-	*v*	26116
3rd best fit	A_{prev}	*T0*	*b*	-	-	26123
4th best fit	-	*T0*	-	-	-	26125
Most complex	*A*	*T0*	*b*	*s*	*v*	26621

Table 3. *3-item Simon task. BIC values comparison for the best fitting model, the next three best fitting models and the most complex model.*

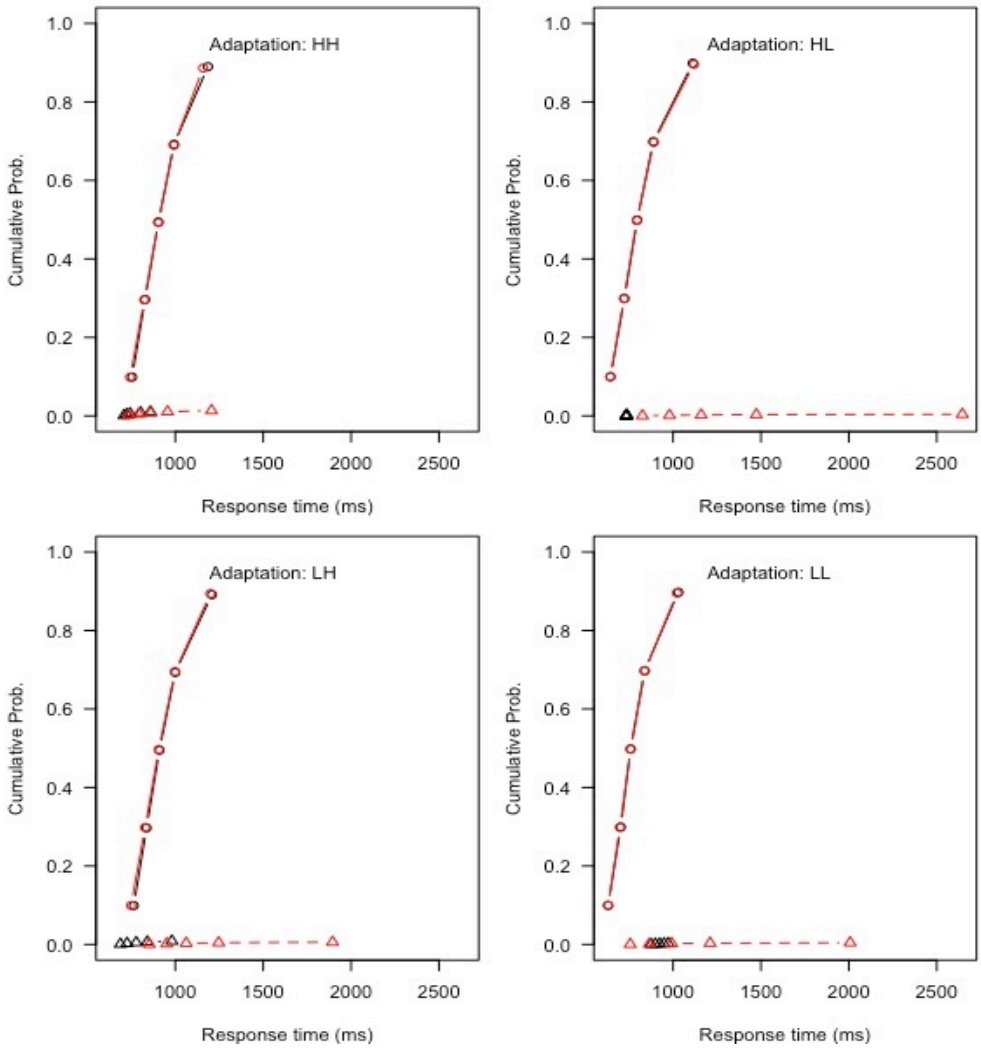


Figure 29. *Defective cumulative distribution plots for the best model fit to the 3-item Simon task data. Response time quantiles for each adaptation condition (HH, LH, HL and LL) were estimated from data (black lines) and predicted by the best-fitting model (red lines) for correct (circles) and incorrect (triangles) responses, and plotted against response probability. The right-most circle (or triangle) represents the quantile below which 90% of the data falls, the next one represents the 70% quantile, the next one represents the 50% quantile and so on.*

Cumulative response time distributions were calculated for the experimental data and for the data predicted by the best model. Figure 29 graphically depicts the fit of the model to the experimental data. As can be seen, the best model very accurately predicts experimental data for correct responses across all levels of adaptation. The

predictions for incorrect responses, however, are not as accurate, particularly for low-conflict trials, where the model greatly overestimates most of the quantiles. This is likely due to very low error rate in the task overall (0.8%) and for low-conflict trials in particular.

The parameter values for the best fitting model (Table 4) were fairly consistent with our predictions. We observed higher response thresholds *b* for HL trials (838.22) than any other trial type. The response threshold was also higher for LH trials (779.51) than HH trials (665.68). These values were complemented by higher start-point distribution values *A* for high-conflict trials (average = 638.75) than low-conflict trials (average = 478.53). Non-decision time *T0* was also generally higher for high-conflict (average = 659.01) than low-conflict (average = 456.21) trials.

Trial type	A	T0	b	v
HH	662.73	701	665.68	1.67
LH	614.78	616	779.51	1.67
HL	494.91	431	838.22	1.67
LL	462.16	480	686.95	1.67

Table 4. *Parameter values of the best-fitting model for the 3-item Simon Task. A is the top of the start-point distribution, T0 represents the non-decision time, b is the response threshold and v is the mean drift rate value for the correct-choice accumulator.*

Given the differences in *A* and *b* values across the adaptation conditions, we also observed a difference in response caution (Figure 30), which was calculated by subtracting the *A* values from the corresponding mean values of *b*. As predicted, HL

trials showed the highest response caution out of all trial types (343.30). Higher response caution values were also observed for LH trials (164.71) than HH trials (2.94).

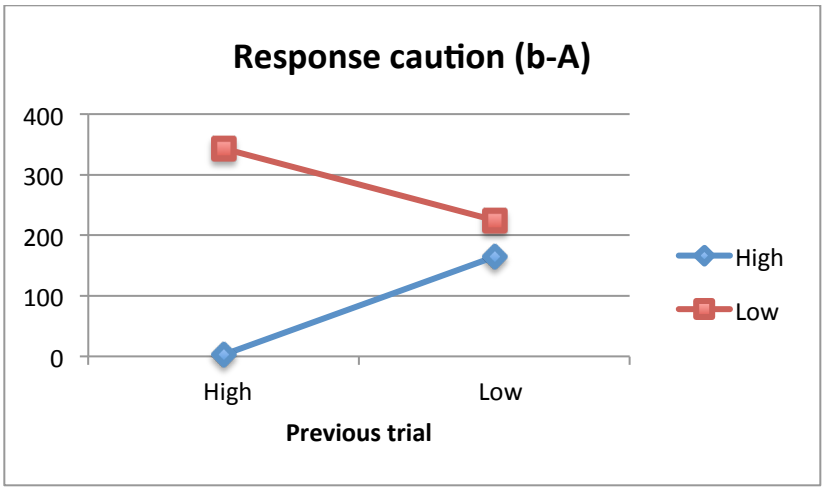


Figure 30. *3-item Simon task. Mean response caution values (calculated as the difference between* b *and* A *for high-conflict and low-conflict trials as a function of preceding trial type (high- or low-conflict).*

In Chapter 3, we showed that the 3-item Simon task showed adaptation effects due to both slower HL trials than LL trials and faster HH trials than LH trials. Given the results of the model fitting, this seems to be reflected in measures of response-caution, with higher response caution for slower trials. In the same chapter, we also showed that the faster HH trials are only observed in the 3-item version of the Simon task, but not in the classic, 2-item version of the task. We investigated whether this difference would also be reflected in the model's parameter values.

2-ITEM SIMON TASK

According to the model selection procedure previously outlined, the best model for the 2-item Simon task was one that allowed the mean of the drift rate distribution *v*, the threshold *b* and the non-decision time parameter *T0* to vary across adaptation levels and *A* to vary only across the current level of conflict (high or low conflict). Based on Schwarz weights comparison, this model is over 500 times more likely than the next best model to be the true model (see Table 5). The next best three models did not majorly differ from the best model in terms of how parameters were allowed to vary, except for the mean of the drift rate distribution, which was equally as likely to vary as to be fixed.

Model	Free parameters					BIC value
Best fit	A_{curr}	*T0*	*b*	-	*v*	19095
2nd best fit	A_{curr}	*T0*	*b*	-	-	19107
3rd best fit	A_{prev}	$T0_{curr}$	-	-	*v*	19110
4th best fit	*A*	*T0*	*b*	-	-	19111
Most complex	*A*	*T0*	*b*	*s*	*v*	19103

Table 5. *2-item Simon task: BIC values comparison for the best fitting model, the next three best fitting models and the most complex model.*

Figure 31 shows the fit between the cumulative response time distributions for the experimental data and for the data predicted by the best model. Similarly to the 3-item version, the best-fit model for the 2-item version predicts correct responses distributions very accurately, but fails to predict incorrect responses distributions as

precisely. As in the case of the previous task, the quantiles for incorrect responses are overestimated by the model. This is also likely due to very low error rates (0.7%).

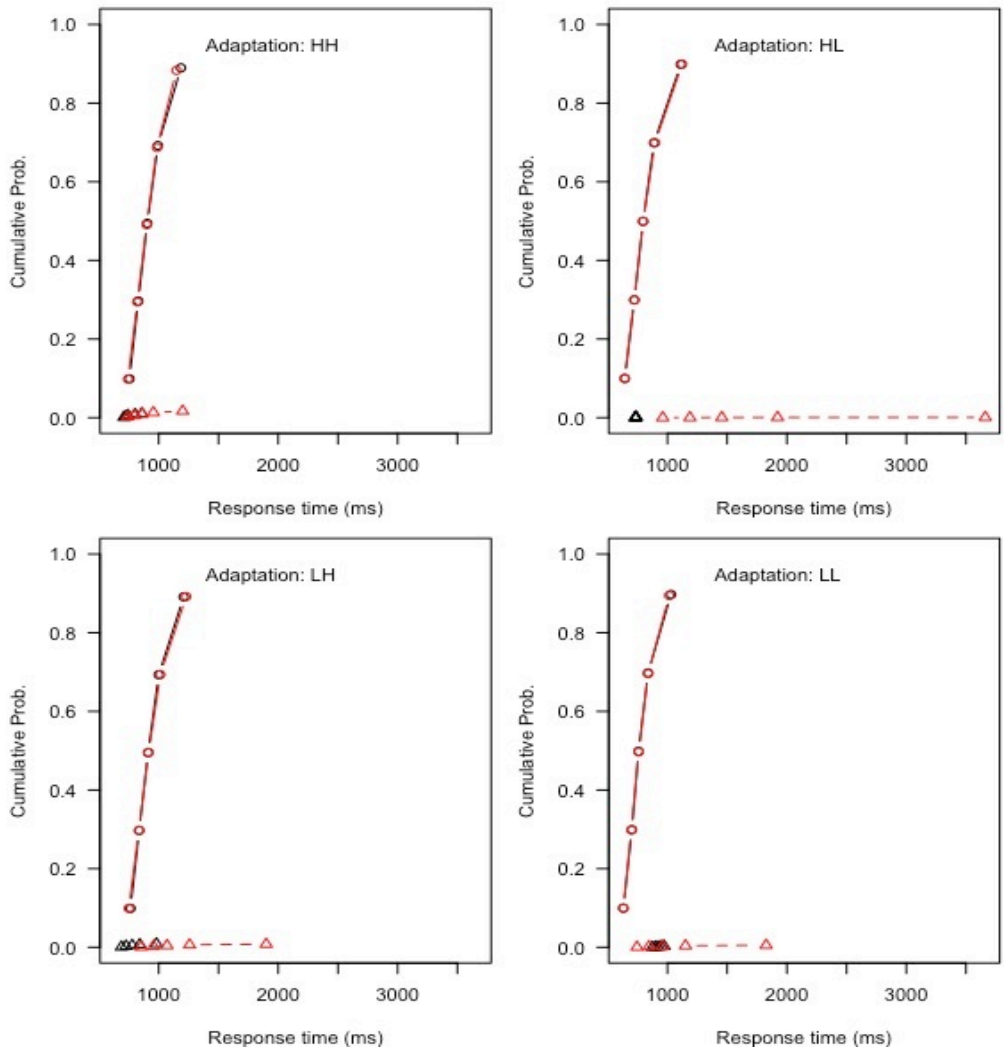


Figure 31. ***Defective cumulative distribution plots for the best model fit to the 2-item Simon task data.*** *Response time quantiles for each adaptation condition (HH, LH, HL and LL) were estimated from data (black lines) and predicted by the best-fitting model (red lines) for correct (circles) and incorrect (triangles) responses, and plotted against response probability.*

The finding that the best 4 models were equally likely to keep *v* fixed or varied can be explained by the fact that, while there are differences in mean drift rate values across the four conditions (see Table 6), the largest differences were actually observed

when comparing high-conflict to low-conflict trials, rather than specific adaptation levels, with higher drift rate values for low-conflict (average drift rate = 2.13) than high-conflict (average drift rate = 1.69) trials. As in the case of the 3-item version, the highest threshold *b* values were observed for HL trials (1095.70) followed by LL trials (959.72). The difference in response threshold between HH and LH trials (18.81) was several times smaller than the one between HL and LL trials (135.98). We also observed higher *A* values for low-conflict (663.82) than high-conflict (564.80) trials. Non-decision time *T0* was higher for high-conflict (average = 643.67) than low-conflict (average = 419.16) trials, with very little discrepancy between HH and LH (difference = 2.85), and slightly higher values for LL than HL.

Trial type	A	T0	b	v
HH	564.80	645	601.88	1.73
LH	564.80	642	620.69	1.66
HL	663.82	403	1095.70	2.06
LL	663.82	435	959.72	2.21

Table 6. *Parameter values for the best-fitting model for the 2-item Simon Task. A is the top of the start-point distribution, T0 represents the non-decision time, b is the response threshold and v is the mean drift rate value for the correct-choice accumulator.*

As in the case of the 3-item version, and once again consistently with predictions, response caution (Figure 32) was much higher for HL trials (431.88) than LL trials (295.89). Response caution values between HH (37.08) and LH trials (55.89) were similar. This is contrasted with the much larger difference in response caution between HH and LH trials in the 3-item version of the task (difference = 161.77).

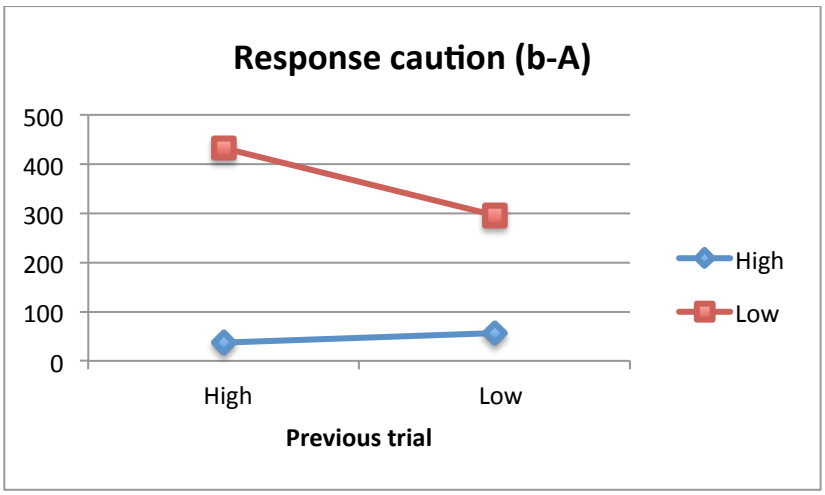


Figure 32. *2-item Simon task. Mean response caution values (calculated as the difference between* b *and* A*) for high-conflict and low-conflict trials, as a function of preceding trial type (high- or low-conflict).*

We can therefore conclude that the presence of faster HH trials, compared to LH trials, that we observed in the 3-item version of the Simon task is likely due to differences in response caution. More specifically, while response-caution is similar for both types of high-conflict trials in the 2-item version, it is much larger for LH trials in the 3-item version. The effect is not due to differences in drift rate values (as these only reflect differences between high- and low-conflict trials), nor other parameters such as *T0* or *A*, which also appears to only reflect differences between conflict, rather than adaptation, levels.

We therefore next investigated whether the phenomenon of faster HH than LH trials in the WCST is also dependent on measures of response-caution or on other parameters.

WCST

The best model to describe the data obtained from the WCST analogue was one that allowed the threshold *b,* and the drift rate parameters (*s* and *v*) to vary across the four adaptation levels, and *A* to vary across the current conflict level (high- or low-conflict), while keeping non-decision time *T0* fixed. Based on Schwarz weights comparison, this model was over 600 times more likely to provide the best fit than the next best model. The next best three models did not majorly differ from the best model in terms of how parameters were allowed to vary (see Table 7).

Model	Free parameters					BIC value
Best fit	A_{curr}	-	*b*	*s*	*v*	20224
2nd best fit	*A*	-	*b*	*s*	*v*	20241
3rd best fit	A_{prev}	-	*b*	*s*	*v*	20244
4th best fit	*A*	-	-	*s*	*v*	20280
Most complex	*A*	*T0*	*b*	*s*	*v*	20426

Table 7. *WCST: BIC values comparison for the best fitting model, the next three best fitting models and the most complex model.*

The parameter values were again used to predict cumulative response time distributions in order evaluate the model fit (Figure 33). The best model very accurately predicts experimental data for correct responses across all levels of adaptation. The predictions for incorrect responses, however, while more accurate than in the case of the Simon tasks, are not as accurate as the predicted values for correct responses, particularly for low-conflict trials, where the model greatly overestimates all quantiles. This is due to the absence of errors for low-conflict trials in the overall task.

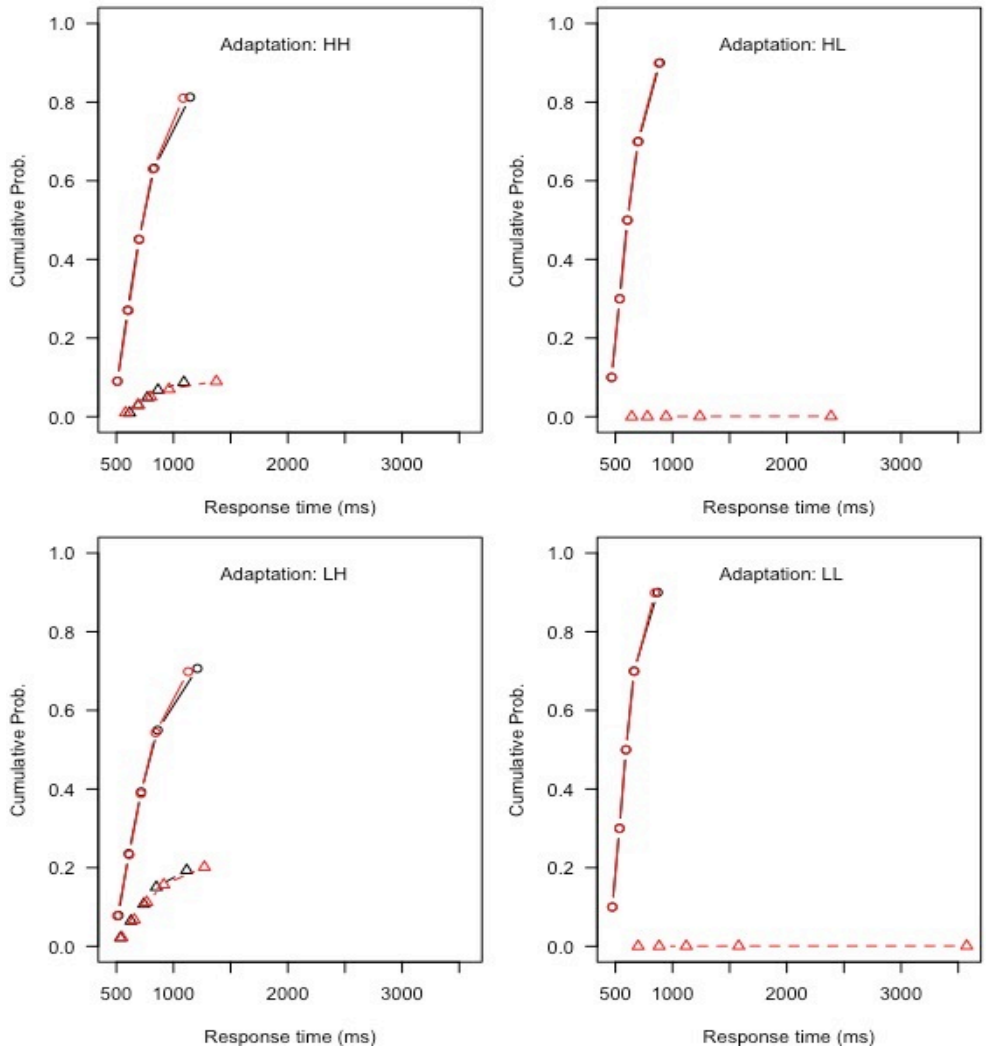


Figure 33. *Defective cumulative distribution plots for the best model fit to the WCST data. Response time quantiles for each adaptation condition (HH, LH, HL and LL) were estimated from data (black lines) and predicted by the best-fitting model (red lines) for correct (circles) and incorrect (triangles) responses, and plotted against response probability.*

The parameter values for the best fitting model are reported in Table 8. Mean drift rate values *v* were larger for low-conflict (average value = 2.49) than high-conflict trials (average value = 0.75). The differences in drift rate values between adaptation levels were small (Figure 34), with HH trials having slightly larger drift rates than LH trials

(0.83 compared to 0.68), and LL trials having larger drift rates than HL trials (2.89 compared to 2.09).

Trial type	A	T0	b	v
HH	383.42	291	551.52	0.83
LH	383.42	291	523.63	0.68
HL	666.07	291	996.78	2.09
LL	666.07	291	1199.4	2.89

Table 8. *Parameter values for best-fitting model for the WCST analogue. A is the top of the start-point distribution, T0 represents the non-decision time, b is the response threshold and v is the mean drift rate value for the correct-choice accumulator.*

Threshold values *b* were higher for low-conflict (average threshold value = 537.57) than high-conflict trials (average threshold value = 537.57), and particularly high for LL trials (1199.4). Values of *A* only varied between high- and low-conflict trials, with higher values for the top of the start-point distribution for low-conflict (666.07) than high-conflict (383.42) trials.

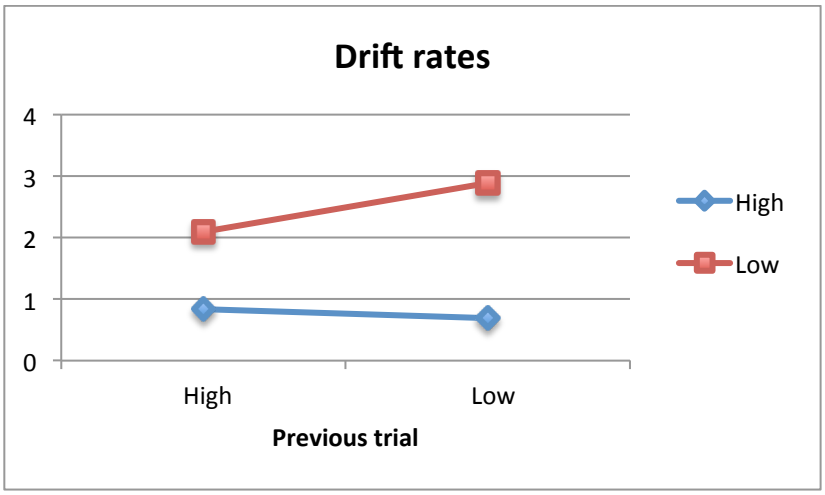


Figure 34. ***WCST.*** *Mean drift-rate distribution values for high-conflict and low-conflict trials as a function of preceding trial type (high- or low-conflict).*

Response caution measures showed a similar pattern to drift rate values (see Figure 35), with higher values for low-conflict (average = 432.02) than high-conflict (average = 154.16) trials, and was particularly high for LL trials (533.33). The difference in response caution between HH and LH was several times smaller (27.89) than the difference between LL and HL (202.62).

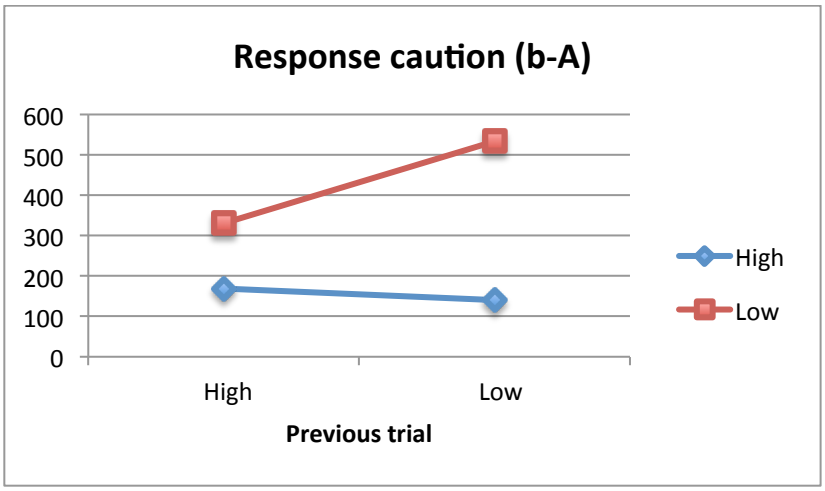


Figure 35. ***WCST.*** *Mean response caution values (calculated as the difference between* b *and* A *for high-conflict and low-conflict trials as a function of preceding trial type (high- or low-conflict).*

Discussion

Our predictions for the Simon tasks were largely confirmed by the model. Adaptation effects in the 3-item version of the task were due exclusively to variations in response-caution parameters, and not to the speed of evidence accumulation, which was fixed for all trial types. In particular, the slower HL trials were due to higher response caution compared to LL trials. This could indicate that participants are

suppressing the habitual response on high-conflict trials, and suppression carries over across trials, reducing facilitation due to habit-exploitation on low-conflict trials. Indeed, the 3-item version of the Simon task also presented higher response caution values for LH trials than HH trials. We originally speculated that this might be due to the fact that a larger stimulus-set might require the formation and maintenance of numerous short-term memory links for the associations between specific stimuli and responses, and that this might lead to more efficient performance on HH trials, once the indirect route, supporting rule-guided behaviour, is more active. However, the LBA fittings reveal that, at least for this version of the task, differences in response time are due exclusively to higher response caution for LH. This was not observed in the 2-item version of the task, but is consistent with the finding that no significant differences in response time between HH and LH trials are reported for the 2-item version, whereas they are observed in the 3-item version (see Chapter 3).

The finding of higher caution values on HL and LH trials is consistent with the idea that a change in conflict levels would lead suppression of the most advantageous route to response (in this case the direct route) and therefore to an increase in caution. Higher hemodynamic activity observed in the ACC and dlPFC on LH trials (Durston, 2003; Kerns, 2006; Kerns et al., 2004) could therefore be an index of an increase in response caution, rather than an index of an increase in the engagement of attentional mechanisms or extra cognitive effort. Indeed, van Maanen and colleagues (2011) found that fluctuations in response caution were correlated with fluctuations in ACC

activation. More specifically, they found that it was trials requiring a *change* between emphasis conditions (speed or accuracy), which therefore triggered a *change* in response caution, that increased activation in ACC, rather than ACC activity being higher when response caution is high (as one might expect to be the case for high-conflict trials). This is consistent with findings indicating higher ACC activation on trials involving a *change* between conflict levels in recent history (i.e. on LH trials, or on high-conflict trials following a sequence of low-conflict trials; see Durston et al. (2003)), rather than when conflict is consistently high (i.e. on HH trials), which may require changes in strategies, such as suppression of the direct route used for response, and that this is reflected in changes in response caution.

If adaptation in the Simon task were due to increased engagement of attentional mechanisms or enhancement of task-relevant information once conflict is experienced, we would expect this to be reflected in higher drift rates following high-conflict trials, as both low- and high-conflict trials would benefit from more efficient processing of task-relevant information, leading to faster evidence accumulation. The best model for the 2-item version of the task did include varying drift rate parameters, although they varied between conflict-levels on current trials, rather than across adaptation levels, and with relatively small differences. This could be due to the fact that adding one item to the stimulus set in this task might make location less informative of the correct response, therefore requiring more time for evidence accumulation, as well as reducing the possible contributions of positive priming, which could speed up evidence accumulation.

Even though faster HH trials are found in both the 3-item version of the Simon task and the WCST, the LBA fits revealed that this might be due to different underlying mechanisms, as response caution in the WCST does not vary for HH and LH trials. One possible explanation is that response caution exclusively reflects the suppression of the direct route. LH trials in the Simon tasks require both the need to suppress the direct route as well as to engage the indirect route. LH trials in the WCST, instead, may only require re-engagement of the indirect route.

We initially predicted that adaptation in the WCST would be reflected in higher drift rate values for HH trials than LH trials, due to biasing of the indirect route, which may result in more efficient evidence accumulation. However, just as we did not observe differences in response caution, we did not observe very large differences in drift rates for HH and LH trials. HH trials did have a slightly higher drift rate, but the difference was minimal compared to the broader high-conflict/low-conflict distinction. The absence of any evident parameter variation between HH and LH trials is perhaps surprising, given that response time difference are robust and consistently replicated (see Chapters 1, 2, 3, as well as Mansouri et al., 2007). One possible explanation is that the effect is 'smeared' across too many parameters to be evident, as all the best models for this task involved a large number of free parameters. It might also be the case that the best model according to our BIC estimates is an overfitted model, which captures noise related to individual variability, and that this could be obscuring changes in

parameter values due to experimental manipulations. This is an issue that is normally overcome by running individual model-fits, which could indicate that the best model varies substantially across participants. However, due to the low number of observations per adaptation levels in our tasks, individual fits would nonetheless be too imprecise to prove conclusive on this issue.

A further possible explanation concerns the fact that the WCST involves an element of visual search. Even assuming that the indirect route is biased after a high-conflict trial, with task-relevant information and rules having been recently re-engaged, the participant still needs to locate the target stimulus on the current trial. Search models (e.g. Smith & Ratcliff, 2009) suggest that, in a spatial cueing task, search for the target usually begins from the cued location, and responses are faster when the target is presented in the cued location than if it is presented in an uncued location. The position of the target stimulus on the previous trial of the WCST might act as a cue for the search to start from, and therefore response time might take longer on trials where the target is in another location, and be faster when the target appears in the same location as the previous target. Prinzmetal and colleagues (Prinzmetal, Taylor, Myers, & Nguyen-Espino, 2011) found that when distractors are present in a spatial cueing task, the search model makes more accurate prediction regarding the pattern of response times observed in this task (i.e. cueing effects increase as a function of the number of possible locations) than evidence accumulation models such as the LBA (which predict that cueing effects would diminish with a larger number of locations because the amount of competition

between different accumulators will be diluted, therefore reducing the cueing effect). As adaptation trials in the WCST involve an element of visual search with distractors (with possible cueing effects from the previous trials, which might be independent of conflict or adaptation levels), it is possible that the dynamics of these processes, which are better described by search models, present a confound that masks the type of conflict-related effects that are better described by decision models such as the LBA. Some studies have indeed suggested that conflict resolution might be modulated by the allocation of attentional resources to different spatial locations (e.g. Chen, Wei & Zhou, 2006).

It is however unlikely that the adaptation effects observed in the WCST are exclusively due to cueing effects, as increases in speed of response are restricted to HH trials exclusively, whereas if cueing were the mechanism responsible for faster responses we would expect trials to be affected regardless of conflict on the current or previous trial. Furthermore, the use of three locations for response greatly diminishes the probability of pairs of trials presenting the target item in the same location. Any cueing or visual search effects are likely to be modulated by conflict levels. Future research could attempt to tease apart the contribution of spatial attention allocation to conflict resolution, for example by modifying the current WCST analogue to present only one test item at a time and require the participant to perform a simple yes/no discrimination as to whether the test item matches a previously presented cue item on the currently relevant dimension (shape or colour).

Due to the, as of yet, limited application of the LBA model to conflict data, it is difficult to draw inferences from our findings, particularly developing an explanation for our inconclusive WCST fits, by comparing our results to the literature. However, despite the limited evidence available to date, we do find some consistency with previous studies. First, the parameter values that best describe the data from our Simon tasks are fairly consistent with those by King and colleagues (2012). They too found that, in a Flanker task - which is also a task we found displaying adaptation effects due to decrease in speed on HL trials (see Chapter 3) - manipulating conflict frequency affected response caution, rather than drift rates. More specifically, they found that increasing conflict frequency resulted in increased response threshold for low-conflict trials. Increasing conflict frequency would, according to our hypothesis, result in the indirect route (i.e. the route supporting rule-guided behaviour) becoming the default route, as this would be more beneficial to performance in the overall task. This would cause low-conflict trials, which are normally facilitated by the use of the direct route, to be affected, showing a decrease in speed. Indeed, increased conflict frequency often results in a reversal of the conflict effect (Hommel, 1994; Logan, 1979; Nieuwenhuis et al., 2003; Sturmer et al., 2002), with slower responses to low-conflict trials than high-conflict trials. The mechanisms behind the reduction in speed on low-conflict trials, when conflict frequency is high, are comparable to the ones reducing speed on HL trials: the direct route is suppressed on the previous H trials, resulting in slower responses.

Such reductions in facilitation, in both our Simon tasks and the Flanker task used by King and colleagues (2012), are reflected in higher response caution.

The main limitations with the application of the LBA to all of our conflict tasks concerned the low error rates, particularly for low-conflict trials, and the relatively low number of observation for each adaptation condition. Both of these issues can lead to overestimation of the drift rate (see Donkin et al., 2009; 2011). Some options to overcome the limitation of low error rates low have been suggested, such as estimating the drift rate based on the proportion of errors rather than their response time (see Ludwig et al., 2009). This option is worth considering for future applications of the model. Fitting the model to vincentized data allowed us to overcome the issue of the low number of observations. However, as previously mentioned, fitting the model to the vincentized data might lead to the inclusion of noise due to individual variability that could result in the best model including variations in parameters that might not be related to the experimental manipulation per-se. For example, one observation that did not meet our predictions, and appears counterintuitive, is the presence of higher response caution for low-conflict than high-conflict trials overall. However, this could be due to an attempt to compensate the high non-decision time observed for high-conflict trials, which could be due to noise related to individual variability.

In conclusion, application of the LBA model to our conflict tasks has revealed that, in tasks where competition is created between a habit and instructed responses,

and a direct route can be exploited to facilitate performance on low-conflict trials, adaptation effects are dependent on adjustments in response caution. Higher response caution is observed on trials that would otherwise benefit from facilitation, when the direct route is suppressed, and this results in slower response times. Adaptation in these tasks is not due to more efficient evidence accumulation following conflict. In tasks where no such facilitation is possible, and the direct route is not exploited, the LBA model has so far proved inconclusive in explaining the potential mechanisms behind adaptation. While drift-rates do play a role in the WCST, as we predicted, there is no evident clear pattern in parameter values that can reflect those observed in response times. This, as discussed, could be due to a number of limitations both with the task and model fitting. Application of the LBA to conflict tasks appears to be nonetheless a promising tool to tease apart and quantify components of response time data that could be responsible for adaptation effects, and how these might correlate with related neural activity.

CHAPTER 5

Effects of Deep Brain Stimulation on conflict-monitoring and conflict-related activity in Local Field Potentials recorded from the ACC

Introduction

As illustrated in previous chapters, the precise role of the ACC in conflict-monitoring is still subject of debate. In particular, the biggest challenge has been the reconciliation of imaging findings that strongly implicate the ACC in processing conflict (Barch et al., 2000; Carter et al., 2000; Badre et al., 2001; Fan et al., 2002; Durston et al., 2003; Hazeltine et al., 2003; Kerns, 2004; Kerns, 2006; van Veen & Carter, 2005; Luks et al., 2007; Nee et al., 2010; Kim, Kroger & Kim, 2011) with lesion (Cohen et al., 1999; Swick & Turken, 2002; DiPellegrino et al., 2007; Sheth et al., 2012; Vendrell et al., 1995; Stuss et al., 2001; Swick & Jovanovic, 2002; Fellows & Farah, 2005) and neurophysiological (Ito et al., 2003; Davis et al., 2005; Nakamura et al., 2005; Emeric et al., 2008; Sheth et al., 2012) findings that provide inconclusive results regarding the necessity of an intact ACC to respond and adapt to conflict, and regarding the presence of conflict-sensitive cells in this region. In this chapter, I will present the results of a pilot study where we attempted to use Deep Brain Stimulation (DBS) implants as a tool to further investigate these issues.

DBS implants represent a unique tool to investigate the role of the ACC in conflict monitoring and adaptation, by combining intracranial neurophysiological recording and brain stimulation within-subjects. They allow direct interference with neural activity at very specific sites, overcoming the limitations of human lesion studies that require reliance on measures of lesion-overlap to infer about the role of specific regions in cognition. They also overcome the limitations of extra-cranial stimulation methods (such as TMS) that are inadequate for reliably targeting deeper cortical regions (see Chapter 2). At the same time, they allow to record localized brain activity directly from the human cortex, overcoming the limited spatial resolution of scalp EEG and the limited temporal resolution of fMRI, as well as the issue of anatomical correspondence that may limit the interpretation of intracranial recordings in animals.

DBS was developed as a neurosurgical procedure to relieve the symptoms of a number of conditions, including Parkinson's (Bittar et al., 2005), depression (Mayberg et al., 2005) and chronic pain (Owen, Green, Stein, & Aziz, 2006). It involves the insertion of stimulating electrodes that function as a 'brain pacemaker', at brain sites that vary depending on the condition to be treated. For the treatment of chronic pain, the two brain areas that can be targeted are the thalamus (Marchand, Kupers, Bushnell, & Duncan, 2003; Owen et al., 2006) and the ACC (Boccard et al., 2013). These electrodes can either present a monopolar configuration (where an implanted pulse generator functions as the anode, while the electrode functions as the cathode) or a bipolar configuration (where one electrode contact is the anode and another contact is the

cathode), as illustrated in Figure 36. Monopolar configurations are generally more efficient as they reach a larger brain volume, whereas bipolar configurations are more localized but can minimize side effects (Bain, 2009). Therapeutic stimulation is usually applied at a frequency ranging between 130 and 185 Hz in the ACC and 5-80Hz in the thalamus.

While DBS has often proven effective in alleviating the negative symptoms of several conditions, the mechanisms by which DBS achieves this, and the precise effects of stimulation on the underlying neural tissue, are still unclear. Many factors interact in generating the clinical effects of DBS, including stimulation settings, cell properties and geometry of both the stimulus field and stimulated elements, but no framework has so far been provided to account for the effects of DBS in different brain regions (Kringelbach & Jenkinson, 2007; Perlmutter & Mink, 2006). The therapeutic effects of DBS often appear to be the same as the effects of surgical ablation of the target structure. For example, the patients in our study were treated to relieve chronic pain symptoms. Reduction in pain perception has been previously achieved by cingulotomy (Cohen, Kaplan, Moser, Jenkins, & Wilkinson, 1999). Effects of DBS and ablation in the treatment of Parkinson's, targeting the globus pallidus, are also similar (Liu & Postupna, 2008). On the other hand, changes in parameters such as the frequency of stimulation have been shown to produce very different effects in chronic pain patients: low frequency stimulation of the thalamus increases pain perception, whereas high-frequency stimulation of the same region alleviates the pain (Owen et al., 2006). This

suggests that DBS can sometimes elicit the same effects as a lesion, or the complete opposite effects, depending on stimulation parameters.

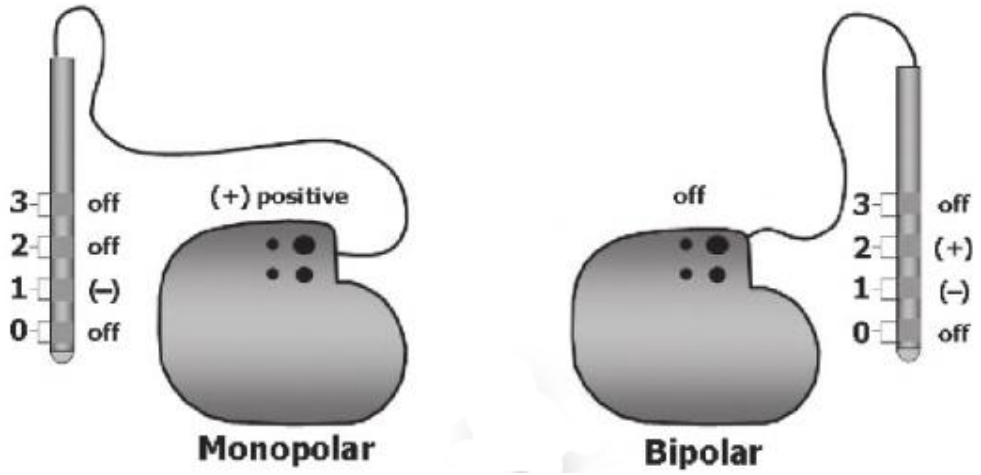


Figure 36. *Examples of a monopolar (left) and bipolar (right) DBS configuration. In a monopolar configuration current flows from the cathode (the active, negative contact) to the anode (the positively charged impulse generator). This allows stimulation to reach a larger volume of tissue. In a bipolar configuration, current flows from the cathode (the active, negative contact) to the anode (another active, positive contact). This restricts stimulation to a much more localized volume of tissue. Any of the contacts can be activated or deactivated, depending on stimulation parameters. (Figure adapted from Isaias and Tagliati (2008), pp. 367).*

Although evidence suggests that ACC ablation and DBS of the ACC produce similar therapeutic effects, no study, to the best of my knowledge, has investigated whether these two methods produce similar effects on behavioural performance. In the current study, we administered the conflict analogue of the WCST, already implemented by Mansouri and colleagues in the macaque (Mansouri et al., 2007) and in our previous human neuropsychological study (Chapter 1), as well as the 3-item version of the Simon task (Chapter 3), to a small cohort of cingulate DBS patients. We aimed to investigate whether direct stimulation of the ACC would have the same effects on the conflict WCST

analogue as ACC damage. Similarly to TMS, this experimental paradigm allowed for within-subjects comparisons, as participants could be tested with the stimulation turned on (therefore interfering with ACC activity) or turned off. This is important not only to establish the potential usefulness of this type of intracranial stimulation in investigating research questions, but also to test whether stimulation may impair normal cognitive functions in patients.

We also wished to compare the effects of DBS on a task that involves conflict between two instructed responses (WCST), with one that involves conflict between an habitual response and instructed responses (Simon), in order to investigate whether interference with ACC activity could affect responses to one type of conflict differently from the other. As illustrated in previous chapters (Chapters 3 and 4), the different type of conflict that is elicited in these two tasks might also lead to adaptation effects that are driven by different underlying mechanisms: slower responses on HL trials for the Simon task only (possibly due to habit/direct route suppression and increased caution) and faster responses on HH trials for both tasks (whose mechanisms are still unclear but appear unaffected by ACC damage – see Chapter 1). Therefore, interfering with ACC activity during these different types of adaptation could help elucidate whether this region is crucial to support one over the other.

We also recorded Local Field Potential (LFP) data during both tasks from the area targeted for stimulation. This allowed us to investigate the presence of conflict-related

activity in the neural population in the target area. There are many advantages to analysing intracranial LFP data to address these questions. First, as previously mentioned, recording activity directly from deep within the cortex, from a small and localized population of neurons, overcomes the problem of source localization and the propagation of signals through tissue and bone, which are limitations of scalp EEG or Electrocorticography recordings. At the same time, LFP data contains information about oscillatory components, representing the level of synchronization or desychronization of activity in the local neuronal population, which cannot be captured with single-cell recordings of spiking activity from only a few cells. This can be particularly valuable when investigating the underlying neural activity of a cognitive process, such as conflict-induced adaptation, which is likely to involve the integration of a substantial amount of information spanning a number of trials.

At present, evidence regarding the presence of conflict-related electrophysiological activity in the ACC seems to highlight discrepancies between the animal and the human literature. Intracranial recordings from the monkey ACC have so far failed to conclusively identify cells that are sensitive to conflict (Ito et al., 2003; Nakamura et al, 2005; Emeric et al., 2008). Some findings from human intracranial recordings, on the other hand, do show that activity appears modulated by the level of conflict in the current trial (Cohen, Ridderinkhof, Haupt, Elger, & Fell, 2008; Davis & Taylor, 2005; Sheth et al., 2012). However, the sensitivity of human ACC neurons has also been shown to generalize beyond conflict to more general aspects of cognition,

such as task-difficulty, familiarity, and errors (C. Wang, Ulbert, Schomer, Marinkovic, & Halgren, 2005).

In a recent study, Cohen and colleagues (2008) recorded intracranial ERPs in two patients carrying out a Flanker task. Electrode implants in both of the patients covered a very large proportion of the mPFC, including areas 24 and 32 of the ACC. They found conflict-related activity both after the cue (beta suppression at dorso-caudal electrodes) and after the response (alpha/theta enhancement at rostro-ventral electrodes), but, as neither patient showed adaptation effects, they did not investigate whether conflict-related activity was modulated by conflict-levels on previous trials. Sheth and colleagues (2012) recorded 59 isolated neurons in dorsal ACC while patients carried out a multiple interference task combining elements of the Flanker, Simon and Stroop tasks. They identified 24 conflict-sensitive neurons that fired preferentially after the cue and 23 conflict-sensitive neurons that fired preferentially after the response. The activity of the cells that fired preferentially to the cue was modulated both by the level of conflict in the current trial and by the level of conflict in the preceding trial. After surgical ablation of this region, only adaptation was impaired, consistent with the hypothesis that while ACC might be processing conflict information, it is not necessary for responses on the current trial. Both studies seem therefore to indicate that the level of conflict in a trial modulates localized cortical activity in human ACC.

The aforementioned studies used adaptation of classic conflict tasks, which create competition between a habitual response and an instructed response. In our study, we aimed to investigate the neural correlates of conflict in a task that involves competition between two instructed responses (the WCST) as well as a task also involving competition between a habit and instructed responses (the Simon task). Since, as previously shown, these two tasks appear to involve different adaptation mechanisms, we were also interested in investigating whether LFP activity would differ between adaptation trials depending on whether they involve the suppression of a habitual response. We predicted that, if the role of the ACC is to detect conflict that is elicited between a habit and an instructed response, we should observe conflict-related activity on high-conflict trials in the Simon task, particularly when there is a change in conflict levels (i.e. when conflict is first detected) that requires the suppression of the habitual response/direct route. Therefore ACC activity should be modulated by conflict particularly on LH trials in the Simon task. As evidence has shown the ACC not to be crucial for responding to high-conflict trials or for trial-to-trial adaptation in the WCST (Mansouri et al., 2007; see also Chapter 1), where conflict is elicited between two instructed responses, we did not expect ACC activity in this task to be modulated by conflict levels.

A second possible hypothesis, concerning the specific ACC site we recorded from, would predict the inverse pattern of results, with the presence of conflict-related activity for the WCST, but not for the Simon task. The electrode implants in our group of

DBS patients, which were inserted for clinical purposes only, aimed to target caudal area 24 of the ACC. According to the proposed rostro-caudal gradient subdivision of the ACC (see Bush, Luu, & Posner, 2000; Milham & Barich, 2005; van Veen & Carter, 2005; Badzakova et al., 2008; Kim, Kroger & Kim, 2011), rostral parts of area 24 should be more sensitive to response-based or emotionally-salient conflict and caudal parts of area 24 should be more sensitive to perceptual/stimulus-based or cognitive conflict. However, another substantial body of literature seems to indicate that such a fine-grained subdivision of ACC is not present, with both rostral and caudal ACC being involved in processing conflict information regardless of the specific type of conflict (Fan, 2003; Weissman et al., 2003; Wendt et al., 2007). None of our tasks involve emotionally salient elements, but the WCST should elicit perceptual/stimulus-based conflict, while the Simon task should elicit response-based conflict. Lie and colleagues (2006) investigated the neural processes underlying the classic WCST using fMRI. They decomposed the task into 4 stages, adding degrees of complexity with each stage. Trials in their baseline stage essentially constituted the equivalent of our low-conflict trials, whereas trials in their most complex stage were equivalent to our high-conflict trials. They found that activation in caudal ACC increased with increasing complexity (therefore from low- to high-conflict) while decreasing in rostral ACC. If we consider caudal ACC as part of a network processing perceptual conflict and rostral ACC as part of a network processing response conflict, the patterns of activation observed by Lie and colleagues suggest that our high-conflict trials evoke perceptual, rather than response, conflict. Therefore, if caudal area 24 is sensitive to perceptual conflict, we predict to

see conflict-related activation only for the WCST, and not for the Simon task. If such a functional subdivision is not present in the ACC, then caudal area 24 should show sensitivity to conflict in both tasks (or only to Simon-related conflict, according to our first hypothesis).

Methods

Participants

5 participants (all male, mean age 55.2) took part in the study. Participants were recruited through the Neurosurgery Department at the John Radcliffe Hospital in Oxford. Participants had undergone surgery for the insertion of Medtronics 3387® or St. Jude Medical 6143® electrodes connected to an implantable pulse generator (Figure 37) with the aim to target area 24.

ID	Age	Gender	Electrode configuration	Electrode placement	Stimulation frequency	Voltage	Etiology
JW	49	M	Bipolar	Bilateral	130 Hz	4.0v	Trauma
PP	72	M	Bipolar	Left ACC	120Hz	3.0v	Stroke
AP	50	M	Bipolar	Left ACC	130 Hz	6.0v	Stroke
WS	42	M	Bipolar	Bilateral	130Hz	9.0v	Trauma
MG	63	M	Monopolar	Bilateral	120Hz	6.9v	Stroke

Table 9. *Participants' demographic information, electrode placement, stimulation settings and etiology of chronic pain symptoms.*

Three participants were fitted with bilateral electrodes and two participants were fitted with unilateral electrodes in the left ACC. In four participants (JW, PP, AP and WS), electrodes presented a bipolar configuration. In one participant (MG) they presented a monopolar configuration. For details about their background information, including etiology of the pain of the patients refer to Table 9. All participants were native English speakers and had normal or corrected-to-normal vision.



Figure 37. *Sample of a standard Deep Brain Stimulation electrode implant compared to the size of a £1 coin: impulse generator (left) and electrode (right). The electrode presents 4 contacts (indicated by the red arrow). In monopolar configurations, the active contacts are referenced to the impulse generator, which serves as the anode. In bipolar configurations, the active contacts are referenced to one other contact.*

Electrode placement and LFP recording procedure

We had access to detailed information about electrode positioning for 3 of our 5 patients (AP, WS and MG). For each patient, electrode placement was confirmed following surgery via CT scan. For AP we had access to his post-operative CT scan. For WS and MG we had access to both the post-operative CT scans and the pre-operative MRI scans. The position of the electrodes within ACC is exemplified in **Figure 38**.

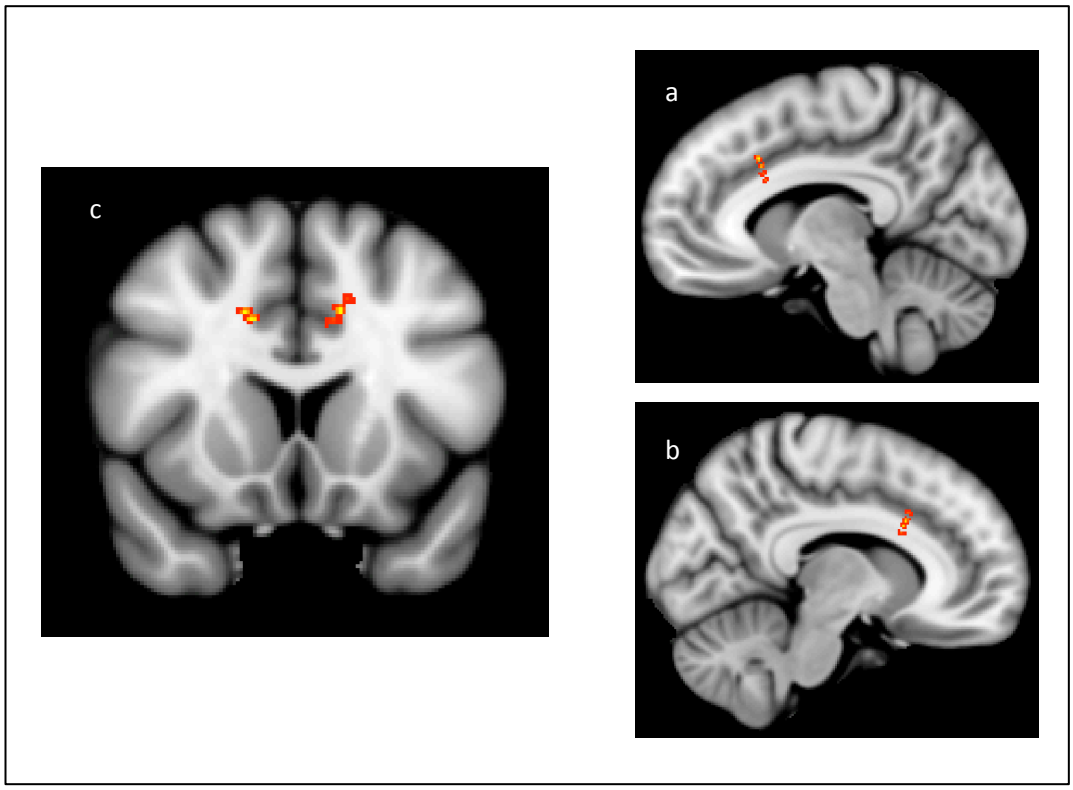


Figure 38. *Electrode placement illustrated on a standard MNI brain. Electrode placement in left ACC for patient WS (a), in right ACC for patient WS (b) and in the coronal view for patient MG (c).*

Each electrode consisted of four contacts. Using FLSview, an electrode mask identifying the location of each contact in the CT scan was created. The CT scan, electrode mask and the pre-operative MRI scan (where available) were then transposed

onto a standard brain and MNI coordinates for each contact were determined (see Table 10).

CONTACTS						
	LEFT ACC			RIGHT ACC		
PATIENT ID	X	Y	Z	X	Y	Z
AP	-11	14	37			
	-11	12	34			
	-10	10	31			
	-10	9	30			
WS	-8	18	35	11	17	34
	-8	17	32	10	16	31
	-7	16	29	10	15	28
	-7	15	26	9	14	25
MG	-14	16	41	18	17	40
	-12	15	37	16	15	37
	-10	14	34	14	15	35
	-8	13	32	12	13	32

Table 10. *MNI coordinates of electrode contacts in the ACC. One set of coordinates corresponds to one contact. Each adjacent pair of contacts represents one recording channel in a bipolar configuration.*

The implants were used to apply stimulation at the settings reported in Table 9 as well as to record Local Field Potential (LFP) signals. In three patients (JW, AP and WS), LFP recordings were carried out via three adjacent pairs of electrode contacts in a bipolar configuration, with each contact referenced to the adjacent contact. Signals were filtered at 0.5-500Hz and amplified (×10,000) using isolated CED 1902 amplifiers (Cambridge Electronic Design, Cambridge, UK), then sampled at 2500 Hz and recorded through CED 1401 MarkII A-D converter (Cambridge Electronic Design) onto a computer using a custom written program in Spike 2 (Cambridge Electronic Design). In one patient, MG, LFP data was recorded from a monopolar configuration, with each contact referenced to an electrode placed on the left mastoid process. Signals were amplified,

digitized and converted using Porti 32 channel (TMS International BV), sampled at a frequency of 2048Hz and converted into bipolar recordings.

Tasks and procedure

Four participants were administered the WCST conflict analogue (as described in Chapter 1) and the 3-item version of the Simon task (as described in Chapter 3). One participant (JW) was administered only the WCST conflict analogue. No modifications to the tasks as described in previous chapters were applied. For each task, participants carried out two blocks lasting approximately 8-10 minutes, with the stimulation turned on. Participants also carried out up to two blocks of each task with the stimulation turned off, depending on their level of comfort. The initial stimulation condition was randomized across participants. During breaks between blocks of the stimulation-off condition, participants were asked to report their level of pain on a scale from 0 to 10 (0 being none and 10 being maximum), in order to ensure the comfort of the patient throughout the session. Testing was suspended if self-reported pain values exceeded 5. Each block provided, on average, between 100 and 110 trials. Behavioural measures of response time and accuracy were recorded both during stimulation on and during stimulation off conditions. LFP data was recorded during stimulation off only.

Behavioural results

As for the other studies, the analysis of behavioural data in this pilot study was carried out on the Speed of Target Selection (STS) values, calculated by taking the reciprocal of the response time data. One patient, PP, was excluded from analysis due to difficulties understanding the requirements of the tasks, which did not provide an adequate number of correct trials, particularly in the high-conflict conditions (36 and 18 for the two conditions in the WCST and 5 trials in the Simon task). Patients JW, AP, WS and MG were included in the analysis of the WCST data. Patients WS and MG were included in the analysis of the Simon task data. Patient AP was excluded from the Simon task analysis, as he was unable to complete the Stimulation Off blocks of the Simon task due to discomfort. Patient JW was not administered the Simon task. Due to small number of participants contributing to the dataset at this stage, which lead to violation of the assumptions of parametric tests, statistical analysis, where possible, was carried out using non-parametric tests.

The number of errors was very low in both tasks. In the WCST, the average number of non-switch errors per session when stimulation was turned off was 15.75, and the average number of non-switch errors when the stimulation was turned on was 4.75. The Simon task presented an even lower number of errors. Patient WS only committed 8 errors in total, and MG made no errors. As such, no meaningful statistical analysis of accuracy data was possible. Nonetheless, as evidence indicates that ACC

lesions in the macaque increase the speed of response on incorrect trials (Buckley et al., 2009; Tanaka, Matsumoto, Mansouri, & Buckley, 2012), we looked at whether patients responded significantly faster on error trials when the stimulation was turned on, as opposed to off. We carried out a non-parametric Wilcoxon Signed Ranks t-test on the speed of response to error trials to compare the stimulation On and stimulation Off conditions. There was no significant difference in the speed of response between the two conditions ($z < 1$) but given the very small numbers of errors and the obvious confound with subjective levels of pain (see below) we do not believe this amounts to robust evidence for any species difference. The slightly lower number of errors when the stimulation was turned on might be due to a higher degree of comfort for the participants during the task, as stimulation substantially decreases their perception of pain.

WCST analogue

First, we looked at whether stimulation had an effect on the overall speed of response. Wilcoxon Signed Ranks paired-samples t-tests revealed no significant difference in speed of response ($z = -1.46$, $p = .250$) between the stimulation Off (mean = .95, sd = .15) and the stimulation On (mean = 1.02, sd = .05) conditions, indicating that stimulation did not affect overall speed of response. As all of our previous studies implementing the WCST were characterized by faster responses to low-conflict than to high-conflict trials, thus indicating the presence of a standard conflict effect, we carried

two 1-tailed Wilcoxon Signed Ranks t-tests in order to investigate whether there was a difference in speed of response to high- and low-conflict trials when the stimulation was turned on as opposed to off. STS (and response-time values, for reference) for each patient across conflict, adaptation and stimulation conditions are reported in Table 11.

ID	**On**		**Off**		**On**		**Off**	
	High	Low	High	Low	HH	LH	HH	LH
JW								
STS	0.893	1.052	0.742	0.899	0.969	0.877	0.886	0.700
RT(ms)	1120	951	1347	1112	1032	1140	1128	1428
AP								
STS	0.956	1.052	0.767	1.012	0.964	0.965	0.925	0.705
RT(ms)	1045	950	1304	987	1037	1036	1081	910
WS								
STS	1.087	1.107	1.144	1.197	1.139	1.055	1.192	1.106
RT(ms)	919	903	874	835	878	948	838	904
MG								
STS	0.973	1.077	0.916	0.965	1.000	0.986	0.959	0.914
RT(ms)	1028	928	1092	1036	910	1014	1041	1094

Table 11. *Individual Speed of Target Selection (STS), and corresponding Response Times (RTs) values for patient in each conflict condition (High and Low), adaptation condition (HH and LH) and stimulation condition (On and Off) for the WCST analogue.*

The analysis revealed a trend towards a significant conflict effect ($z = -1.83$, $p = .063$), for both stimulation conditions, with higher STS values for low-conflict (mean = 1.02, sd = .13, in the stimulation Off condition and mean = 1.07, sd = .03, in the stimulation On condition) compared to high-conflict trials (mean = .89, sd = .18, in the stimulation Off condition and mean = .98, sd =.03, in the stimulation Off condition), as illustrated in Figure 38. There was no significant difference in the size of the conflict

effect (i.e. the difference in speed between high- and low-conflict trials) between the two stimulation conditions ($z < 1$).

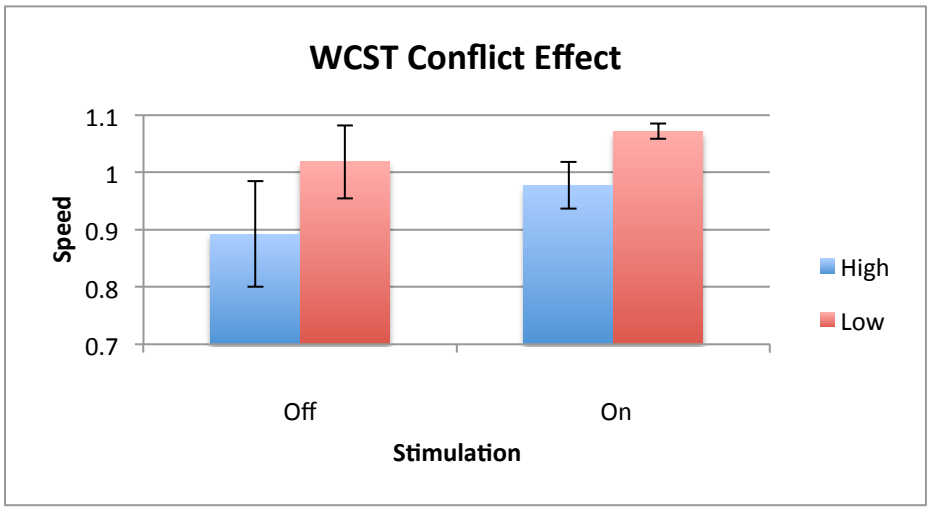


Figure 39. *Effects of stimulation on the speed of response during high- and low-conflict trials in the WCST analogue.*

As in the case of conflict, given that all our previous applications of the WCST have been characterized by adaptation effects due to faster speed of response on HH compared to LH trials, we carried out further 1-tailed Wilcoxon Signed Ranks paired-sample t-tests to investigate the effects of stimulation on the adaptation effect. Once again, there was a trend towards a significant difference between speed of response on HH and LH trials in the stimulation Off condition ($z = -1.83$, $p = .063$), with higher STS values on HH (mean = .99, sd = .14) than LH (mean = .86, sd = .19) trials. This difference was not significant when the stimulation was turned on ($z = -1.46$, $p = .25$). This effect seems to be characterized by a numerical increase in STS values for LH trials, more so

than for HH trials (see Figure 40). However, a further t-test revealed that this increase in speed of response for LH trials was non-significant (z = -1.45, p = .25).

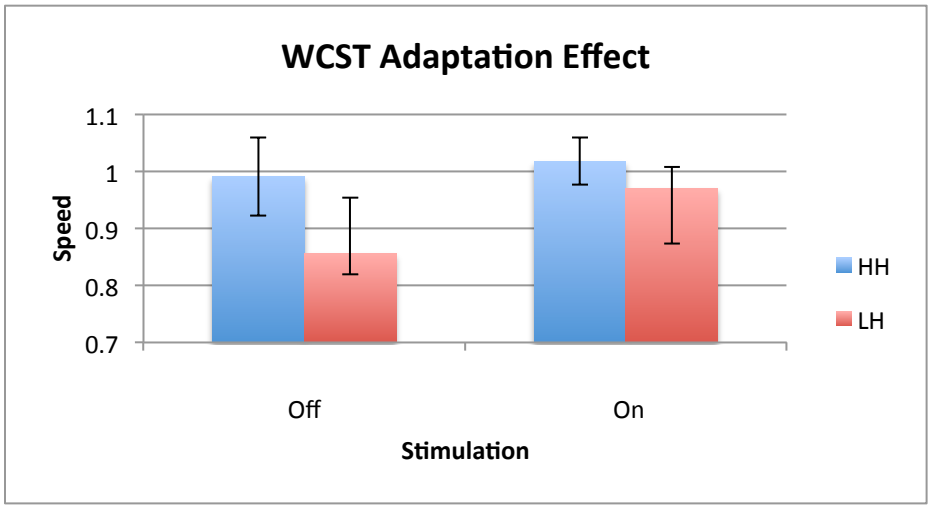


Figure 40. *Effects of stimulation on the speed of response during HH and LH trials in the WCST analogue.*

Simon Task

As only two patients successfully completed both stimulation On and Off conditions of the 3-item Simon task, statistical analysis on the behavioural data could not be carried out. Nonetheless, we report mean STS (and response-time values, for reference) for each patient, across conflict and stimulation conditions in Table 12.

Both patients showed a numerical difference between the speed of response to low-conflict and high-conflict trials, with higher STS values for low-conflict than high-conflict trials, indicating the presence of a classic conflict effect (mean difference

between high- and low-conflict trials = 0.114, in the stimulation Off condition). Stimulation did not appear to affect the size of the conflict effect (mean difference between high- and low-conflict trials = 0.107, in the stimulation On condition).

ID	Stimulation ON					
	High	Low	HH	LH	HL	LL
WS						
STS	0.840	0.915	0.837	0.850	0.918	0.937
RT	1190	1092	1198	1176	1089	1066
MG						
STS	0.852	0.992	0.831	0.873	0.980	1.013
RT	1173	1008	1204	1146	1020	987
	Stimulation OFF					
	High	Low	HH	LH	HL	LL
WS						
STS	0.863	0.948	0.854	0.872	0.941	0.949
RT	1158	1055	1171	1146	1063	1053
MG						
STS	0.872	1.017	0.879	0.869	0.996	1.030
RT	1146	983	1138	1151	1003	970

Table 12. *Individual Speed of Target Selection (STS), and corresponding Response Times (RTs) values for each patient in each conflict condition (High and Low), adaptation condition (HH, LH, HL and LL) and stimulation condition (On and Off) for the Simon task.*

Regarding adaptation effects, results appear inconclusive (refer to Table 12, columns four to six). In the stimulation Off condition, WS showed numerically higher STS values for LH trials than for HH trials, unlike the effects normally observed in this task (see Chapter 3). He also showed slightly lower STS values for HL trials compared to LL trials, as normally observed in this task. When stimulation was turned on, both HH and LH trials showed a decrease in speed of response, but the numerical difference between the two types of trials trials appeared unchanged, with faster speed on LH trials than HH

trials. Stimulation also appeared to affect the speed of response on HL and LL trials, with a larger decrease in STS values for HL trials than LL trials, therefore increasing the difference in speed between the two types of trials.

MG showed the opposite pattern of results. In the stimulation Off condition, he showed slightly higher STS values for HH trials than for LH trials and lower STS values speed for HL trials, compared to LL trials. Both effects are normally observed in this task. In the stimulation On condition, there was a substantial decrease in speed of reponse on HH trials, but not on LH trials. There was also a decrease in speed of response to both HL and LL trials, but the difference in speed between the two appeared unchanged.

Intracranial recordings analysis

Analysis of the intracranial data is illustrated in this section, for both our tasks. Data from three patients (AP, WS and MG)[1] was included in the WCST analysis and data from two patients (WS and MG) was included in the Simon task analysis. Because of variability in electrode positioning and referencing across patients, analysis of LFP data was carried out separately for each participant.

[1] Patient JW was excluded from this analysis. As he was our very first pilot patient, he was tested on an older version of the WCST, which did not contain trigger pulses that we implemented in newer version of the task in order to mark, within the recorded LFP data files, the different types of trials (conflict levels and correct/incorrect) and the exact time of presentation, trial onset and response.

The data was analysed offline using Spike2 software, custom-written scripts in Matlab (Version 8.0, Mathworks, USA) and the Field Trip analysis toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011). Data was pre-processed by applying a comb filter to remove the 50Hz mains noise and its harmonics. Error trials were removed and epochs centred around the event of interest were then extracted from the continuous recordings. Events of interest were defined as the time of conflict-onset (i.e. the time at which the sample item or fixation cross is touched by the participant, in order to initiate the trial, and test stimuli are presented on the screen), which will be referred to as stimulus-locked from here on, and as the time of response, which will be referred to as response-locked from here on. Trials were then visually inspected for artifacts. Trials containing technical artifacts, such as blocking or sudden jumps in voltage, were removed and only artifact-free trials were included in the analysis. The total number of trials included in the analysis was 198 WCST trials for AP, 308 WCST trials and 118 Simon trials for WS, 237 WCST trials and 218 Simon trials for MG.

Time-frequency decomposition was conducted in Field Trip using Hanning-windows of frequency-dependent length for frequencies below 30Hz, and a multitaper method with a series of Slepian tapers for frequencies above 30Hz. Raw power was converted into decibels (dB - by the following equation: log10(power/baseline)*10) in order to allow for easier comparison. The baseline was defined as the average frequency power in a time window -0.5 to -0.3 seconds before trial onset. We applied the same baseline to both stimulus-locked and response-locked data, in order to avoid

masking of pre-response activity due to inclusion in the baseline, similar to Cohen et al. (2008). The time-locked averaged intracranial Event Related Potentials (ERPs) for each condition were also computed. ERP were baseline corrected using -0.5 to -0.3s pre-event baseline, using separate baselines for stimulus-locked and response-locked data.

WCST analogue

Figure 41 illustrates the electrode average time-frequency decomposition and single-channel intracranial ERPs recorded in the left ACC for patients AP and WS, respectively, for the high- and low conflict conditions (for recordings from WS's right ACC electrode see Appendix 3, Figure A10).

The main consistent finding across the two participants was represented by an event-related potential around 200-100 ms prior the time of trial-onset and prior to response. This was accompanied by enhanced oscillatory activity at all lower frequencies and, at high-gamma (50-100Hz). The same pattern of activity was also observed in WS's right ACC electrode. The time at which this activity is present prior to trial onset corresponds approximately to the time the sample stimulus is presented on the screen, and time 0 corresponds to the time at which the participant touches the screen to initiate the trial, following which the test stimuli are presented on the screen.

As the type of activity at this time very closely resembles the activity present at 100-200ms before response is made, again, by touching the screen, it is unlikely that

this activity reflects ACC's response to the presentation of the visual stimulus, but is more likely to represent activity related to the initiation of voluntary movement. Both patients also presented an increase in power across all frequencies from around 200ms after trial onset for WS in his left ACC electrode, and around 400ms after trial onset for AP. A similar increase in power was also present in WS's right ACC electrode. As this resembled the motor-related activity observed prior to screen touch, it might represent activity related to the preparation of movement before response.

Patient WS did not show any discernable activity that differed between conflict conditions in either hemisphere. Patient AP, however, showed some beta band (20-30Hz) suppression starting from the time of trial onset and more sustained for high-conflict than for low-conflict trials. During the feedback period, patient AP showed enhanced beta activity (18-25Hz), which started immediately after response and was again sustained for a longer period after high-conflict trials compared to low-conflict trials.

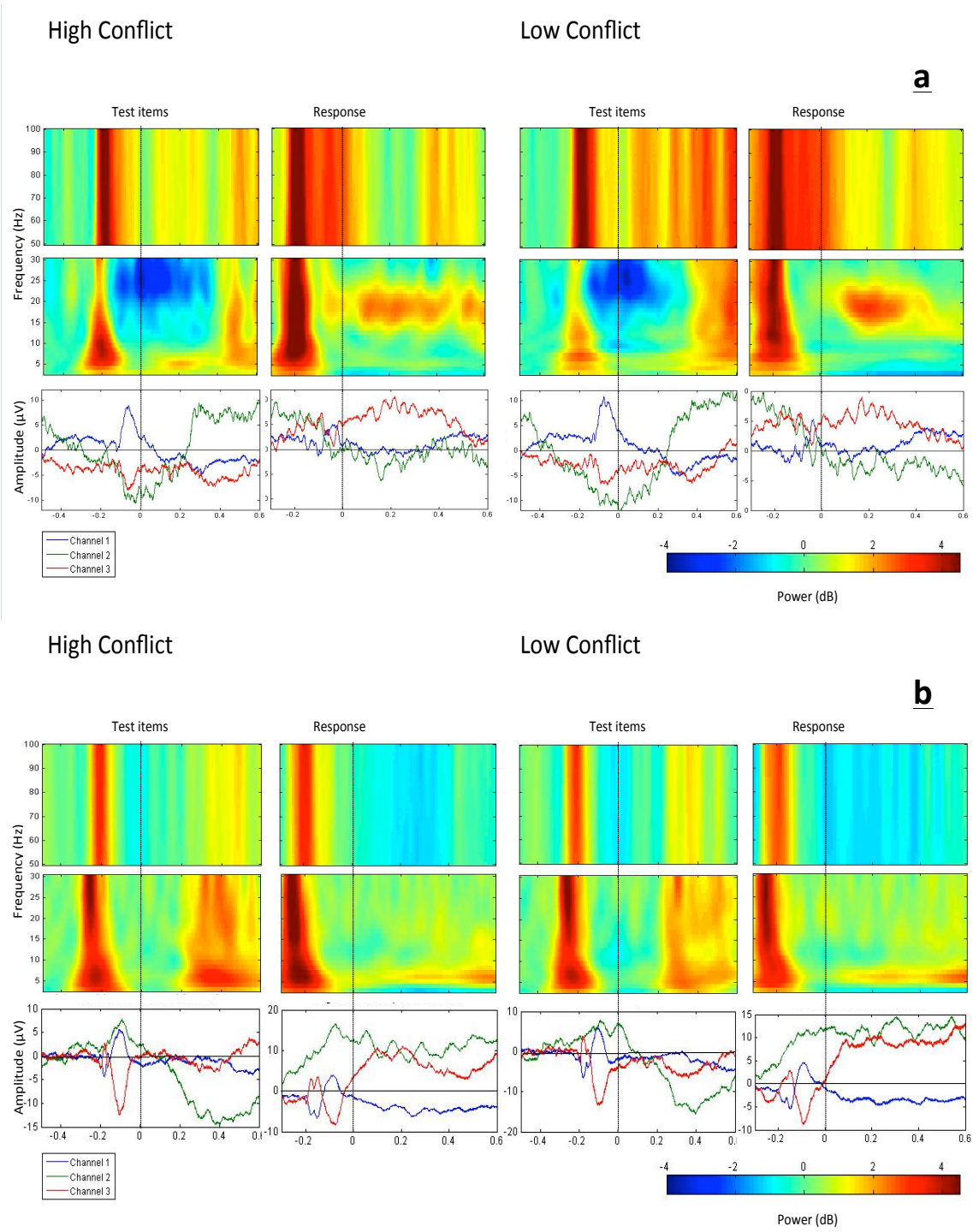


Figure 41. *Time-frequency decomposition and event-related potential activity in the left ACC for patient AP (a) and WS (b) on high- and low-conflict trials in the conflict WCST analogue. Stimulus-locked and response-locked activation plotted on adjacent columns. As response time varied, data is plotted from -0.5s pre-test stimuli onset to 0.6s post-test stimuli onset for the stimulus-locked data and from -0.3s pre-response to 0.6s post-response for the response-locked data. During the inter-trial interval (ITI), the choice-item stayed on the screen for one second after a correct response. Therefore, up to 1s post-response represents the feedback period, whereas from 1s to 1.5s post-response (corresponding approximately to -0.5s to 0.2s pre-trial onset) represents blank-screen ITI period.*

Given the presence of beta desynchronization in patient AP, we then looked at ACC activity for stimulus-locked data on adaptation trials in this patient (LH and HH, see Figure 42). Again, AP presented what appears to be motor-related ERP and oscillatory activity 200-100ms prior trial onset and response, and around 400ms after trial onset. He also presented a post-stimuli decrease in beta power (20 to 30Hz) starting from the time of trial onset and more sustained on LH than on HH trials.

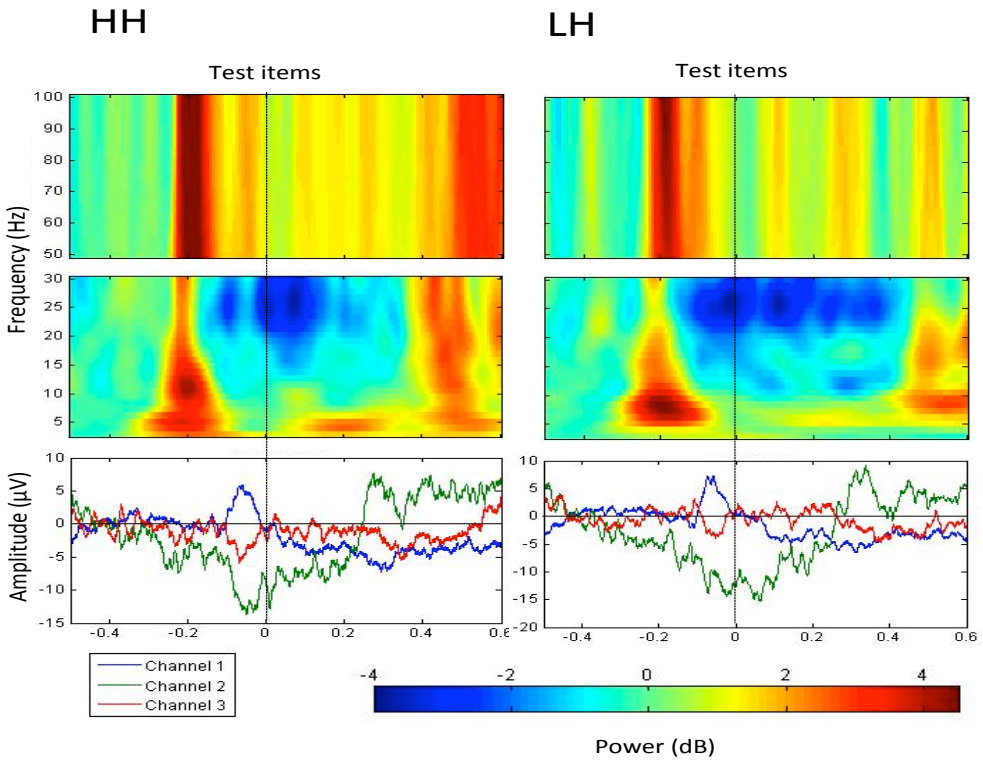


Figure 42. *Time-frequency decomposition and event-related potential activity for patient AP on HH and LH adaptation trials in the conflict WCST analogue. Data is plotted from -0.5s pre-test stimuli onset to 0.6s post-test stimuli onset.*

Patient MG did not present any discernable consistent ERP activity or oscillatory activity, and recordings from his right-hemisphere microelectrode presented a

substantial amount of noise, which rendered the data rather uninformative (see Appendix 3, Figures A11 and A12). Data from this patient was therefore not analysed further. There are a number of possible explanations for this. Referencing in this patient was to a contact on the mastoid process, rather than to another contact within ACC, as in the case of AP and WS. Recordings in the latter two patients are therefore localized to differences within a more restricted region. The different recording setup might have been more sensitive to environmental noise than in the other two patients. A further explanation could be due to the orientation of the electrode within the neural tissue. As can be seen from the MNI coordinates reported in Table 10, the orientation of MG's microelectrodes presents a less steep angle than that of AP and WS, which might possibly influence the recordings.

Simon task

Figure 43 illustrates the electrode average time-frequency decomposition and single-channel intracranial ERPs for patient WS for the high- and low conflict conditions of the Simon task. Similarly to the WCST, patient WS again showed an increase in power across all frequency ranges around 200ms prior to trial onset (i.e. screen touch) and response, in both left and right ACC, likely an index of motor-related activity. He presented some alpha suppression (8-12Hz) around 200ms after trial onset in the right ACC electrode, only for high-conflict trials.

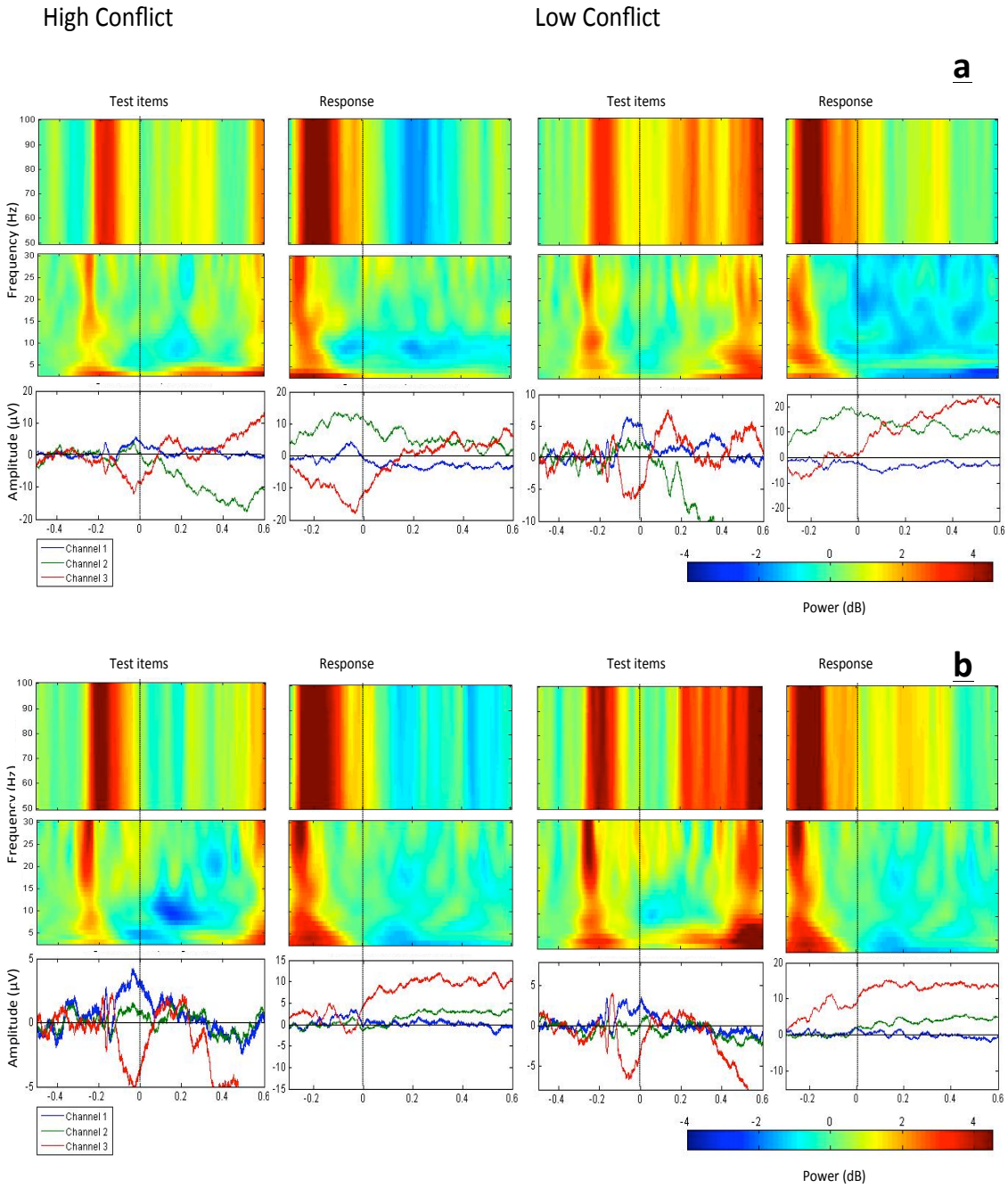


Figure 43. *Time-frequency decomposition and event-related potential activity in the left (a) and right (b) ACC electrodes of patient WS on high- and low-conflict trials in the Simon task.*

Discussion

Statistical analysis on the speed of response for the WCST indicated a tendency for participants to be faster at responding on low-conflict than on high-conflict trials (i.e. a conflict effect), regardless of stimulation condition, and that the magnitude of the difference in speed between the two trials was unaffected by stimulation. This suggests that stimulation of the ACC is unlikely to affect responses to conflict on the current trial in the WCST. These results are in line with our human neuropsychological investigation (Chapter 1) and the corresponding data from monkey neuropsychological investigations (Mansouri et al., 2007), and therefore with the findings that ACC stimulation and ablation have comparable effects (Cohen, Kaplan, Moser, Jenkins, & Wilkinson, 1999; Liu & Postupna, 2008). Due to the small number of patients who successfully completed the Simon task, statistical analysis was not possible on this sample. However, inspection of the numerical values also seemed to show no effects of stimulation on the size of the conflict effect during the Simon task. Findings from both tasks would therefore be consistent with the body of literature showing that ACC damage does not impact responses to conflict on the current trial (Fellows & Farah, 2005; Sheth et al., 2012; Stuss et al., 2001; Swick & Jovanovic, 2002; Vendrell et al., 1995), regardless of whether the task involves competition between two instructed responses (WCST) or a habit and instructed responses (Simon).

The effects of ACC stimulation on adaptation are more difficult to pinpoint. In the WCST there appeared to be a tendency towards faster responses on HH trials than on LH trials, as it is normally observed in this task, only when the stimulation was turned off. This is inconsistent with the hypothesis that the effects of DBS and lesion on performance on this task are comparable (since we did not find an effect of ACC lesion on adaptation in the WCST - see Chapter 1 and Mansouri et al., 2007, for the comparable non-human primate data). However, numerically, this effect seemed to be driven by an increase in speed of response on LH trials in all but one patient (WS), rather than a decrease in speed of response on HH trials (HH trials also showed a numerical increases in speed in all patients except for WS, although not as large as the increase on LH trials). It seems likely that, while the effects of DBS on adaptation might differ from those of lesion, they seem to somehow aid performance on LH trials, rather than negatively affecting responses on this task. Therefore, this would not necessarily be inconsistent with the hypothesis that conflict-resolution in the WCST is not dependent on ACC processing, as interfering with its activity seems to speed up responses on trials that should be most sensitive to conflict costs (LH). We observed similar indications of increases in speed on LH trials following TMS to dlPFC (Chapter 2). One possible explanation for these findings is that stimulation of the ACC interferes with signalling from the ACC to the dlPFC and this, in turn, produces similar effects to interference with dlPFC activity. The small number of participants, combined with the variability in electrode placement and type of stimulation, precludes us from reaching any definite

conclusions at this stage, but this is a finding that might be better clarified in the future, with the addition of more patients to the dataset.

Preliminary analysis of the intracranial LFP data has not revealed consistent conflict-related modulation across patients. However, AP's results might indicate the presence of oscillatory activity that is consistent with the literature. AP showed beta suppression following the onset of test stimuli in the WCST, particularly for high-conflict trials and specifically LH trials. This is a similar pattern to that observed by Cohen and colleagues (2008) in the Flanker task and it would be consistent with the hypothesis that high-conflict trials in our WCST analogue elicit stimulus-based conflict effects (which are thought to trigger activations in more caudal ACC regions, see Bush, Luu, & Posner, 2000; Milham & Barich, 2005; van Veen & Carter, 2005; Badzakova et al., 2008; Kim, Kroger & Kim, 2011) that might trigger the recruitment of cognitive control on subsequent trials. Indeed, beta-band suppression was also more prominent for LH trials (i.e. the first trial during which conflict is detected) than on HH trials (i.e. a trial where cognitive resources should already be engaged from the previous trial).

Only one of the two patients included in the WCST intracranial analysis showed this pattern of activation, which limits our interpretation. However, this might be somewhat related to his behavioural performance. AP showed much larger conflict and adaptation effects compared to all three of our other patients. He was also the patient who showed the largest effect of stimulation on adaptation, with the elimination of the

adaptation effect due to a numerical increase in the speed of response on LH trials. As previously mentioned, this increase in speed of response on LH trials resembles that observed during TMS to dlPFC in the same task, and might be an indication of 'relayed' interference from the ACC to the dlPFC. ACC-dlPFC interaction during conflict tasks is very frequently reported in imaging literature (e.g. Durston, 2003; Egner & Hirsch, 2005; Kerns, 2006; Kim et al., 2012) and TMS to the dlPFC has been shown to increase blood-flow responses in the ACC (Paus, Castro-Alamancos, & Petrides, 2000), indicating that stimulation to this region can affect activity in the cingulate. The hypothesis that stimulation of the ACC could similarly affect activity in the dlPFC, resulting in similar behavioural effects, is therefore a realistic possibility. Oscillatory synchronizations and desynchronizations have been associated with communication between the ACC and other brain areas, such as the hippocampus (Remondes & Wilson, 2013; Totah, Jackson, & Moghaddam, 2013). Beta oscillatory activity in the ACC during the WCST, especially given the similar effects of ACC and dlFPC stimulation on performance during trials that show particularly strong conflict-related beta modulation, could reflect ACC-dlPFC communication in aiding adaptation.

The finding that there may be some conflict-related oscillatory modulation in the ACC for a task involving competition between two instructed responses (WCST), and that stimulation of this region in the patient showing this type of oscillatory activity affects adaptation, is inconsistent with the hypothesis that ACC is only involved when competition is created between a habitual response and an instructed response.

However, AP did not carry out the Simon task with the stimulation turned off, so we cannot conclusively rule out the hypothesis that his oscillatory activity might be differently modulated or more pronounced in this type of task and that stimulation might affect performance differently in the two tasks. WS, the only patient included in our analysis who carried out both tasks, failed to show any clear conflict-related activity in either task, which prevents us from drawing any conclusions on this issue.

Discrepancies between the results across patients could be due to differences in electrode positioning. AP's electrode was more posterior and slightly more dorsal than the corresponding electrode in WS's left ACC (see Figure 44 and Table 12). Indeed, electrodes showing beta suppression in Cohen and colleagues' study (2008) were located slightly more caudally than the ones in most of our patients. Similarly, Sheth and colleagues (2012) appear to have recorded from more dorsal ACC regions than the ones we recorded from in most of our patients. AP's electrode might therefore be located closer to the dorso-caudal parts of ACC that are active during conflict trials, compared to our other patients.

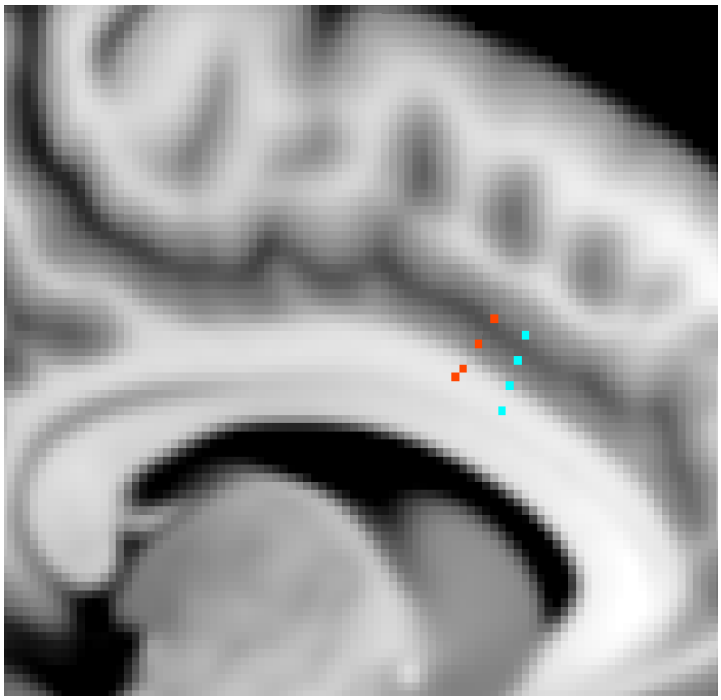


Figure 44. *Comparison of Y electrode coordinates between patient AP (red) and WS (blue), mapped onto a standard MNI brain.*

This potential explanation raises the question of whether, and how much, conflict-related activity in our recordings might be masked or confounded by motor-related activity. For example, in AP, beta suppression starts during the movement to touch the screen in order to initiate the trial. However the same type of suppression is not present following the movement to touch the screen for response, and is more sustained in the high-conflict condition, suggesting that it might be conflict-modulated, rather than related to movement. One explanation for the presence of beta suppression prior to the start of the trial is that it might be related to anticipatory activity, rather than conflict-detection, perhaps as a function of previously processed feedback. Beta desynchronization in the ACC has recently been suggested to reflect the preparation of a response during early, but not late, learning phases (Altamura, Carver, Elvevåg,

Weinberger, & Coppola, 2013), a stage during which feedback processing is crucial to performance, and beta coherence between brain areas had been correlated with the degree of stimulus expectancy (Gross et al., 2006), which also modulates anticipatory behaviour. In AP, beta oscillatory activity was present both during trial and during feedback. Suppression of beta in the decision-making period was followed by a beta increase in the feedback period, and both suppression and enhancement were more sustained on high-conflict than low-conflict trials. When we look at adaptation trials, however, HH trials, which follow feedback periods characterized by sustained beta enhancement, present a shorter period of beta suppression compared to LH trials. One possibility is that the onset of beta-suppression prior to the screen touch initiating the trial (and during the trial itself) is related to the amount of feedback-related beta enhancement on the previous trial, and that this might represent an index of feedback-dependent, rather than conflict-dependent, signaling that drives adaptation on subsequent trials.

Unfortunately, dissociating motor from cognitive activity has proven difficult in this particular dataset due to significant confounds imposed by the use of touchscreens for response. Requiring subjects to touch the screen in order to start the trial, which appears to lead to a substantial ERP as well as increase of oscillatory activity at all frequencies, might be overshadowing any conflict-related processing during the very early stages of the trial, unless it is particularly strong (like in the case of AP). Nonetheless, some methodological changes could still allow us to further test our

hypotheses by recording from this same area in the future. Using button-boxes for response, which would greatly reduce the amount of limb movements, could represent a much more suitable option. A potentially even better option is to administer eye-tracking versions of these tasks. Cingulate motor areas (CMA) project directly onto other frontal cortical motor areas and the spinal cord (Paus, 2001), which directly control limb movements, including finger movements, and stimulation of the CMA regions have been shown to evoke movement in the limbs (Dum & Strick, 2002). However, stimulation of the CMA does not seem to trigger eye movements, and connectivity to areas controlling saccades appears much weaker (Schall & Emeric, 2010). Therefore, using eye-tracking versions of the task, as well as removing self-initiation of the trials, could substantially minimize any motor-related activity and allow us to discern with much more clarity whether any conflict-related modulation is present in caudal area 24 of the ACC. We hope to have the opportunity to implement these changes in future testing with DBS patients.

A further confound was presented by the referencing used in the group of patients. Of the patients whose intracranial recordings were analyzed, electrode contacts were referenced to each other for two of the patients (bipolar recordings) and to the mastoid in one patient (monopolar recordings, converted into bipolar recordings offline). Furthermore, the angle of implant of the microelectrodes was different between the bipolar and monpolar configurations. Only the two patients with bipolar configurations showed event-related activity, be it motor-related or conflict-related. It is

unclear whether the differences in recordings are due to environmental noise or differences in referencing, rather than electrode positioning. These issues could be better clarified should we have access to more patients with monopolar configurations in the future. Finally, one unavoidable confound in this patient population is that of pain or discomfort when the stimulation is turned off. While we ensured that patients were as comfortable as possible during these blocks, and interrupted testing as soon as pain and discomfort exceeded low-to-mild levels, comfort presents a covariate to stimulation, which cannot be controlled for in chronic-pain patients. Nonetheless, performance did not appear to be significantly modulated in a manner that would be consistent with confounding effects of discomfort (e.g. severely worse performance when the stimulation was turned off), and testing was interrupted when discomfort reached levels that might have been problematic for the patient's performance (as in the case of AP during the Simon task).

To conclude, preliminary findings from our pilot study investigating the effects of intracranial stimulation to caudal ACC regions, while currently inconclusive due to limited sample size and confounds in the electrophysiological data, have nonetheless provided a rich variety of data to help inform further hypotheses about the role of ACC in conflict tasks, as well as methodological improvements, that could be investigated in further DBS research. We can tentatively conclude that stimulation does not seem to affect speed of response on high- versus low-conflict trials, regardless of whether the task involves competition between instructed responses (WCST) or between a habit and

instructed responses (Simon). Some modulatory activity of stimulation on adaptation effects in the WCST seems present, and appear comparable to those obtained from dlPFC transcranial magnetic stimulation. One patient showed beta-band suppression high-conflict trials, particularly LH trials, in the WCST that appeared to be correlated to his response speed. Furthermore, he showed beta-enhancement during the feedback period, which was larger for high-conflict than low-conflict trials. As beta-suppression onsets before the start of the trial, one possibility is that beta synchronization during feedback drives anticipatory activity in the ACC on subsequent trials, and that anticipatory activity is reflected in the degree of beta desynchronization. Furthermore, given the finding that stimulation on trials showing conflict-related beta modulation affected adaptation in a similar fashion as TMS to parts of the dlPFC, we hypothesized that beta oscillations in this region might be an index of ACC-dlPFC interaction. We hope to further investigate these issues in DBS patients in the future.

CHAPTER 6

Effects of Frontopolar Cortex lesions on abstract rule acquisition and integration in the macaque monkey

Introduction

One fundamental aspect of conflict resolution is the ability of the cognitive system to select a response in the face of uncertainty, caused by the presence of many possible competing responses, distractors and/or by contextual changes. Behavioural adaptation, in particular, relies on the ability to consistently ignore task-irrelevant information across time in order to optimize performance. In some settings, however, distractors may represent a potential source of new, possibly larger, rewards that may be worth exploring. The exploration of alternative options is a crucial aspect of the flexibility needed to maximize behavioural outcomes.

As seen in previous chapters, the ACC and dlPFC are thought to be involved in alerting the cognitive system about changes in the level of conflict or uncertainty and in implementing the cognitive control necessary to "stick to course" in the presence of distractors or competing options. Evidence indicates that the frontopolar cortex (FPC), might be involved in exploring the value of those other, unchosen options and promoting behavioural flexibility (Boorman, Behrens, Woolrich, & Rushworth, 2009;

Buckley et al., in preparation; Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006; Koechlin, Danek, Burnod, & Grafman, 2002; Ramnani & Passingham, 2001; Strange, Henson, Friston, & Dolan, 2001).

The FPC is the largest area of the PFC, making up 1.2% of the total brain volume in humans. Relatively to the rest of the cortex, it is twice as large in humans as it is in other great apes (Semendeferi et al., 2001). This area has large number of connections with most other PFC areas, the ACC and the anterior temporal cortex, and is possibly the only PFC area that lacks connections with more primary sensory areas (Petrides & Pandya, 2007; Yeterian et al., 2012; Ramnani & Owen, 2004). It receives input from multi-modal association areas as well as limbic areas (Petrides & Pandya, 2007). The relatively recent development of FPC in primate evolution, as well as its disproportionate size in humans compared to other primates, and its dense interconnections with other PFC areas, has made it a primary candidate to be the structure responsible for the type of higher-order behaviour that characterizes human cognition, such as complex reasoning and abstraction.

From a functional perspective, FPC remains the least well-understood area of the human brain. Its activation, as well as de-activation, has been reported across a wide range of complex cognitive tasks, such as relational reasoning and integration of higher-order, abstract representations (Badre & D'Esposito, 2007; Badre, 2008; Christoff et al., 2001; Koechlin, Ody, & Kouneiher, 2003; Kroger et al., 2002; Reynolds, McDermott, &

Braver, 2006), maintaining and processing internal subtasks and states (Koechlin, Basso, Pietrini, Panzer, & Grafman, 1999; Reynolds et al., 2006), multitasking and reallocating attention (Gilbert et al., 2006), episodic and prospective memory (Braver & Bongiolatti, 2002; Gilbert et al., 2006; Reynolds et al., 2006; Rugg et al., 1998), and learning and exploration (Boorman et al., 2009; Daw et al., 2006; Koechlin et al., 2002; Ramnani & Passingham, 2001; Strange et al., 2001). Lesion studies in humans have indicated that damage to FPC impairs the ability of patients in dealing with complex or volatile situations (Burgess et al., 2000; Dreher et al., 2008). Despite the apparent role of FPC in complex cognitive tasks, single-cell recordings indicate that neurons in this region encode relatively simple information. They are sensitive to the preferred direction of a saccade, but only during (or close to) the feedback period of a trial, and they are particularly sensitive to positive feedback, but not to the rewarding properties of the feedback (Tsujimoto, Genovesio, & Wise, 2010). The specific contribution of FPC to these very diverse tasks has therefore been difficult to pinpoint.

One of the most prominent accounts of FPC function proposes a role for this region in relational reasoning and integrating multiple abstract representations to select a response. The cascade model of PFC function (Koechlin et al., 2003; Koechlin & Summerfield, 2007) postulates that PFC is organized along a hierarchical rostro-caudal gradient, with more anterior areas of PFC processing and manipulating increasingly more general, abstract and temporally distant information. FPC, being situated at the

extreme anterior termination of this gradient, is thought to process and integrate abstract plans, second-order relations, and use contextual signals to guide behavior.

Ramnani & Owen (2004), in their review of frameworks of FPC function, evaluated most of the main functional accounts of FPC (memory, relational integration, processing of internal states and branching) according to three criteria that the accounts were required to meet in order to be considered adequate to explain FPC's functional properties: converging evidence (evidence is consistently obtained by different methodologies), information-specificity (the framework specifies what kind of information is being processed) and functional and anatomical specificity (the task, or aspects of the task, reliably and exclusively recruits FPC). The only account that met all these criteria suggests a role of FPC in the relational integration of two or more cognitive operations. The authors proposed that FPC is actively engaged when decision-making requires the application of two or more cognitive operations or rules. Experimental work by Ramnani and colleagues provide support for the idea that FPC is recruited when subjects are required to integrate two or more unrelated rules in order to solve a problem (Ramnani & Miall, 2003) and during the acquisition of a second rule after overtraining on the first (Ramnani & Passingham, 2001).

A more recently developed alternative framework of FPC function, which was not included in Ramnani and Owen's review, proposes a role for this region in exploring the value of unchosen options, promoting behavioural flexibility and learning.

Neuroimaging evidence suggests that that FPC is robustly engaged during early stages of learning and increasingly disengages over the course of the learning process (Strange et al., 2001; Koechlin et al., 2002). Learning is characterized by a period of enhanced uncertainty, until the rule, strategy or skill has been acquired adequately enough to allow for confident performance. Yoshida & Ishii used a hidden Markov model (HMM) to identify correlations between regional BOLD signal and measures of hidden state uncertainty in a navigation task. The uncertainty in their study was manipulated by varying whether subjects were given direct instructions as to what direction to follow in a maze (low uncertainty) or to identify their location solely on the basis of visual cues and choose the next step out of a number of possible routes to reach their ultimate target (high uncertainty). They found that FPC activity correlated with a measure of current-state uncertainty (Hidden Current Position entropy), with higher activation for higher uncertainty (Yoshida & Ishii, 2006). More recently, Boorman and colleagues (2009) used a Bayesian model to investigate variations in FPC BOLD signal during a Two-Armed Bandit task, where subjects were given no explicit instructions regarding optimal switching between two possible options, but had instead to rely on their internal estimates of the varying probability of reward associated with selecting each option at a given point in time. They found that FPC activation correlated with the accumulated evidence in favor of switching from one option to the alternative, unchosen option. These findings strongly implicate the FPC in exploring and evaluating unchosen options to facilitate learning.

The vast majority of the evidence for the different accounts of FPC function comes from neuroimaging. As we saw in the case of ACC function, while imaging can help inform on whether FPC is involved in various cognitive processes, it cannot establish whether it is necessary for said processes. Furthermore, lesion studies can help dissociate competing frameworks of FPC function. In this study we sought to investigate whether the relational integration or the exploratory/evaluation account of FPC more adequately describe its role in cognition, by assessing the effects of FPC ablation in non-human primates. We trained two groups of macaque monkeys (a FPC lesion group and a control group) on a decision-making task that required the acquisition and combination of two abstract rules ('same' and 'smaller than') in order to select the correct option out of 4 possible stimuli, and obtain a reward. The animals were first trained to apply a simple matching-to-sample rule ('same as'), followed by learning to apply a second abstract rule ('smaller than'), and finally to integrate both rules. Using this experimental paradigm, we tested whether FPC is necessary for the acquisition of new rules once a different rule has been previously established as a successful problem-solving strategy (therefore supporting the exploratory/evaluation framework) or whether it is necessary to integrate these abstract rules after acquisition, in order to perform the correct response (therefore supporting the organizational hierarchy framework).

Acquiring a new abstract rule, after the consolidation of a previous strategy involving a different rule, involves the exploration of new options other than the ones that were previously preferred. Combining two abstract rules once they both have been

successfully acquired, on the other hand, does not involve exploration, but integration of previously acquired knowledge about the environment. Therefore, if FPC is necessary for rule integration, we predicted that the performance of FPC lesion animals will be significantly worse than that of control animals when required to apply two abstract rules to solve a problem. On the other hand, if FPC is necessary only to explore new options and/or extract the value of unchosen options, we predict that the FPC lesion animals will be impaired during the acquisition of the second, novel, abstract rule, but their performance will be comparable to control animals, when eventually required to apply two established abstract rules to solve a problem.

Methods

All procedures were licensed in compliance with the United Kingdom Animals Scientific Procedure Act, 1986. Seven female macaque monkeys (*Macaca Mulatta*) were trained and tested as part of this study. All animals had training experience on other experimental tasks. Four animals had previously received surgical ablation of the Frontopolar Cortex, and three animals had not undergone any previous surgical procedure and therefore were assigned to a control group. Prior to surgery, the animals had been matched on pre-operative performance and then randomly assigned to one of two groups (FPC or Control). On average, 16 months passed between surgery and training on the current task.

Surgery

The animals were sedated, intubated and anesthetized using Isoflurane (1.0 – 3.0%). An operating microscope was used to aid the operation and all operations were carried out in sterile conditions. A bone flap was raised over the right and left FPC and the dura mater was cut and reflected, exposing the anterior-posterior extent of the posterior sulcus. The lesion was created using a small-gauge metal aspirator insulated to the tip. The posterior extent of the lesion on the dorsal surface was 2mm posterior to the anterior tip of the principal sulcus. The posterior extent of the lesion on the orbital and medial surfaces was limited by an imaginary line 2mm posterior to the anterior tip of the principal sulcus. All cortex anterior to these limits was removed by aspiration. Only grey matter was removed with the exception of the most anterior part of the lesion. The dura mater was then sewn back, the bone flap replaced, and the wound closed in layers. The operated animals rested for 20 days after surgery before beginning post-operative training.

Apparatus and stimuli

Subjects were trained and tested individually in a dark experimental cubicle. They sat, unrestrained, in a transport cage at a distance of approximately 30cm from a touchscreen, and could reach through the gaps in the transport cage bars to touch the items on the screen to make their responses. A pellet-delivery system was set up to deliver flavoured reward pellets (Monkey Delight Tab 190mg, supplied by TestDiet™) whenever a correct response was made. The pellets were delivered into a food well

positioned underneath the screen, that the subject could also reach through the bars of the transport cage. Additionally, a metal lunchbox (200mmx100mmx100mm) was also positioned underneath the screen and opened immediately following completion of the task to provide the animal's daily diet, consisting of fruit, nuts, and wet monkey chow. Water was provided in the colony room ad libitum. The task was programmed using Turbo Pascal (Borland), run in DOS on a desktop PC and presented on a 20.1" colour touchscreen (TFT LCD TS200H GNR). The stimulus set consisted of cliparts in two different sizes (10 degrees of visual angle in width by 10 degrees in height for the larger items and 5 by 5 degrees for the smaller items), and they were presented against a black background.

Training and testing procedures

During training, the animals were exposed to a series of 7 different training stages. Stages 1 to 5 trained the animal to perform a simple matching-to-sample task, therefore learning a 'same as' rule. Stages 6 and 7 involved the acquisition of a new abstract rule, 'smaller than'. Finally, Stage 8 involved the combination of the rules 'same as' and 'smaller than'.

Stages 1-5

During Stage 1, a randomly selected clipart, the sample item, was presented in the centre of the screen. Once the subject touched the image on the screen, the sample item disappeared and two items appeared on the screen: one match item, identical to

the sample item, and one unrelated, non-match item. These items were presented randomly in 2 of the 4 corners of the screen, 10 degrees of visual angle diagonally to the center item. The subject was rewarded with a pellet for selecting the match item, therefore being reinforced for applying a simple matching-to-sample (or 'same as') rule. Once a correct response was performed, 8 seconds elapsed before the following trial, to allow time for the subject to ingest the pellet. If an incorrect response was performed, no reward was provided and a 16 seconds delay was introduced before the following trial. Stage 2 was procedurally identical to Stage 1, but the sample item remained on the screen during the decision-making period, effectively becoming a simultaneous matching-to-sample task, as opposed to a delayed matching-to-sample task. Stage 3 also followed the same procedure as the previous stages, with the addition of an extra, non-match item to the possible choices. In Stage 4, a further non-match item was added, bringing the total number of possible choices up to four. In Stage 5, the four possible choices included two match items in two different sizes (same size as the sample item or smaller than the sample item) and two non-match items which were of the same size as the sample item, exposing the subject to the 'smaller than' item for the first time. The subject was rewarded for selecting the same-size, match item. Stages 1 to 5 were each carried out by the subject over several consecutive days, until an accuracy criterion of 80% or more was achieved.

Stages 6-7

During Stage 6 (see Figure 45), the subject was required to learn to apply the 'smaller than' rule for the first time. The four possible choices included two match items, one of the same size and one smaller than the sample item, and two non-match items, both of the same size as the sample item. The subject was rewarded for picking the smaller item out of the four. During this stage, the subject carried out 100 trials a day over 3 consecutive days. Training continued in Stage 7, with the introduction of help-errors. When the subject made more than 5 errors in a row, the computer did not present the big-match item on the next trial, in order to encourage the acquisition of the new, 'smaller than' rule. This stage was carried out over several consecutive days until an accuracy criterion of 80% or more was reached. The subject was then moved on to the final stage.

Stage 8

During Stages 6 and 7 the animal could perform the correct response by selecting the smaller item even if she disregarded the identity of the item, as there was only one small item in the set. In Stage 8, one item was identical in both size and identity to the sample item (big-match), one item shared the same identity as the sample item but was smaller in size (small-match), one item was unrelated in identity to the sample item but in the same size (big non-match), and one item was unrelated in identity to the sample, but related in identity to the big non-match item, and was smaller in size (small non-match). The subject, therefore, needed to integrate the two rules, 'same as' and 'smaller

than', in order to select the correct item, as simply selecting the smaller item would cause the animal to perform at chance. An example of a typical trial is provided in Figure 45. Successful performance on the task was defined as an accuracy criterion of 80% or more correct trials over the course of one experimental session.

Reaction times and accuracy for all stages were recorded and analysed.

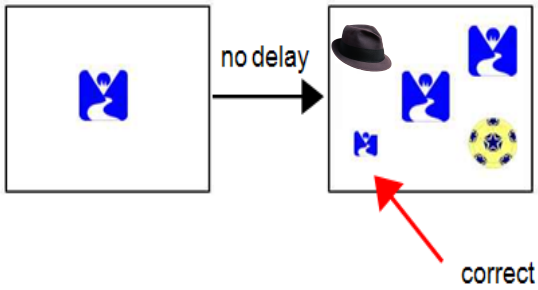


Figure 45. *Example of a simultaneous matching-to-sample trial during Stage 6 of the task, where the animal was first required to apply the 'smaller than' rule.*

Results

Data analysis was carried out on errors and on response time data. Analysis on percentage values was carried out on the ARCSIN transformation of the data.

Independent-sample t-tests were carried out on the reaction time data, for both correct and error trials, from the overall stages, as well as across sessions within each stage. There were no significant differences in response time between FPC and control animals at any phase of training or testing, for either type of trial ($p > .05$).

In order to investigate the role of FPC during the acquisition of the two-rule strategy, we analysed the percentage errors during Stage 6, where the animals began the acquisition of the 'smaller than' rule. We carried out a 3x2 mixed ANOVA with Session (1, 2 and 3) as the within-subject factor and Group (FPC and CON) as the between-subject factor. There was a significant main effect of Session ($F_{(2,10)} = 6.91$; $p = .013$), with the percentage of errors decreasing across sessions, as well a significant interaction between Session and Group ($F_{(2,10)} = 5.22$; $p = .028$). There was no significant main effect of Group ($F_{(1,5)} = 1.61$ $p = .26$). We investigated the significant interaction further with independent-sample t-tests. There was a significant difference between FPC and Control animals in the decrease in errors from Session 1 to Session 2 ($t_{(5)} = 3.18$, $p = .024$), with Control animals showing a larger decrease (mean = 23.3%, sd = 0.16) than FPC animals (mean = -.03%, sd = .07), as can be seen in Figure 46. This was confirmed by non-parametric Wilcoxon Signed Ranks t-tests ($z = -2.12$, $p = .034$). The percentage error for control animals decreased from around 80% to 50%, whereas it remained stable for FPC animals.

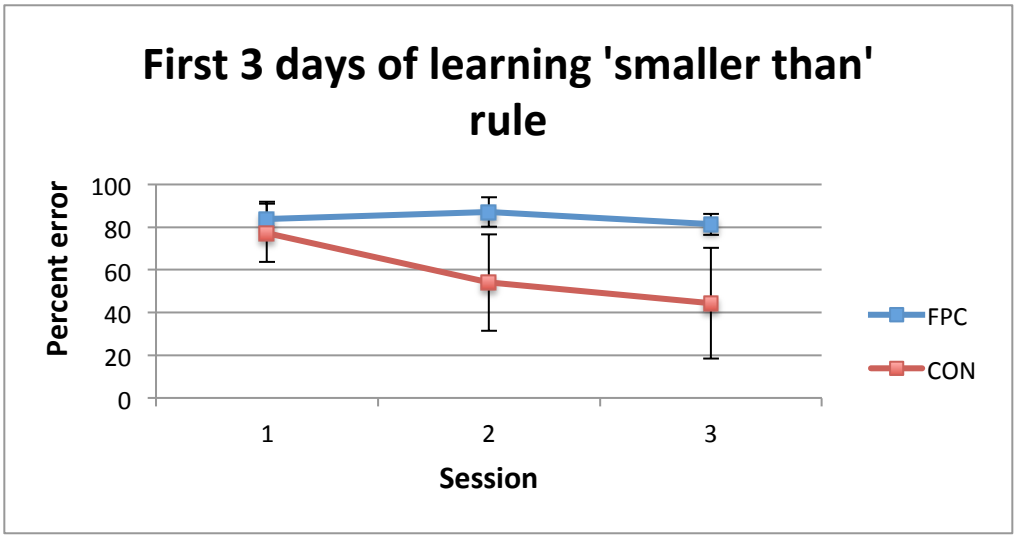


Figure 46. *Percent error across sessions for the two groups (FPC animals, blue, and Controls, red). Error bars represent Standard Error of the Mean (SEM).*

Animals could commit two types of errors: Big-Match errors (i.e. trials in which the animal applied the same rule as during the first 5 stages of training; or exploitative errors) and Non-Match errors (i.e. trials in which the animal selected an item that was unrelated to the previously acquired rule, but was not the correct item; or exploratory errors). We therefore carried out an analysis on the proportions of errors during Stage 6 that were Big-Match errors or Non-Match errors, across the three sessions, to investigate whether either type of error was more prevalent for either group.

The error data was entered in a 2x3x2 ANOVA with Error Type (BigMatch or NonMatch) and Session (1, 2 and 3) as the within-subject factors and Group (FPC or CON) as the between subject factor. The only significant main effect was that of Type ($F_{(1,4)}$ = 262.04, p = .000), with a higher rate of Big Match errors (mean = 76.62%, sd = .08) than Non-Match errors (mean = 20.61%, sd = .09), overall (Figure 47). There was

also a significant interaction between Type and Group ($F_{(1,4)}$ = 10.54, p = .03). Independent-sample t-Tests revealed a significant difference in the proportion of Big-Match errors between FPC and control animals ($t_{(5)}$ = 3.31, p = .021), with a higher proportion of Big-Match errors for FPC animals (mean = 80.63%, sd = .02) than controls (mean = 71.61%, sd = .04). There was also a significant difference in the proportion of Non-Match errors between FPC and control animals ($t_{(5)}$ = 4.08, p = .01), with a higher proportion of Non-Match errors for control animals (mean = 23.36%, sd = .02) than FPC animals (mean = 16.09%, sd = .03). Both differences were also confirmed by non-parametric t-tests (p = .034).

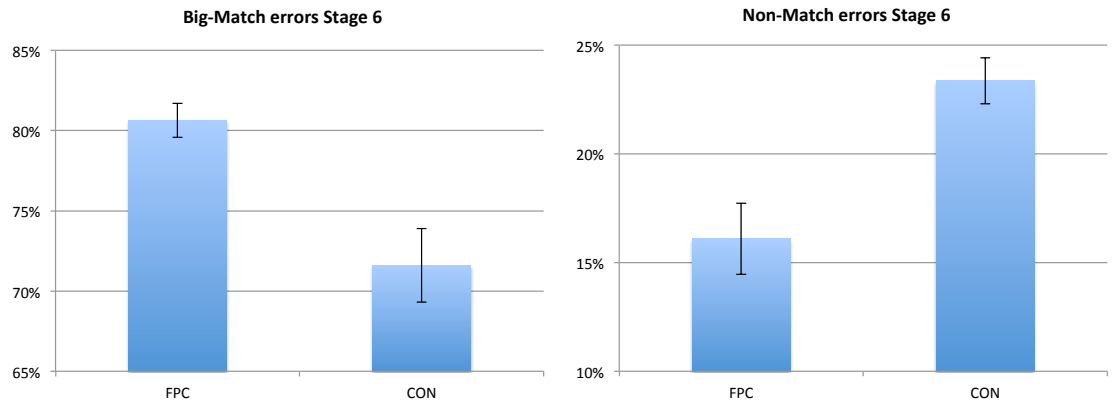


Figure 47. *Percentage of Big-Match (exploitative) and Non-Match (exploratory) errors during Stage 6 for the two groups (FPC = FPC animals, CON = Control animals). Mean percentage error is plotted along the y-axis. Error bars represent SEMs.*

Given the different prevalence of the two types of errors in the two groups, we then looked at whether control animals were more likely than FPC animals to commit a Non-Match error (i.e. an exploratory error) following a Big-Match error or another Non-Match error. A one-way ANOVA revealed no significant difference between the two

groups (p > .05), although there was a trend towards a significant difference for the proportion of Non-Match errors following Big-Match errors (t(6) = 4.47, p = .088), with higher proportion of Non-Match errors for control animals (mean= 15.22%, sd = .03) than FPC animals (mean = 10.29%, sd = .03), as shown in Figure 48.

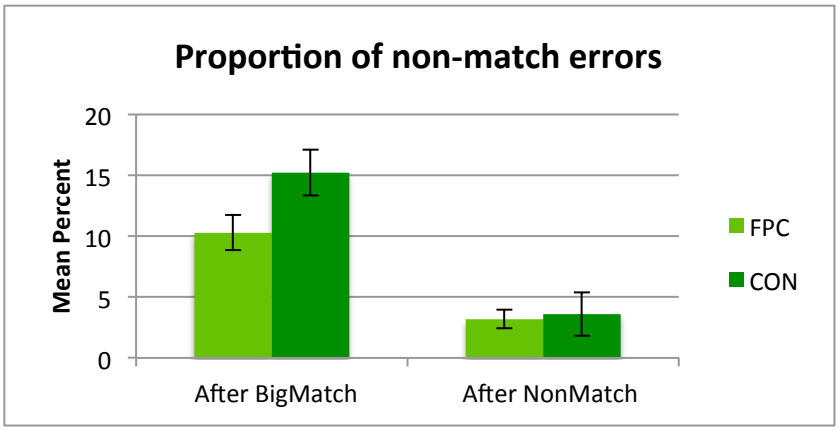


Figure 48. *Proportion of Non-Match (exploratory) errors after a Big-Match (exploitative) errors or a Non-Match (exploratory) error, for the two groups (FPC = FPC animals, CON = Control animals). Error bars represent SEMs.*

Once correction trials were introduced during Stage 7 (which removed the Big-Match option from the items set following 5 consecutive errors), performance between the two groups was undistinguishable and all animals reached criterion (>90% correct) within 3 sessions.

In order to investigate the role of FPC in the integration of two abstract rules, we carried out independent-sample t-tests on the number of errors accumulated by animals in order to reach criterion during Stage 8, where animals had to apply both abstract rules concurrently in order to perform the correct response. There was no significant

difference in the number of errors between the FPC and the control group (t(5) = 1, p = .362) (Figure 49).

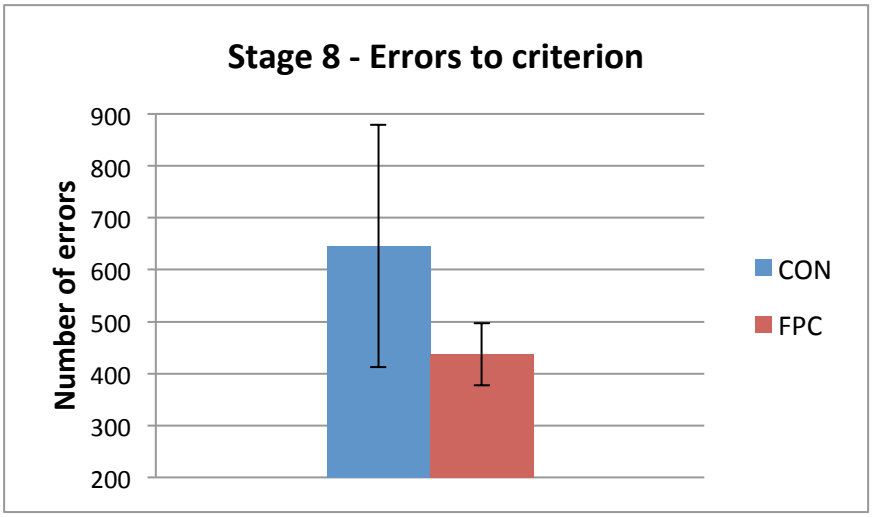


Figure 49. *Number of errors to reach criterion for the two groups (FPC = FPC animals, CON = Control animals), during Stage 8, where they were required to combine the two previously acquired abstract rules for correct responses. Error bars represent SEMs.*

Discussion

The animals in this study learned to apply two different abstract rules ('same as' and 'smaller than') during separate, consecutive training stages and were subsequently required to combine the two rules in order to select the correct option out of a set of 4 items. Comparing the performance of FPC lesion animals to the performance of control animals, we observed a decrease in errors in the control group that was absent in the FPC group during the early stages of acquisition of the second abstract rule, when learning was driven by exploration. Once correction trials were introduced, preventing the possibility of committing an exploitative error (i.e. the selection the Big-Match item,

therefore continuing to apply the previously acquired rule), both groups reached criterion very rapidly. Subsequently, once both rules had been successfully acquired, FPC animals were not impaired, compared to the control group, in their ability to integrate the two abstract rules to select the correct option. This suggests that FPC is not necessary for rule integration, as suggested by Ramnani and colleagues (Ramnani & Miall, 2003; Ramnani & Owen, 2004) and the proponents of the cascade model of PFC function (Koechlin & Summerfield, 2007), but is instead important for aiding the acquisition of new rules and/or response strategies, consistently with findings indicating a role of FPC in learning (Koechlin et al., 2002; Ramnani & Passingham, 2001; Strange et al., 2001). Furthermore, the selectivity of impairment to the training stage that did not include correction trials suggests that the role of FPC in learning might be linked to self-initiated exploration, i.e. the voluntary disengagement from previously rewarded behaviour in order to explore alternative options, rather than exploration that is imposed on the animal by the absence of previously rewarded options, and therefore the impossibility to exploit past choices. Once correction trials are introduced and the animal is forced to explore, performance between FPC and control animals is undistinguishable.

Stage 6 represented the first time animals were required to select the Small-Match item instead of the Big-Match item that had been rewarded during previous stages. Differences in performance between the FPC and control groups emerged when we compared improvement across sessions: between the first and second sessions of

Stage 6, the error rate of control animals dropped from 80% to 50%, whereas it remained constant around 80% in FPC animals during all three sessions. These findings suggest that FPC is important during the early stages of rule acquisition when another rule has been previously successfully reinforced. This is consistent with Ramnani and Passingham's findings (2001), which showed increased activation in FPC when participants had to acquire a new rule after a previous rule had been overlearned. It is also consistent with evidence of higher FPC activation during the early stages of learning, which declines with improvements in task performance, reflecting acquisition of skills or strategies (Koechlin et al., 2002; Ramnani & Passingham, 2001; Strange et al., 2001).

Second-rule acquisition was not impaired in FPC animals once they were prevented from making an exploitative error (i.e. continue applying the same rule that had previously been rewarded). This indicates that FPC animals were able to learn which item was now rewarded as efficiently as control animals if they were forced to make a different choice than the one they were most familiar with. The impairment was therefore restricted to the stage where learning was dependent on self-initiated exploration. The drop from 80% to what is effectively chance level, in control animals, indicates that one session is enough for these animals to learn, when faced with a reward prediction error, that the previously reinforced rule is no longer rewarded, and that new options need to be explored and evaluated. Indeed, our results show that, during Stage 6, although exploitative errors were still very prevalent in both groups, FPC

animals were more likely to commit this type of error than controls, by picking the previously rewarded 'big match' item, and were less likely than controls to select the non-match items, indicating a different explorative tendency. This tendency was not driven by simple switching from an exploitative choice to a exploratory choice on a trial-by-trial basis, as controls were not more likely than FPC animals to make a Non-Match error following a Big-Match error. Consequently, control animals were also more likely than FPC animals to make correct responses by occasionally selecting the previously unrewarded smaller-than item, as shown by the decrease in overall errors over the first two sessions. This suggests that FPC plays a crucial role in disengaging from previously rewarded behaviour in order to voluntarily explore alternative options and learn new behavioural strategies.

The exploration of alternative options and the estimation of the potential value of such options are two processes that contribute to learning and decision-making in a volatile environment, and they are functions that have both been attributed to FPC (Boorman et al., 2009; Daw, Niv, & Dayan, 2005). Exploration is necessary to identify which options are being rewarded at any point in time, but it can also be a time consuming process, especially when the number of alternative options is large. Estimating the value of alternative options based on the information already acquired from the previously chosen options can minimize the need for exploration, making task acquisition faster and more efficient.

While different processes, exploration and evaluation are closely related to one another. An agent cannot rely exclusively on internal evaluation, as this requires access to information about at least a subset of potential choices in order to estimate the values of others, and this can only be gathered through exploration. Similarly, estimating which current options offer the best pay-off or whether a previously preferred option has been devalued, promotes exploration of different options. We can therefore see how these two processes are strongly dependent on one another and therefore might not be easy to dissociate.

Our task provides evidence for different exploratory tendencies between FPC-lesion animals and animals with an intact FPC. This is evidence in support a role of FPC in exploratory behaviour. However, this difference was evident only when exploration had to be initiated voluntarily. The voluntary initiation of exploratory behaviour may very well be dependent on the ability of animals to use information about the value of all current available options to determine when to try out a different response than the one that had been previously rewarded. This is consistent with findings by Boorman and colleagues (2009), who showed that human FPC's activation is correlated with the amount of evidence in favour of switching to a different option, rather than being more active during exploration per se. Therefore our results are in line with both the exploratory and evaluative view of FPC function, but cannot provide definite evidence in favour of one over the other, due to the nature of the task.

Dissociation between exploration and evaluation might be possible by applying a paradigm involving a larger number of options: if the number of options is small, the animal can explore all of them before discovering relatively quickly which one is rewarded. On the other hand, trying out all possible options in a larger set would not be the most efficient way of maximizing reward, making alternative option evaluation crucial for rapidly improving performance. If FPC were necessary for exploration, controls might be more likely to select a more varied spectrum of responses, whereas the variety of choices of FPC animals would be more limited, and a tendency to commit more exploitative errors may become more evident. On the other hand, if FPC were necessary for estimating the value of alternative options, we might observe the opposite pattern of results, with control animals showing more rapid improvements in performance combined with a more limited variety of responses, while FPC animals would be heavily reliant on extensive exploration.

If we assume that FPC is crucial for tracking and evaluating alternative options, rather than for exploratory behaviour per-se, it is not surprising that the effects of damage are more evident during the very early stages of learning, when the animal has only had limited amount of time to explore and gather information about new options and therefore any improvement (or disengagement from previously rewarded behaviour) is heavily reliant on re-evaluating previously rewarded options and/or inferring about the value of new options as a result of a reward-prediction error. As exposure to alternatives increases, so does the opportunity to explore and gather

information about these alternatives by trial-and-error, and uncertainty is reduced. Therefore, increasing the number of potential responses while restricting the time allowed for learning should help dissociate evaluation and exploration (and FPC role in either) even further than simply increasing the number of alternatives.

We have seen that once the new rule has been acquired, both groups perform similarly when required to integrate the two rules, and that the impairment in acquisition is very selective to the very early stages of learning. Similar effects of FPC lesions on the very early stages of learning have been observed in a number of recent studies by Buckley and colleagues (in preparation), who also found that control animals show sharp and rapid improvements in performance after the first session on a new task, rule or problem set, when compared to FPC-lesion animals, both during object-in-scene and object-discrimination learning (Buckley et al., in preparation). Our results indicate that these effects also extend to the acquisition of new abstract rules.

A second limitation of the task is that animals might be learning the combination of 'same' and 'smaller than' as a response strategy as early as Stage 6. Stage 8 (which is defined as the stage during which the animals are trained to combine 'smaller than' with the older 'same' rule for the first time) presents animals with two 'smaller than' options, only one of which shares the same identity as the sample. However, during Stages 6 and 7 -which are defined as the stages during which the second rule is acquired- animals are already being trained to pick the item that is both smaller and of

the same identity as the sample, as there is only one 'smaller than' option and this shares the same identity as the sample. This raises the question of whether animals are effectively learning a new simple rule (i.e. 'smaller than') during Stage 6 or already learning to combine 'same' and 'smaller than' and acquiring this as a more complex rule of its own, and therefore can solve the problem in Stage 8 by simply continuing to apply this rule. The correct option in Stage 6 needs to share the same identity as the sample to ensure that the animals do not unlearn 'same'. One way to avoid the unlearning problem could be to give animals 4 options, each of a different size, but with only one being smaller than the sample item. This item could randomly vary in whether or not it matches the sample item: with this procedure, animals would learn to pick the item that was 'smaller than' the sample, regardless of identity, and therefore avoid confounds. Regardless, the impairment in FPC groups seems to be specifically selective to the learning of a new kind of rule, rather than to integration of two rules. Even if the animals were already learning to combine two rules during Stage 6, rather than learning a simple rule to then combine with another simple rule in Stage 8, it is only the very early stage of acquisition that is affected, and this seems to be closely linked to exploratory tendencies.

The option-evaluation framework of FPC function can also help explain findings of FPC activation in humans across a broad spectrum of complex tasks, as well as start to delineate its contribution to conflict tasks.

Both learning and complex tasks require tracking and evaluating multiple options simultaneously in order to acquire a new skill set or during goal-directed behaviour. Therefore, an involvement of FPC in the continuous evaluation of alternatives and, consequently, in promoting exploration of these alternatives, is consistent with findings of higher FPC activation both during learning and with increased task complexity involving relational, abstract integration (Christoff et al., 2001; Kroger et al., 2002; Koechlin et al., 2003; Reynolds, McDermott & Braver, 2006; Badre & Esposito, 2007; Badre & Wagner, 2008), tracking internal states (Koechlin et al., 1999; Reynolds, McDermott & Braver, 2006) and multitasking (Gilbert et al., 2006a). It is also consistent with findings that lesions to FPC in humans impair performance in highly volatile situations (Burgess et al., 2000; Dreher et al., 2008).

Tracking the value of alternative options involves distributing attentional resources across all different options. FPC activity has indeed been found to correlate with the amount of uncertainty between multiple, simultaneously tracked, options (Yoshida & Ishii, 2006). While this is crucial to promote behavioural flexibility that can be highly beneficial to the agent, especially to facilitate learning, allocating attention to many alternatives can be detrimental in circumstances requiring high levels of cognitive control, where alternative options may become distractors. If FPC lesions contribute to making the agent less likely to consider alternative options, it follows that they might be less sensitive to distractors. Resistance to distractors may increase performance in conditions of uncertainty and conflict. Indeed, lesion studies have shown that lesions to

the macaque's FPC enhanced conflict-induced behavioural adaptation, as well as performance during tasks that were interrupted by other tasks or by distractors (Buckley et al., in preparation).

To conclude, FPC lesions were found to affect acquisition of a second abstract rule only when acquisition was dependent on self-initiated exploration. This is consistent both with the view of FPC as a region involved in exploratory behaviour as well as the evaluation of alternative options. Lesions to FPC did not impair the ability of animals to integrate two abstract rules, which contradicts findings of higher FPC activation in humans when tasks require relational integration. However, these findings can be reconciled by considering that complex tasks requiring relational integration might often involve tracking and continuously evaluating many different options, and therefore the role of FPC in complex tasks might be related to processing information about all possible alternatives. This indicates that FPC is a crucial structure for supporting behavioural flexibility, and that it may interact with areas involved in cognitive control to allow the cognitive system to respond optimally to contextual changes in the environment.

GENERAL DISCUSSION

SUMMARY AND IMPLICATIONS OF THE BEHAVIOURAL AND MODELING FINDINGS

DISSOCIABLE COMPONENTS OF COGNITIVE CONTROL

Our behavioural studies (Chapter 3) confirmed that adaptation effects as often defined in the literature (i.e. reductions in the size of the conflict effect following exposure to conflict), which have been suggested to result from the implementation of cognitive control in order to reduce detrimental effects of conflict on performance (Botvinick et al., 2001), are not a unified phenomenon. All of our tasks presented reductions in the size of the conflict effect following exposure to conflict. However, these reductions were driven by either a decrease in speed of response on low-conflict (HL) trials or an increase in speed of response on high-conflict (HH) trials (or both).

Whether the former or the latter most contributed to reductions in the size of the conflict effect seemed to be heavily dependent on whether or not the task allowed the exploitation of a previously acquired automatic response (i.e. a habit) on low-conflict trials. If the task did not allow the exploitation of a habit to respond to low-conflict trials, as in the case of the WCST, reductions were driven exclusively by an increase in speed on high-conflict trials. On the other hand, if the task allowed the exploitation of a

habit, reductions were driven primarily by slower speed of response on low-conflict trials, as in the case of the classic Simon task or the Flanker task.

Testing a version of the Simon task with three items, instead of the classic two-item version, further revealed that adaptation effects can be driven both by a reduction in speed on low-conflict trials, and by an increase in speed on high-conflict trials, within the same task. This may be due to the fact that, within this particular version of the task, competition is created between a habit and an instructed response (i.e. the need to suppress the tendency to respond to the location of the item, in order to implement rule-guided behaviour that requires response to a different location), as well as competition between two instructed responses: once response to the location (i.e. the habit) is suppressed, there are two possible competing options left, both determined by instructions. Therefore, adaptation effects in this version of the Simon task were driven by the mechanisms observed in tasks that allow the exploitation of a habit (like the classic Simon task), as well as the ones observed in tasks where conflict is elicited between two instructed responses (like the WCTS).

I wish to argue that it might be more appropriate to explicitly distinguish between these two types of adaptation effects (reduction in HL speed and increase in HH speed), rather than measuring adaptation as a reduction in the size of the conflict effect, as these effects might point to very different underlying cognitive mechanisms. Cognitive

control could drive adaptation via two mechanisms: increased caution and enhanced attention/task-relevant processing (resulting in decrease of HL speed and increase of HH speed, respectively). These, in turn, might be reflected in the modulation of two routes to responses: a direct route supporting automatic/habitual behaviour and an indirect route supporting rule-guided/instructed behaviour.

Caution and slower responses on HL trials

Stoffels (1996) found that when high- and low-conflict trials are intermixed within the same block, only low-conflict trials are affected, showing a reduction in the speed of response. He hypothesized that this might be due to the fact that, once intermixed, low-conflict trials are responded to by exploiting an indirect route that supports instructed behaviour, which is more advantageous for responding to high-conflict trials. As high-conflict trials normally exploit that route, they would not be affected, whereas low-conflict trials, which normally exploit a direct route supporting automatic, habitual behaviour, would show a decrease in speed of response. Sturmer and colleagues (2002) suggested that a similar mechanism might be at play during trial-to-trial adaptation effects, with the direct route being suppressed following a high-conflict trial, resulting in a reduced conflict effect. Our findings of reductions of speed on HL trials, but no effect on the speed of HH trials, for tasks that involve the exploitation of a habit, allow such a distinction, and are consistent with the idea of the suppression of the direct/habitual route to response.

We were further able to quantify the reduction in the speed of HL trials by applying the LBA model (Brown & Heathcote, 2008) to the response time data. This revealed that the suppression of the direct route could be explained, as predicted, by an increase in caution, which is measured in the model as the difference between response threshold and the start-point of evidence accumulation (see Chapter 4). Increased caution might induce the suppression of habitual responses in context that are not advantageous, like high-conflict trials, and subsequent proactive suppression would result in slower responses on HL trials. When a difference in response time between HH and LH trials was observed, higher caution was also observed on LH trials. However, this was only the case for the task involving habit suppression (the 3-item Simon task), and not for the WCST. Therefore, caution is likely to represent one of the mechanisms underlying adaptation effects, and it may be a particularly prominent feature of adaptation effects specifically in tasks that require the suppression of habitual responses (and of the route to response that supports them), when these are no longer advantageous.

Faster responses on HH trials: attention and enhanced task-relevant processing or practice effects and automaticity?

While our findings regarding caution and suppression of the direct route are fairly conclusive, it has been more difficult to tease apart the mechanisms that might be

driving the increase in speed of response on HH trials. Attempts to describe these effects in terms of higher speed of evidence accumulation (quantified by the drift rate parameter of the LBA model), perhaps driven by enhanced attentional and task-relevant processing mechanisms, have proven inconclusive. While this might be due to some methodological limitations (as discussed in Chapter 4), another possibility is that increases in speed on HH trials are not caused by faster evidence accumulation due to enhanced task-relevant processing or attention, but to other mechanisms.

In our studies, increased speed on HH trials was found only for tasks that created competition between two instructed responses: the WCST and the 3-item Simon task. However, while adaptation effects *restricted* to HH trials were found only in a task that does not allow for the exploitation of a habit to respond on low conflict trials, adaptation effects in classic conflict tasks that do allow for the exploitation of a habit are nonetheless not always exclusively restricted to slower responses on HL trials (as also demonstrated by the 3-item Simon task). While it is possible that faster responses on HH trials in classic tasks that employ a small stimulus set could be partly explained by repetition priming, many studies have shown increases in speed on HH trials, even when controlling for exact stimulus repetitions (e.g. Egner, Delano, & Hirsch, 2007; Egner & Hirsch, 2005; Kerns et al., 2004; Kerns, 2006; Ullsperger et al., 2005), suggesting that priming is not the only explanation for this phenomenon. An alternative hypothesis for this type of adaptation effects could be represented by a compromise between priming

and enhanced attentional mechanisms, in the form of practice effects on a micro-scale level.

The effects of practice on task-performance and skill acquisition, and how they can lead to the development of habits and automaticity, have been extensively investigated (see Boettiger & D'Esposito, 2005; Cohen, Dunbar, & McClelland, 1990; Dickinson, 1985; Jenkins, Brooks, Frackowiak, & Passingham, 1994; Kornblum, Hasbroucq, & Osman, 1990; Lungu, Binenstock, Pline, Yeaton, & Carey, 2007; Ramnani & Passingham, 2001; Shiffrin & Schneider, 1977). Generally speaking, practice leads to increased accuracy and speed, as effortful goal-directed behaviour (which, during the early stages of learning, requires strong recruitment of cognitive control) gradually becomes effortless and automatic. Increases in speed and accuracy on high-conflict trials have also been reported, for example when the frequency of conflict trials is increased (e.g. Hommel, 1994; Logan, 1979; Nieuwenhuis et al., 2003; Sturmer et al., 2002), or as the participant becomes familiar with the task or is overtrained (Milham et al., 2003; Heink et al., 1997). Adaptation effects on HH trials, such as the ones observed in the WCST or 3-item version of the Simon task in this thesis, are also defined by similar increases in speed. It is therefore possible that conflict-induced behavioural adjustments resulting in faster responses on high-conflict trials might be practice effects manifesting themselves at different levels: at the microscale level for trial-to-trial adaptation effects (faster responses on HH trials) and at the macroscale level for frequency effects that affect overall task performance (faster responses on high-conflict trials overall).

The findings of the absence of modulation of the conflict effect as a consequence of changes in the frequency of conflict in the proportional version of our WCST (see Chapter 3) could be consistent with this hypothesis, and possibly indicate that it is not exposure to conflict per-se that induces behavioural adaptation, but rather the possibility to develop a new response strategy to reduce the effects of conflict.

In the WCST, increasing the percentage of high-conflict trials effectively only increases exposure to conflict, but does not provide the participant with more practice with one particular rule over the other. Participants are presented more frequently with trials where they need to select between two competing items (high-conflict) than with trials where they are presented with two identical items (low-conflict). However, the two competing responses are still rewarded with equal frequency: the criterion for rule-switch remains the same regardless of the proportion of high-conflict trials, and 'match-by-shape' is just as likely to be rewarded as 'match-by-colour' in the overall task. Therefore, when the frequency of high-conflict trials is increased, the cognitive system should heavily favour the indirect route (which supports rule-guided behaviour, as this is the most advantageous route). Nonetheless, it should not favour one rule over the other, as they are both equally as likely to be advantageous at different points during the task.

Results from this experiment showed that the conflict effect is not affected by increasing the frequency of high-conflict trials, suggesting that even when the reliance on the indirect route, or the need for cognitive control, is increased, high-conflict trials still incur response-speed costs. The crucial difference between the WCST and classic conflict tasks is that, in the latter, increasing the frequency of high-conflict trials might encourage the development of a new, perhaps temporary, habit. This may explain why increasing the frequency of high conflict trials often leads to a reversal of the conflict effect, where responses to high-conflict trials become faster than low-conflict trials: the most rewarded behaviour would be recognized as more advantageous and therefore repeatedly exploited through the task, possibly becoming temporarily automatized. It would be interesting to investigate whether inducing the development of a habit in the WCST (e.g. by rewarding 'match-by-colour' more often than 'match-by-shape'), we would observe a somewhat similar reversal in the conflict effect. This hypothesis is still consistent with the routes-to-response account (Stoffels, 1996; Sturmer et al., 2002), but, rather than reversal of the conflict effect in response to frequent conflict being explained in terms of sustained suppression of the direct route and enhanced biasing of the indirect route, it could be explained as the encoding (or transfer) of a new behavioural strategy from the indirect to direct route. Development of automaticity for a new advantageous response might initially rely, more or less exclusively, on the indirect route. As the behaviour becomes automatized through practice, it might rely less and less on the indirect route and more on the direct route.

If the effects of conflict frequency on the speed of response to high- and low-conflict trials represent macroscale types of practice/adaptation effects, trial-to-trial adaptation effects that result in increased speed on HH trials might also be explained as a similar type of mechanism on a microscale level. When high- and low-conflict trials are intermixed in the same block, the first high-conflict trial after a low-conflict trial (LH) represents a novel and uncertain situation, where the previously rewarded behaviour is no longer advantageous and needs to be suppressed in favour of a different response. The same effortful processes that are generally used to support the encoding of a new automatic response during the early stages of practice (and which are possibly reliant on the indirect route to response) might then be engaged in order to exploit the most advantageous response strategy from one trial to the next. The second high-conflict trial in a sequence (HH) will not present the same degree of novelty and uncertainty as the first trial in a sequence, but might benefit from having practiced the new advantageous response (or response strategy) on the previous (LH) trial, therefore resulting in faster and more accurate responses. Evidence has shown that the first high-conflict trial in a series is slower as a function of the number of preceding low-conflict trials, and high-conflict trials become increasingly faster as a function of the number of preceding high-conflict trials (Durston, 2003), suggesting that performance on high-conflict trials gradually improves in a manner similar to that observed as a result of increased automaticity of a skill, or of overtraining (Cohen et al., 1990; Milham et al., 2003; Poldrack et al., 2005).

This might explain why increasing the frequency of high-conflict trials does not affect the overall conflict effect on the WCST, but does abolish the adaptation effect. While increasing the number of high-conflict trials does not create a context where one rule could be automatized over the other in the overall task (as previously explained), it does encourages exploitation of one rule over the other during the blocks between one rule-switch and the other: the number of trials where the currently rewarded rule needs to be applied in order to perform a correct response (high-conflict) is higher than that of trials where the rule is not necessary (low-conflict). Repeated application of the currently rewarded rule might cause it to become automatized and increase speed of response on LH trials within the block, reducing the difference between LH and HH trials and cancelling out the adaptation effect. In a task with an equal number of high- and low-conflict trials, the necessity for repeated application of the currently rewarded rule is restricted to sequences of high-conflict trials (as low-conflict trials can be resolved by applying either rule, or no rule at all), and therefore could explain the faster responses on HH trials. The finding of faster HH trials in tasks where multiple arbitrary stimulus-response associations determine response, such as the Simon task, suggests that this phenomenon is unlikely to represent only the exploitation of an abstract rule to provide a short-cut to response.

If adaptation effects due to increase in HH speed were due to practice effects, it could also help explain why we were unable to quantify these effects within the drift rate parameter of the LBA model. To the best of my knowledge, the LBA has only once

been applied to studying practice effects. In this application, Heathcote and Hayes (2012) found that practice effects seemed to modulate different parameters in different participants, with 3 out of 4 participants showing no effects of practice on the drift rate. Therefore, if increased speed of response on HH trials is due to practice effects on a microscale level, it is unlikely to be reflected in the drift rate, at least according to this application. However, this was a technical paper whose primary goal was that of comparing the conclusions drawn from the application of the LBA and the Ratcliff Diffusion model, and it was carried out on a very small sample. Further investigation would need to be carried out in order to test this claim.

Default route, advantageous and disadvantageous responses

One of the main questions regarding the routes-to-response framework of Sturmer and colleagues (2002) concerned the issue of whether the cognitive system has a 'default' mode, i.e. it always starts off with suppression of the direct route (therefore making the indirect route a default route to response) or with the direct route free from suppression (therefore making the direct route the default route to response). Evidence from our battery of behavioural tasks (Chapter 3) and from our TMS study (see Chapter 2) seems to suggest that the presence of an a-priori default route to response is unlikely. A more realistic scenario might be that the default route is more flexible, and represents the route that is the most advantageous given the current macroscale context (for example, due to task demands or conflict frequency). If a task elicits

competition only between two instructed responses, rather than between a habit and an instructed response, such as in the case of the WCST analogue, the direct route would be unlikely to prove advantageous, and therefore the indirect route might be selected as default. Viceversa, in the case of most classic conflict tasks, which are designed to elicit conflict between a habit and an instructed response, if the frequency of low-conflict trials is high or equal to that of high-conflict trials, then the direct route might be selected as default, as it would prove advantageous in a substantial number of trials. If low-conflict trials are rare, on the other hand, the direct route would not be advantageous on most trials, and therefore the indirect route may be selected as a default. This hypothesis is consistent with Miller and Cohen's (2001) view that, when there are no task demands, the stronger connections (i.e. the default route) between a stimulus and a response would generally prevail over weaker ones. However, as they pointed out, connections can be strengthened to the point of automaticity. Therefore task demands themselves can actively modify the strength or weakness of these connections, and therefore determine what route will prevail (i.e. default), even if only temporarily within a specific context.

Furthermore, as illustrated early, with overtraining or practice, a particularly advantageous response might initially be encoded and supported by the indirect route and, with time, become less reliant on the indirect route and more reliant on the direct route. In classic conflict tasks that are designed by creating competition between a habit an instructed responses, when the proportion of high- and low-conflict trials is equal,

the exposure to high-conflict trials might not be sufficient to automatize instructed-behaviour as the most advantageous response strategy overall. Low-conflict trials are likely to continue exploiting the direct route, as they tend to rely on habits that have already been automatized, since they are generally advantageous in most every-day life contexts (such as reading words instead of naming the colour they are printed in). When the proportion of high-conflict trials is higher, or when a longer sequence of high-conflict trials is presented, on the other hand, the response strategy that is most advantageous will have more of a chance to automatize, and could result in the elimination or even the reversal of the classic conflict effects. This indicates that the automatization of a response might actually be a relatively fast process, under certain conditions. Therefore, the 'default' route (i.e. the most advantageous route) might change at different stages in the task: the indirect route will be exploited when the response has yet to be automatized, and the direct route will be exploited as the response becomes more automatic.

CONCLUSIVE REMARKS ON THE THEORETICAL FRAMEWORK

To summarize our behavioural findings, consistent with the CMCC model (Botvinick et al., 2001), the detection of conflict has been found to drive reductions in the size of the conflict effects in all tasks, regardless of whether conflict is elicited between a habit and an instructed response or between two instructed responses. However, we were able to dissociate two types of conflict-induced behavioural

adaptation that contribute to the reduction of the conflict effect: a decrease in speed on HL trials and an increase in speed on HH trials. The former is likely to be due to increases in caution, which can be quantified by LBA model parameters, which, in turn, may drive the suppression of the direct route supporting habitual/automatic behaviour (Hommel et al., 2004; Sturmer et al., 2002) and it appears to contribute to the reduction of the conflict effect only in tasks where the exploitation of a pre-existing habit is possible. The mechanisms driving the latter are still unclear, but one possibility is that they might represent microscale practice effects that result from repeated exploitation of a new advantageous response (or response strategy).

These microscale and macroscale adjustments, driven by context and continuous fluctuations in environmental contingencies, are an example of the interplay between cognitive control and flexibility that is a fundamental feature of the cognitive system. I would like to argue that it is important to distinguish between the two types of adaptation mechanisms that seem to contribute to the reduction of conflict effects. If we are interested in looking at the neurophysiological substrates of cognitive control and flexibility, it may help clarify whether specific areas are involved in one mechanism over the other. As we will see in the next section of the discussion, some of our evidence could already indicate the role of different prefrontal regions in supporting different aspects of adaptation. Furthermore, if these two distinct mechanisms can be quantifiable by model parameters, as it seems to be the case for caution driving adaptation on HL trials, we could have access to a way to directly correlate the

recruitment of specific mechanisms (and their effects on response times) with the activity in specific brain areas. As we will see in the next section, some evidence already points to a correlation between caution, as quantified by the LBA model, and brain activity in regions also involved in conflict-related processing, such as the ACC.

SUMMARY AND IMPLICATIONS OF THE INVESTIGATION INTO THE NEURAL SUBSTRATES OF COGNITIVE CONTROL AND FLEXIBILITY

ACC AND MONITORING FUNCTION

Results from our neuropsychological investigation into the effects of ACC damage on the conflict analogue of the WCST (Chapter 1) revealed that patients with ACC lesions showed neither difference in the speed of response to conflict trials nor any impairment in measures of adaptation, compared to healthy controls. This seems to indicate that human ACC is not crucial for conflict-processing or adaptation when the task involves conflict between two instructed responses. While the findings of no impairments in the ability to respond to high-conflict trials is fairly common in ACC patients (e.g. Fellows & Farah, 2005; Sheth et al., 2012; Stuss et al., 2001; Swick & Jovanovic, 2002; Vendrell et al., 1995), some studies have reported impairments in adaptation following ACC lesion (diPellegrino et al., 2007; Sheth et al., 2012). Sheth and

colleagues (2012), in particular, provided very convincing evidence for the necessity of an intact ACC for adaptation, as their study involved targeted surgical removal of a part of ACC that was shown to present cells that were responsive to adaptation trials. However, our study, as well as some of the aforementioned ones, does not agree with this conclusion.

Reconciling findings in neuropsychology

One possible explanation concerns whether damage extends into area 32' of the ACC, which has been proposed as the area responsible for conflict-monitoring in humans (Cole et al., 2009). In our neuropsychological investigation, only a limited amount of the lesions encroached onto area 32'. Therefore, one explanation for the discrepancy might be related to the fact that most of the damage affected subdivisions of the human ACC that are not recruited for processing conflict. Our pilot study investigating the effects of intracranial ACC stimulation in the WCST (Chapter 5) seemed to indicate that stimulation to more caudal areas than the ones in our neuropsychological study might have an effect on adaptation in this task. Our LFP analysis also indicated that there might be very fine-grained conflict-processing subdivisions within ACC, as electrodes located slightly more posterior and more dorsal in one patient than another appeared to record conflict-related activity, whereas the more anterior electrodes did not. Nonetheless, other studies investigating the effects of fairly

substantial lesions to caudal and dorsal human ACC on conflict-monitoring failed to find an effect of ACC damage on adaptation (e.g. Fellows & Farah, 2005).

The discrepancies between results in the neuropsychological literature could also be explained in terms of the nature of the tasks used, and how they might relate to the two different types of adaptation effects that can be observed (i.e. increases in HH speed or decrements in HL speed), and the mechanisms that support them (i.e. caution versus attention or enhanced processing of task-relevant information). Conflict in the WCST is elicited between two competing instructed responses, and adaptation in this task is due to an increase in speed of HH trials. In classic conflict tasks, where conflict is elicited between a habit and an instructed response, adaptation is likely to be driven also (or mostly) by a decrement in speed of HL trials. As described in the previous section of the general discussion, these different types of adaptation could be reliant on different underlying mechanisms that support cognitive flexibility via the interplay between a direct and an indirect route to response. Furthermore, the modulation of these routes might be sensitive to contextual contingencies that can determine which response strategy is more advantageous at a specific point in time, as well as promote automatization of the most advantageous response. Consequently, whether or not ACC damage affects adaptation mechanisms could be heavily dependent on methodological differences, such as whether slight variations in the experimental design influence the interplay between the competing responses and the automatization of response strategies.

For example, Stuss and colleagues (2001) found no effects of ACC damage on performance on the incongruent trials of the Stroop task. However, they only administered patients with the incongruent version of the task. If participants are only exposed to incongruent trials, they might be either exploiting the indirect route as the default route to response, or be encoding colour-naming (as opposed to word-reading) as a temporary automatic response, as this is the only response that is rewarded during the task. Therefore, the competition between the two routes could be significantly reduced in this context.

Fellows and Farah (2005) found no effect of ACC damage on the ability of participants to adjust their behaviour depending on the percentage of high-conflict trials, also in a Stroop task. However, they administered participants with the mostly high-conflict block first, and only later with the mostly low-conflict block. This might have encouraged participants to automatize the colour-naming response during the first block, therefore reducing the effects of competition between the habitual word-reading response and the instructed colour-naming response in the subsequent block. Indeed, studies that have investigated the effects of practice on task performance found that, even though behavioural measures of conflict were constant across different stages of training, ACC activity decreased as the task became over-practiced (Erickson et al., 2004), often quite rapidly (Milham et al., 2003), as the participants became more and more conditioned to respond to high-conflict trials.

Therefore, it seems plausible that, if competition is elicited only between instructed responses, or in contexts where exploitation of the indirect route is generally advantageous (as in the case of our study or Stuss and colleagues' study), or in contexts where the automatization of an instructed response can reduce the competition between the response and a previously acquired habit (as in the case of Fellows & Farah's study), ACC damage does not affect conflict processing nor adaptation. However, if the context allows the exploitation of a habit to facilitate performance on low-conflict trials, and conflict is elicited between the habit and an instructed response, ACC damage could affect adaptation (diPellegrino et al., 2007; Sheth et al., 2012).

This suggests that the ACC may only be crucial for performance in contexts where competition is elicited between an arbitrary instructed response (or a response that is rarely advantageous), and a habitual response (or one that would be regarded as more frequently advantageous). Recent evidence has indeed indicated that ACC activity is related to the relative advantage of switching away from the best long-term option when faced with a choice at a specific point in time (Boorman, Rushworth, & Behrens, 2013). A habitual response could represent a type of long-term, default, option that needs to be switched away from in favour of a different response that is advantageous specifically within the current demands of the task or even the specific demands of a trial. Inability to assess the advantage of switching away from such an option could result in impairments in adaptation.

If this were the case, some further questions could be explored in future research. Why is ACC crucial for adaptation only when competition is elicited between more frequently advantageous (generally habitual) and less frequently advantageous (generally instructed) responses? Furthermore, we tentatively established that, in tasks that elicit competition between a habit and an instructed response, adaptation effects are often driven by reduction in the speed of HL trials (Chapter 3), which can be quantified as an increase in caution in the LBA model (Chapter 4). Therefore, could the role of the ACC in conflict tasks be better defined in terms of a region that modulates caution during behavioural adaptation?

Possible alternative ACC functions during conflict tasks

Besides conflict, ACC has been found to be involved in situations of environmental volatility (e.g. Behrens, Woolrich, Walton, & Rushworth, 2007; Ridderinkhof & Ullsperger, 2004; Rushworth & Behrens, 2008), during cognitively demanding tasks (e.g. Davis, Hutchison, Lozano, Tasker, & Dostrovsky, 2000; Davis & Taylor, 2005; Silton et al., 2010) or when faced with novel, surprising, unexpected outcomes or rare events (e.g. Badre & Wagner, 2004; Braver, Barch, Gray, Molfese, & Snyder, 2001; Hayden, Heilbronner, Pearson, & Platt, 2011; Ito et al., 2003; Nee, Kastner, & Brown, 2010; Wilk, Ezekiel, & Morton, 2012). All four of these contexts, as well as conflict, are situations that increase the uncertainty regarding the appropriate

response. This makes feedback processing particularly important in order to determine the appropriate course of action.

Uncertainty, errors and caution

Uncertainty increases the likelihood of errors (i.e. negative feedback), which is an event ACC has also been shown to be responsive to (e.g. Braver et al., 2001; Hayden et al., 2011; Holroyd et al., 2004; Ito et al., 2003; Lapish, Durstewitz, Chandler, & Seamans, 2008; Nee et al., 2010; Wang, Ulbert, Schomer, Marinkovic, & Halgren, 2005; West & Travers, 2008).

The Error Related Negativity (ERN) is an event-related potential that occurs after an error has been committed. It has been attributed a generator in the medial-prefrontal areas that include the ACC (Bartholow et al., 2005; Keil, Weisz, Paul-Jordanov, & Wienbruch, 2010; van Veen & Carter, 2002; West & Travers, 2008), and has been found to be attenuated by ACC damage (Swick & Jovanovic, 2002; Swick & Turken, 2002; Yeung & Cohen, 2006). The ERN appears to increase in amplitude during high-conflict trials (Swick & Turken, 2002; van Veen & Carter, 2002) and has been suggested to represent post-response conflict on error trials by the proponents of CMCC model (Yeung, Botvinick, & Cohen, 2004). Some evidence, however, seems to indicate that it does not to correlate with conflict as defined by the CMCC model, i.e. in terms of the

amount of competition between two concurrently active responses, or Hopfield energy (Burle, Roger, Allain, Vidal, & Hasbroucq, 2008; Carbonnell & Falkenstein, 2006).

A related, alternative proposed function of ACC, the processing of increased error likelihood (Brown & Braver, 2008), could help reconcile these findings. Although ERN amplitude might not be determined by conflict defined as response-competition, it might be determined by the increased error likelihood on high-conflict trials when choice is between a habit and an instructed response. If a pre-existing habit (e.g. reading a word) is highly conditioned as a behaviour that would lead to a positive outcome in most contexts, it may be more likely to lead to an error in a context where it is no longer advantageous, such as high-conflict trials in most classic conflict tasks.

High error-likelihood would also lead to increased caution. As per the previously discussed hypothesis, slow responses on LH trials and HL trials during conflict tasks that create conflict between a habit and an instructed response might be related to an increase in caution, which is quantifiable by LBA parameters. Activity in the ACC has been found to correlate with changes in the amount of caution as quantified by the LBA model, during speed-accuracy tradeoffs (van Maanen et al., 2011). Sheth and colleagues (2012) found that the activity of human ACC neurons was correlated with the changes in conflict levels from one trial to the next. They suggested that, in situations where cognitive demands remain constant, ACC signalling facilitates responses, whereas in situations involving rapidly changing demands, it promotes accuracy by slowing down

responses. Thus, ACC seems to be support adaptive processes akin to speed-accuracy tradeoffs. ACC damage has also been correlated with impaired post-error slowing (Tanaka, Mansouri & Buckley, in preparation), which is another context where caution may be increased in order to reduce the likelihood of future errors.

Uncertainty, reinforcement history and counterfactual prediction errors?

However, the ACC has been found to be sensitive not only to negative feedback (i.e. errors), but also positive feedback (i.e. rewards). In conditions of uncertainty, it is also particularly difficult for the agent to determine which action is likely to lead to a reward, and efficient processing of information pertaining to rewards, when obtained, is crucial to identify new correct responses.

The ACC appears to specifically encode associations between actions and rewards (Rudebeck et al., 2008). Its neurons present a firing profile that closely resembles that of dopaminergic neurons (Niv, Duff, & Dayan, 2005; Schultz, 2002, 2006; Smith, Li, Becker, & Kapur, 2006), which have the highest number of terminations in the cortex within the ACC (Paus, 2001). ACC cells, just like dopamine cells, are sensitive to reward. They are particularly responsive to unexpected rewards during early stages of conditioning, with responsiveness to reward declining through time (Ito et al., 2003; Matsumoto, Matsumoto, Abe, & Tanaka, 2007) and increasing in response to the absence of an expected reward (Ito et al., 2003) or to errors (Emeric et al., 2008; Ito et

al., 2003; Nakamura et al., 2005), effectively performing a similar type of temporal-difference learning to that observed in dopamine neurons (see Suri & Schultz, 1999), and encoding reward-prediction errors (Kennerley, Behrens & Wallis, 2011).

Consequently, the ACC has been suggested to play a more general role in reward-driven learning and behaviour and in evaluating the outcome of actions, rather than in behavioural adjustments that are specifically made in response to errors (Bush et al., 2002; Kennerley & Wallis, 2009; Kennerley et al., 2006; Nishijo et al., 1997; Walton et al., 2007). For example, Kennerley and colleagues (2006) showed that ACC lesions in the macaque did not impair the animals' ability to adjust their behaviour on trials immediately following an error. Rather, when faced with a reward-prediction error following a previously advantageous response, animals were unable to optimally use information about sustained recent positive feedback to identify new advantageous responses. The results of our LFP analysis (Chapter 5) indicated that enhancement of ACC's beta oscillations during the positive feedback period were modulated by the level of conflict in the trial, and this, in turn, appeared to modulate the degree of anticipatory beta suppression on subsequent trials. Cohen and colleagues (2008) and Sheth and colleagues (2012) also reported ACC activity during the positive feedback period following a response during conflict tasks. This is all consistent with a role for the ACC in processing feedback, both positive and negative, in order to proactively modulate conflict-induced behavioural adaptation.

Given the suggested role for the ACC in exploiting recent reinforcement history in order to drive behavioural adaptation towards new (or less frequently advantageous) options, its involvement in situations where error likelihood is high and in behavioural adjustments that help minimizing the occurrence of errors, and its response profile that resembles the reward-prediction behaviour of dopamine neurons, one related hypothesis that could be investigated is whether, during conflict tasks, the ACC is involved in the generation of counterfactual reward-prediction errors. One possibility is that the ACC internally simulates whether an action that, based on reinforcement history, is generally likely to lead to a reward (such as a habit) would generate a reward-prediction error given the current context (such as a high-conflict trial), and, if it does, help adjust behaviour accordingly, by switching away from habitual responses towards new behavioural strategies. The ACC might therefore play a similar role to the Confirmation Function suggested by Kornblum (1990), where an action that is about to be performed needs to be confirmed in order to test whether it will lead to an error, and changed if necessary.

Some authors have suggested that the ERN could represent a temporal-difference reward-prediction error signal originating in the ACC (Burle et al., 2008), which lasts until remediation of the error takes place (Holroyd, Yeung, Coles, & Cohen, 2005). The N2, an ACC-generated component that has been suggested as the equivalent of the ERN on correct high-conflict trials (van Veen & Carter, 2002; Yeung et al., 2004) could represent the neural signal of this 'simulated' reward prediction error. This

component has been found to reverse when conflict frequency is increased, with larger amplitude for low-conflict trials than for high conflict trials (Nieuwenhuis et al., 2003). ACC BOLD signal showed a similar pattern of reversal in a study by Wilk and colleagues (2012), and response times have also been shown to reverse when the percentage of high-conflict trial is high (e.g. Hommel, 1994; Logan, 1979; Nieuwenhuis et al., 2003; Sturmer et al., 2002). As we hypothesized earlier, this could be a context where the automatization of the instructed response is possible, and therefore would represent the go-to response for the task overall, leading to a possible reward-prediction error on low-conflict trials.

Counterfactual ability, the ability to internally simulate and compare alternative scenarios, is a fundamental property of the human cognitive system, has been also attributed to the other two prefrontal areas which are often active in conjunction with ACC in conflict tasks, dlPFC (see Knight et al., 1995) and FPC (Boorman et al., 2011). Some data has begun to emerge regarding the encoding of fictive rewards in the monkey's ACC neurons (Hayden, Pearson, & Platt, 2009), suggesting that this area might be responsive to 'virtually generated' reward signals. Whether similar type of information is encoded in the human brain, and how it could be exploited to guide the estimation of counterfactual prediction-errors during conflict tasks, is a topic that could be investigated further.

In conclusion, while the ACC may carry out a monitoring function, this is likely to be a more general and flexible monitoring function concerned with changes in environmental contingencies and feedback processing in situations of uncertainty regarding the appropriate response, which require adjustments in behaviour in order to exploit the most advantageous response given the current context, particularly with the aim to minimize errors from action selection, rather than to strictly perform conflict detection and/or processing.

DLPFC SUPPORTS BEHAVIOURAL ADAPTATION, BUT WHAT ARE THE MECHANISMS?

The ACC projects to dlPFC via the cingulate fasciculus and dlPFC projects back to the ACC, directly onto cingulate motor areas in the caudal parts of ACC (Paus, 2001). These are approximately the areas we recorded from in our DBS study (Chapter 5) and that, when stimulated during the WCST, appear to induce similar types of effects as TMS to more superior dlPFC regions (Chapter 2). We hypothesized that the role of ACC might be better defined as a structure that responds to changes in environmental contingencies that increase the uncertainty of obtaining a reward by exploiting a generally advantageous response, and that it might modulate the ability of an agent to switch away from such default option by processing actual (as well as, possibly, counterfactual) feedback. This information is then likely made available to dlPFC via the cingulate fasciculus. dlPFC is implicated in a range of functions (temporal, representational and control functions – see General Introduction) that could exploit

ACC signalling in order to guide activity-flow along pathways that will ultimately result in the selection of contextually appropriate motor-outputs, whose outcome will then be processed and encoded by the ACC. Therefore, the network of bidirectional projections between ACC and dlPFC represents an ideal candidate to support the type of behavioural adjustment required to maximize the outcome of actions in a changing environment. However, the degree to which these two regions interact, as well as the need for each or both of these structures to be intact, for modulating goal-directed behaviour might be context-dependent.

As previously illustrated, ACC might be crucial only in contexts where adaptation is driven by an increase in caution in order to suppress a habitual response that is likely to lead to a negative outcome in that context. The dlPFC has often been implicated in the inhibition of task-irrelevant information and habits (Jahanshahi et al., 1998; Kadota et al., 2010; Knight, Staines, Swick, & Chao, 1999; Knoch, Brugger, & Regard, 2005; Mitchell, Rhodes, Pine, & Blair, 2008; Treichler, 1973; see also General Introduction). This is consistent with a role of the dlPFC in suppressing the direct route in contexts where it is no longer advantageous, resulting in the reduction of conflict effects in conflict tasks that rely on the exploitation of an habitual response (i.e. where the direct route is default). TMS to the dlPFC has been shown to impair adaptation in the Simon task, affecting the reduction in conflict effects normally observed after exposure to conflict (Sturmer et al., 2007). Findings from out neuropsychological (Chapter 1) and TMS (Chapter 2) studies confirmed the need of an intact and functioning dlPFC in order

to perform conflict-induced behavioural adaptation. Both damage and interference with activity in this area resulted in a reduced adaptation effects in the WCST, unlike ACC lesions. The findings that an intact ACC is not crucial to adaptation in this task, while dlPFC is, might be related to different types of modulations that could contribute to adaptation.

We saw how reductions in the size of conflict effects are also observed in tasks that do not involve the need to suppress the direct route (such as the WCST - see Chapters 3 and 4 - or tasks that create competition between neutral trials, as in Chen & Melara's study (2009). Unlike the ACC, an intact dlPFC appears to be necessary for adaptation effects regardless of whether the task requires suppression of the direct route in favour of enhanced indirect route processing or not.

It is unlikely that the indirect route would ever need to be suppressed on low-conflict trials during the WCST. No habitual response can be exploited to speed up response time in this task, and therefore there should be no competition between the two routes that requires suppression of one over the other. Rather, conflict or competition in this task is elicited within the same route, the indirect route. This may indicate that the role of dlPFC is to specifically suppress irrelevant information or competing responses even when they share the same route to response. However, if this function were compromised, we would expect high-conflict trials to be affected in a similar manner, regardless of whether they follow a low-conflict or a high-conflict trial.

Instead, evidence from our neuropsychological (Chapter 1) and TMS (Chapter 2) studies seems to indicate that compromising dlPFC function might modulate the responses on high-conflict trials slightly differently depending on the nature of the previous trial. The adaptation effects are abolished both by a decrease in speed of response on HH trials (consistent with the idea that dlPFC modulates suppression of irrelevant responses in order to aid adaptation) but might also be due to an increase in speed of response on LH trials (which is inconsistent with this idea). This indicates that recent conflict-history and changes in the level of conflict have an effect on dlPFC-driven adaptation effects. Indeed, increased dlPFC activity has been reported to be higher on high-conflict trials after a sequence of low-conflict trials (Durston, 2003), and for the first pair of high-conflict trials (Durston, 2003; Kerns, 2006; Kerns et al., 2004; Kim, Chung, & Kim, 2012) but to then gradually decline as the number of consecutive high-conflict trials increases (Durston, 2003), despite the fact that all these contexts require suppression of irrelevant responses.

We hypothesized that this might be due to a possible role of dlPFC in maintaining information about recent conflict history online. If dlPFC requires a signal from the ACC in order to trigger inhibition of disadvantageous responses, we would expect lesions to the ACC to cause impairments in behavioural adjustment. However, as we have seen in Chapter 1, and as some other neuropsychological studies have shown (e.g. Fellows & Farah, 2005), this is not always the case. This suggests that dlPFC can process conflict-related information independently of ACC. Mansouri and colleagues

(2007) found that cells in dlPFC might encode information about the level of conflict in the previous trial and several studies have shown activity in dlPFC that very closely resembles ACC activation during conflict tasks, albeit on a slightly lagged timeframe (Durston, 2003; Egner & Hirsch, 2005a; Kerns, 2006; Kerns et al., 2004; Kim et al., 2012). Some studies have shown that ACC and dlPFC activation is modulated by informative cues that warn about the level of conflict to be expected on the subsequent trial. During the target period, ACC activation decreases when the cue is informative of the level of conflict, regardless of the level of conflict experienced on the trial, whereas dlPFC remains active for high conflict trials. When the cue is uninformative, activation commonly observed in conflict tasks is observed (Luks et al., 2006, Aarts et al., 2008). This suggests that dlPFC does encode conflict-related information independently from ACC signalling.

If the dlPFC creates and maintains a memory trace for recent conflict levels in working-memory, disruption of this process could cause the cognitive system to reset onto the default route to response for the current context. In the case of most conflict tasks, this would be the direct/habitual route to response, and would result in an abolition of the adaptation effect due to the release of the direct route in contexts where it would normally be suppressed (i.e. HL or LH trials). As the direct route is never advantageous or exploited in the WCST, the indirect route that supports rule-guided behaviour would serve as the default route, and the loss of the memory trace would reset the cognitive system onto this route. This could explain why faster responses on

LH trials are sometimes observed as part of the reduction in the adaptation effects when stimulation is administered: loss of the conflict-history memory trace would reset the system to the default route for the task, in this case, the indirect route that supports rule-guided behaviour, which would be advantageous on LH trials. At the same time, loss of a memory trace indicating the presence of conflict on a previous trial might reduce the proactive enhancement of further indirect route biasing that would normally result in faster responses on HH trials, causing them to drop to LH levels.

However, one issue with this hypothesis is that stimulation at any point during the intertrial interval should have the same effect on performance, since it would cause a loss of the memory trace. Sturmer and colleagues (2007), instead, found that only TMS during a time-window closer to stimulus onset disrupted performance. One possibility is that the effects of stimulation at an earlier time-point might dissipate and allow for the reinstatement of the memory trace, or reinstatement of default route suppression, by the time of new trial onset, perhaps aided by ACC signalling. A further possibility is that dlPFC might be involved in supporting automatization processes during the early stages of adjustment from one type of response to another.

We saw earlier how adaptation effects in the WCST are due exclusively to faster speed of response on HH trials, and how this could represent microscale practice effects that may constitute the early stages of automatization of a response (in the case of the WCST, the currently rewarded rule). Lungu and colleagues (Lungu et al., 2007) provided

evidence that the ACC/dlPFC network involved in conflict-monitoring and cognitive control displays changes in its patterns of activation that are modulated by practice and improvements of performance during a continuous task. They asked subjects to track a sine wave on a screen by moving a cursor so that upward finger movements moved the cursor upward and downward finger movements moved the cursor downward (compatible condition) or viceversa (incompatible condition). They found that both ACC and dlPFC activation decreased during later stages of training, accompanied by an increase in subcortical activation (specifically in the caudate nucleus). The changes in activation correlated with the participants' improvements in performance during the incompatible condition, until it became statistically indistinguishable from the compatible condition. More specifically, improvements in performance on their task were restricted to faster tracking during the incompatible condition (therefore similar to the effects observed during the WCTS, where HH trials increase in speed), and not a decrease in speed on the compatible condition (therefore unlike the decrease in HL speed observed in conflict tasks involving the suppression of a habit).

This pattern of decreased ACC and dlPFC activation was also observed during sequences of high-conflict trials (Durston, 2003). Furthermore, Wilk and colleagues (2012), found that increasing the frequency of high-conflict trials gradually decreased both ACC and dlPFC activity for this type of trials. As previously discussed, this is a scenario during which it might be possible for the instructed behaviour, which becomes more advantageous as the number of high-conflict trials is increased, to become

automatized as a new, temporary habit. Finally, in our study, conflict was elicited between two instructed responses ('match-by-colour' or 'match-by-shape'), in a context where the direct route to response would not prove advantageous, as no habit can be exploited to facilitate performance on low-conflict trials.

Decreased prefrontal activation, also accompanied by modulations in striatal activity, is often observed during the process of automatization of a motor skill and improvements in performance (Boettiger & D'Esposito, 2005; Doyon & Benali, 2005; Doyon, Penhune, & Ungerleider, 2003; Poldrack et al., 2005), and is thought to underpin practice effects. The similarities in the patterns of activation during adaptation and automatization of a novel action set would be consistent with the idea that an instructed response can become automatized, possibly reducing uncertainty, likelihood of errors (and counterfactual prediction errors) and, consequently, reducing activity in the ACC. Therefore, one possibility is that dlPFC could be supporting the microscale practice effects underlying increases in speed from LH to HH trials. This would not necessarily be affected by interference with dlPFC activity during the inter-trial period, as practice effects would only be evident as the action is actually being performed.

Taking the findings from Chapters 1 and 2, which looked at the effects of lesions and TMS of the dlPFC on adaptation during the conflict version of the WCST, and integrating them with the findings from Chapters 3 and 4, which looked at the possible dynamics underlying different types of adaptation, as well as with the literature on the

proposed mechanisms of dlPFC function in conflict, we can conclude that dlPFC is certainly a crucial structure to support behavioural adjustments and cognitive control, but that it may do so via three possible complementary mechanisms. First, suppression and inhibition of irrelevant information or competing responses, regardless of whether competition is between routes or within routes (and possibly biasing relevant information), which affects responses on high-conflict trials. Second, maintenance of recent conflict history in working memory in order to detect changes that require such suppression, which affects responses on LH and HH trials. Third, aiding automatization of a currently advantageous response, which affects responses on HH trials. Taken together, these three mechanisms might explain why interference with dlPFC function affects performance on high-conflict trials differently depending on the nature of the previous trial, and could provide a more detailed explanatory account of how cognitive control is implemented to flexibly adjust behaviour.

A number of hypotheses could be investigated in future research regarding the role of the dlPFC in behavioural adaptation, especially using TMS. Given the correlation found in the imaging literature between activation in medial PFC (especially pre-SMA/ACC), and dlPFC during adaptation, one question concerns whether stimulation of mPFC areas might affect specific trials differently from dlPFC stimulation. Preliminary findings from our DBS study (Chapter 5) indicate that, for the WCST, intracranial stimulation of the ACC did have a numerical effect on the speed of LH trials, in a similar manner to TMS of the dlPFC. However, unlike TMS to dlPFC, ACC stimulation did not

appear to have even a numerical effect on the speed of response on HH trials. This might be due to the specific underlying dynamics of adaptation on LH vs. HH trials in the WCST.

Application of the LBA model to response time data collected from the WCST and the 3-item Simon task (Chapter 4) seems to indicate that different mechanisms might underlie the difference in speed between HH and LH trials in the two tasks. Differences in speed between HH and LH trials in the 3-item Simon task seem to be related to higher caution on LH trials, whereas in the WCST they do not seem to be influenced by caution. We hypothesized that, in the latter task, adaptation might be more dependent on practice effects, as no habit needs to be suppressed in this context. Therefore, while caution might nonetheless play a role because of increased difficulty when a high-conflict trial follows a low-conflict trial (LH), its influence might be significantly less than in a task that requires the suppression of a habit. HH trials in the WCST might be even less dependent on caution, as not only there is no need to suppress a habit on these trials, but they might benefit from practice effects following a high-conflict trial that reinforced the currently rewarded rule.

Therefore, if dlPFC is involved in modulating practice effects, stimulation of this region would be expected to influence trials that benefit from these effects (such as HH trials), which we did observe in our TMS study. On the other hand, if the one of the roles of the ACC is to trigger modulations in caution (as indicated by the studies of van

Maanen and colleagues), effects of stimulation to this region should only affect trials that strongly depend on adjustments in the level of caution. This would be the case of LH trials, particularly when they involve suppression of a habit. This framework could explain the numerical increase in speed of response on WCST LH trials after ACC stimulation, but its failure to reach significance.

We were not able to properly investigate the difference between effects of ACC stimulation in the WCST and the Simon task due to the small sample size, particularly in the latter task. One possibility is that stimulation of the ACC might affect speed of response on LH trials significantly more for tasks involving suppression of a habit than tasks such as the WCST. While TMS is not an ideal tool to investigate the effects of stimulation to the ACC, it might still be possible to test whether these hypothesized effects could be obtained by the stimulation of more superficial mPFC areas, such as pre-SMA, which has been found to be play a role in responses in the presence of conflict (Luks et al., 2007; Mars et al., 2009; Nachev, Rees, & Parton, 2005; Soutschek et al., 2013). Further research could also implement a paired-pulse paradigm to TMS over these two areas, in order to investigate the functional connectivity between these regions during conflict tasks and the effects of time-lagged stimulation of mPFC and dlPFC on response.

FPC, EVALUATION AND A PREFRONTAL NETWORK FOR COUNTERFACTUAL ABILITY?

Conditions of uncertainty and increased error likelihood appear to recruit the ACC and the dlPFC in order to flexibly modulate contextually advantageous behaviour. ACC may be involved in testing whether or not a response that is frequently correlated with a reward is likely to lead to an error in a given context, and signal the need for adjustment to a different response if it does. dlPFC may support the implementation and automatization of the new advantageous response during the early stages of adjustment following a change in context. If these two structures are involved in signalling the need for adaptation, and adapting, respectively, that means some part of the cognitive system must be involved in determining what new behaviour it needs to adapt to. The evaluation and exploration of alternative options is necessary in order to determine the most appropriate response in a given context. As illustrated in Chapter 6, FPC is a likely structure to support this function.

Damage to the FPC affected the acquisition of a new response strategy during the very early stages after a previously advantageous response strategy stopped being rewarded. Control animals showed a rapid decline in errors, compared to FPC animals, only during the very first sessions that required the exploitation of a new abstract rule, indicating that they were able to more efficiently infer which of the possible options was now advantageous. FPC might therefore represent the structure that helps determining which new response is now rewarded, once ACC has signalled that a previously

advantageous response is now likely to lead to errors. It could achieve this both through self-initiated exploration and internal evaluation of alternative options.

FPC's involvement in the exploration of alternatives (see Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006) could represent the equivalent of the ACC's involvement in responding to the occurrence of errors or positive feedback: responding to information from the environment about the outcome of an action. However, we speculated whether ACC might be also involved in counterfactual ability by performing an action-confirmation function, in order to evaluate the risk or benefit of exploiting a generally advantageous response (often a habit) before execution, therefore minimizing the risk of actual errors. The ability to perform such evaluations internally is particularly useful in situations where a large number of options are available, as learning only through exploration and trial-and-error would be costly and time consuming. While human ACC cells have recently been shown not to be sensitive to the number of concurrent possible responses (Sheth et al., 2012), FPC has been shown to be highly active in situations where a large number of options or relations have to be evaluated at the same time (Badre & D'Esposito, 2007; Christoff et al., 2001; Kroger et al., 2002). Therefore, FPC might be an important structure to determine which option or response is more advantageous out of many possible alternatives in a quick and efficient manner, through the simultaneous evaluation of several counterfactual scenarios. FPC has indeed been recently implicated in the processing of counterfactual prediction errors, alongside dorsomedial prefrontal areas that may include parts of the ACC (Boorman et al., 2009,

2011). This, of course, could also aid learning, a process during which FPC has been shown to be very active (Koechlin et al., 2002; Koechlin & Hyafil, 2007; Ramnani & Passingham, 2001; Strange et al., 2001), similarly to dlPFC (Boettiger & D'Esposito, 2005; Lungu et al., 2007; Poldrack et al., 2005).

The human cognitive system has developed throughout evolution to support faster, more efficient and higher-order processing. Therefore, one question concerns whether there are specific cytoarchitectonic differences between species that could be responsible for such a gap in processing power. One obvious answer is of course the larger size of human PFC, and FPC in particular, and the larger number of synaptic connections, compared to the monkey brain. However, recent evidence seems to suggest that the relative size of PFC is not much larger in humans than it is in non-human primates (Passingham, 2002; Semendeferi et al., 2002). Another reason for this gap could instead be found in the different types, or numbers, of cell present across species. In the case of this prefrontal and medial-prefrontal network for counterfactual ability and cognitive flexibility, spindle neurons might represent such inter-species difference.

Spindle neurons have been suggested to underpin intuition in humans (Allman, Watson, Tetreault, & Hakeem, 2005), which is a faculty that involves internally processing many variables at the same time, often in uncertain conditions, in order to reach a decision, very similarly to the counterfactual ability hypothesized for ACC and

FPC. ACC spindle neurons have been shown to project to the FPC, possibly to rely error likelihood information to this area, and the size of FPC in great apes seems to decline as a function of the decrease in density of spindle cells in the ACC (Allman et al., 2002). Furthermore, spindle cells are rich in dopamine D3 receptors, which may be responsible for the signalling of reward expectancy under uncertain conditions in spindle cells (Allman et al., 2005). Spindle cells are also found in human dlPFC, which is another area that, as we saw, is heavily involved in situations of uncertainty and high error likelihood such as conflict and early stages of learning, where counterfactual ability can greatly speed up the selection of what is the most advantageous response in a given context. While these cells have been extensively reported in human ACC, and dlPFC, in the macaque, they have so far been found most often in the insula, and appear to be scarcer and more dispersed in the monkey's ACC (Evrard, Forro, & Logothetis, 2012; Fajardo et al., 2008; Nimchinsky et al., 1999). Spindle neurons could potentially represent a crucial underpinning of human counterfactual ability and possibly the reason for some of the discrepancies found between human and non-human primate research. Further research could help clarify the presence and number of these neurons in the macaque ACC and investigate their role in this potential prefrontal network for counterfactual ability.

LIMITATIONS OF THE CURRENT RESEARCH

A few final remarks concern some methodological limitations with the current research that could be overcome in the future.

What appears evident from the series of study presented in this thesis is that the conflict-analogue of the WCST has revealed the presence of different types of adaptation, and that, consequently, care must be taken in the future when inferring about the role of different neural substrates in conflict-monitoring and behavioural adaptation. While the WCST has proven valuable in identifying these differences (and can continued to be applied in the future to investigate them further), it would be ideal to attempt to minimize the differences between this and other conflict tasks in order to ensure that the different effects obtained are due to the cognitive processes we are interested in, rather than other confounds.

For example, the WCST is a visually rich task, and the visual properties of high-conflict trials in the WCST are different from those of low-conflict trials. During high-conflict trials, all the items on the screen differ from each other in some way, whereas during low-conflict trials only two of the items differ from the others. In classic conflict tasks, high- and low-conflict trials generally share similar visual properties. For example, in the Simon task, a coloured dot appears on the screen in the same manner in high-

and low-conflict trials. In the Stroop task, it is always a word that is presented to the participant. The Flanker task, on the other hand, also presents the same kind of visual differences between high- and low-conflict trials as the WCST: they both involve ignoring distractors on high-conflict trials, although the low-conflict trials of the Flanker task present no distractors at all, unlike the WCST. It might be worth investigating how these different types of presentation might affect adaptation. Conceptually, the Flanker task is more similar to the WCST than it is to the Simon or Stroop tasks: there is no habitual response that would encourage participants to press a left or a right-hand button in response to a particular letter on low-conflict trials. Nonetheless, there might be a spontaneous tendency to react to the information relied by the majority of the stimuli that needs to be suppressed on high-conflict trials, and this suppression could therefore carry over onto low-conflict trials following a high-conflict trials (HL), resulting in slower responses. Furthermore, the possibility of using an abstract rule to select the appropriate response in the WCST, as opposed to an arbitrary stimulus-response mapping in the Flanker task or the Simon task, could provide a shortcut to response that can speed up responses on high-conflict trials. It would be interesting to design versions of the WCST where low-conflict trials are defined by all identical items, therefore making them similar to low-conflict Flanker trials, or by simply removing one of the test items that would share a property with the sample in a high-conflict trials, therefore maintaining the visual similarity between high- and low-conflict trials as in the Stroop or Simon tasks.

The idea that adaptation effects can be due two different mechanisms, depending on whether HH or HL trials are affected, also poses a 'chicken and the egg' type of conundrum concerning the change in conflict levels. Is it the case that HL trials slow down compared to LL trials, or LL trials speed up, compared to HL trials? Similarly, do HH speed up or do LH trials slow down? Stoffels' (1996) evidence that it is only low-conflict trials that area affected, when high- and low-conflict trials are intermixed in the same block, combined with Jacobi and colleagues' (2001) findings that increasing the percentage of high-conflict trials in a Stroop task only influenced word-reading (i.e. the habit) processes, indicate that it is likely to be the case that differences in speed between HL and LL are due to slower responses on HL. Stoffels also indicated that high-conflict trials were unaffected by being intermixed with low-conflict trials, suggesting that adaptation effects on high-conflict trial are likely due to HH trials increasing in speed, rather than LH trials being slower due to the effects of previous conflict levels. This however contradicts the findings from our TMS study (Chapter 2) indicating some type of behavioural modulation following low-conflict trials. The issue might be clarified by investigating different types of tasks with TMS, such as the 3-item version of the Simon task, which also involves differences in LH and HH response speed, to test whether the effects of stimulation on LH trials might be due to the nature of the WCST analogue or might underlie a more general type of adaptation mechanism. Determining the mechanisms driving adaptation during sequences of high-conflict trials could also shed some light on this issue. While differences in speed between HH and LH represent

a genuine type of adaptation effect, its mechanisms, as illustrated in this thesis, are still unclear, and require further investigation.

Possibly the greatest limitation of the methodology applied in this body of work concerns the use of touchscreens. Touchscreens, which we implemented in all tasks in order to minimize differences in task design between human and non-human primate research, might be the underlying cause of the extremely low error rates, which precluded us from carrying out meaningful accuracy analysis, as well as possibly limiting the explanatory power of the LBA model application. Touchscreens are likely to provide participants with more time to abort and correct their responses, compared to button presses and, especially, compared to the type of saccade responses often used in the animal literature. Furthermore, searching and selecting a response on the screen involves very large limb movements, which generated great confounding motor activity in our LFP recordings, limiting inferences about conflict-related activation in the ACC. As cingulate motor areas only synapse directly onto motor areas concerned with limb and finger movement, and the spinal cord (Dum & Strick, 2002), but are much less connected with areas controlling eye-movements (Schall & Emeric, 2010), using eye-tracking versions of the tasks might be a more suitable option to investigate conflict-related activation in the ACC from intracranial recordings.

FINAL COMMENTS

To conclude, we have shown that human ACC is not always necessary for conflict-monitoring and conflict-induced behavioural adaptation, but that, rather, its role in modulating behaviour during situations where conflict is present might be better defined as signalling the need to move away from a default, generally advantageous, behaviour, in response to changes in context and environmental fluctuations. This could be achieved through the processing of feedback in order to proactively anticipate adjustments in response strategies and, possibly, simulating predictions about the application of default behaviour to the current context. This may drive adaptation through increases in caution when a prediction error is obtained. dlPFC, on the other hand, is necessary for the implementation of adaptation, and this could be achieved through several mechanisms, including inhibition of the default behaviour, maintenance of a memory trace of recent environmental contingencies or automatization of a new advantageous response. Further research could aim at better distinguishing between these processes and the role dlPFC plays in them. Finally, FPC appears to be necessary to support the simultaneous evaluation and/or exploration of multiple possible response options in order to rapidly determine which one is the new advantageous response, once a previously option stops being rewarded. Given their roles in reward and error prediction, the implementation and automatization of goal-directed behaviour and the evaluation and the exploration of alternatives, these three structures and their interconnections could represent a prefrontal network for counterfactual ability, a

fundamental property of the human cognitive system that underpins the extensive cognitive flexibility that sets our species apart from other primates.

REFERENCES

Aarts, E., Roelofs, A., & van Turennout, M. (2008). Anticipatory activity in anterior cingulate cortex can be independent of conflict and error likelihood. *The Journal of neuroscience*, *28*(18), 4671–8.

Akçay, C., & Hazeltine, E. (2008). Conflict adaptation depends on task structure. *Journal of experimental psychology. Human perception and performance*, *34*(4), 958–73.

Allman, J. M., Hakeem, A., & Watson, K. (2002). Two phylogenetic specializations in the human brain. *The Neuroscientist*, *8*(4), 335–46.

Allman, J. M., Watson, K. K., Tetreault, N. A., & Hakeem, A. Y. (2005). Intuition and autism: a possible role for Von Economo neurons. *Trends in cognitive sciences*, *9*(8), 367–73.

Allport, A. (1987). Selection for action: Some behavioural and neuropsychological considerations of attention and action. In H. Heuer & A.F. Sanders (Eds.), *Perspectives on perception and action* (pp. 395-419). Hillsdale, NJ: Erlbaum.

Altamura, M., Carver, F. W., Elvevåg, B., Weinberger, D. R., & Coppola, R. (2013). Dynamic cortical involvement in implicit anticipation during statistical learning. *Neuroscience letters*, *558*, 73–77.

Asaad, W. F., Rainer, G., & Miller, E. K. (2000). Task-Specific Neural Activity in the Primate Prefrontal Cortex. *Journal of Neurophysiology*, 451–459.

Asaad, W. F., Rainer, G., & Miller, E. K. (1998). Neural activity in the primate prefrontal cortex during associative learning. *Neuron*, *21*, 1399–1407.

Baddeley, A.D., & Hitch, G. J. (1974). Working Memory. In G.H. Bower (Ed.), *The Psychology of Learning and Motivation*, Vol. 8 (pp. 47-89). London: Academic Press.

Baddeley, A. D. (1996). Exploring the Central Executive. *The Quarterly Journal of Experimental Psychology*, (1), 5–28.

Badre, D. (2008). Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends in cognitive sciences*, *12*(5), 193–200.

Badre, D., & D'Esposito, M. (2007). Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. *Journal of Cognitive Neuroscience*, 2082–2099.

Badre, D., & Wagner, A. D. (2004). Selection, integration, and conflict monitoring; assessing the nature and generality of prefrontal cognitive control mechanisms. *Neuron*, *41*(3), 473–87.

Bain, P.G. (2009). *Deep Brain Stimulation*. Oxford: Oxford University Press.

Barch, D. M., Braver, T. S., Sabb, F. W., & Noll, D. C. (2000). Anterior Cingulate and the Monitoring of Response Conflict: Evidence from an fMRI Study of Overt Verb Generation. *Journal of Cognitive Neuroscience*, *12*(2), 298–309.

Barr, M. S., Farzan, F., Rusjan, P. M., Chen, R., Fitzgerald, P. B., & Daskalakis, Z. J. (2009). Potentiation of gamma oscillatory activity through repetitive transcranial magnetic stimulation of the dorsolateral prefrontal cortex. *Neuropsychopharmacology*, *34*(11), 2359–67.

Bartholow, B. D., Pearson, M., Dickter, C. L., Sher, K. J., Fabiani, M., & Gratton, G. (2005). Strategic control and medial frontal negativity: beyond errors and response conflict. *Psychophysiology*, *42*(1), 33–42.

Bauer, R. H., & Fuster, J. M. (1976). Delayed-matching and delayed-response deficit from cooling dorsolateral prefrontal cortex in monkeys. *Journal of comparative and physiological psychology*, *90*(3), 293–302.

Beam, W., Borckardt, J., Reeves, S., & George, M. (2009). An efficient and accurate new method for locating the F3 position for prefrontal TMS applications. *Brain stimulation*, *2*(1), 50–54.

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature neuroscience*, *10*(9), 1214–21.

Bittar, R. G., Burn, S. C., Bain, P. G., Owen, S. L., Joint, C., Shlugman, D., & Aziz, T. Z. (2005). Deep brain stimulation for movement disorders and pain. *Journal of clinical neuroscience, 12*(4), 457–63.

Bludau, S., Eickhoff, S. B., Mohlberg, H., Caspers, S., Laird, a R., Fox, P. T., Amunts, K. (2013). Cytoarchitecture, probability maps and functions of the human frontal pole. *NeuroImage*.

Boccard, S. G. J., Pereira, E. C., Moir, L., Van Hartevelt, T. J., Kringelbach, M. L., Fitzgerald, J. J., Aziz, T. Z. (2013). Deep brain stimulation of the anterior cingulate cortex: targeting the affective component of chronic pain. *Neuroreport*, 1–6.

Boettiger, C., & D'Esposito, M. (2005). Frontal networks for learning and executing arbitrary stimulus-response associations. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *25*(10), 2723–32.

Boorman, E. D., Behrens, T. E. J., Woolrich, M. W., & Rushworth, M. F. S. (2009). How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron*, *62*(5), 733–43.

Boorman, E. D., Behrens, T. E. J., & Rushworth, M. F. S. (2011). Counterfactual choice and learning in a neural network centered on human lateral frontopolar cortex. *PLoS biology*, *9*(6).

Boorman, E. D., Rushworth, M. F. S., & Behrens, T. E. J.(2013). Ventromedial prefrontal and anterior cingulate cortex adopt choice and default reference frames during sequential multi-alternative choice. *The Journal of neuroscience, 33*(6), 2242–53.

Botvinick, M. M. (2007). Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cognitive, Affective, & Behavioral Neuroscience*, *7*(4), 356–66.

Botvinick, M. M., Braver, T., & Barch, D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624–652.

Botvinick, M. M. (2008). Hierarchical models of behavior and prefrontal function. *Trends in cognitive sciences*, *12*(5), 201–8.

Braver, T. S., Barch, D. M., Gray, J. R., Molfese, D. L., & Snyder, A. (2001). Anterior cingulate cortex and response conflict: effects of frequency, inhibition and errors. *Cerebral cortex*, *11*(9), 825–36.

Braver, T. S., & Bongiolatti, S. R. (2002). The role of frontopolar cortex in subgoal processing during working memory. *NeuroImage*, *15*(3), 523–36.

Brodmann, K. (1912). Neue Ergebnisse uber die vergleichende histologische Lokalisation der Grosshirnrinde mit besonderer Berucksichtingung des Stirnhirns. *Anat. Anz. 4 41 (Suppl.),* 157-216.

Brown, J. W., & Braver, T. S. (2008). A computational model of risk, conflict, and individual difference effects in the anterior cingulate cortex. *Brain research*, *1202*, 99–108.

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: linear ballistic accumulation. *Cognitive psychology*, *57*(3), 153–78.

Brown, S. D., & Heathcote, A. (2005). A ballistic model of choice response time. *Psychological review*, *112*(1), 117–28.

Buchsbaum, B. R., Greer, S., Chang, W. L., & Berman, K. F. (2005). Meta-analysis of neuroimaging studies of the Wisconsin card-sorting task and component processes. *Human brain mapping*, *25*(1), 35–45.

Buckley, M. J., Mansouri, F. A., Hoda, H., Mahboubi, M., Browning, P. G. F., Kwok, S. C., & Tanaka, K. (2009). Dissociable components of rule-guided behavior depend on distinct medial and prefrontal regions. *Science*, *325*(5936), 52–8.

Bunge, S., Kahn, I., Wallis, J. D., Miller, E. K., & Wagner, A. D. (2003). Neural circuits subserving the retrieval and maintenance of abstract rules. *Journal of neurophysiology*, *90*(5).

Burgess, P. W., Veitch, E., de Lacy C., & Shallice, T. (2000). The cognitive and neuroanatomical correlates of multitasking. *Neuropsychologia*, *38*(6), 848–63.

Burle, B., Allain, S., Vidal, F., & Hasbroucq, T. (2005). Sequential compatibility effects and cognitive control: does conflict really matter? *Journal of experimental psychology. Human perception and performance*, *31*(4), 831–7.

Burle, B., Possamaï, C.-A., Vidal, F., Bonnet, M., & Hasbroucq, T. (2002). Executive control in the Simon effect: an electromyographic and distributional analysis. *Psychological research*, *66*(4), 324–36.

Burle, B., Roger, C., Allain, S., Vidal, F., & Hasbroucq, T. (2008). Error negativity does not reflect conflict: a reappraisal of conflict monitoring and anterior cingulate cortex activity. *Journal of Cognitive Neuroscience*, *20*(9), 1637–55.

Burman, K. J., Reser, D. H., Richardson, K. E., Gaulke, H., Worthy, K. H., & Rosa, M. G. P. (2011). Subcortical projections to the frontal pole in the marmoset monkey. *The European Journal of Neuroscience*, *34*(2), 303–19.

Burman, K. J., Reser, D. H., Yu, H.-H., & Rosa, M. G. P. (2011). Cortical input to the frontal pole of the marmoset monkey. *Cerebral cortex*, *21*(8), 1712–37.

Bush, G., Luu, P., & Posner, M. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in cognitive sciences*, *4*(6), 215–222.

Bush, G., Vogt, B., Holmes, J., Dale, A. M., Greve, D., Jenike, M., & Rosen, B. R. (2002). Dorsal anterior cingulate cortex: a role in reward-based decision making. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(1), 523–8.

Carbonnell, L., & Falkenstein, M. (2006). Does the error negativity reflect the degree of response conflict? *Brain research*, *1095*(1), 124–30.

Carter, C. S., Macdonald, A. M., Botvinick, M., Ross, L. L., Stenger, V. A., Noll, D., & Cohen, J. D. (2000). Parsing executive processes: strategic vs. evaluative functions of the anterior cingulate cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(4), 1944–8.

Carter, C., & van Veen, V. (2007). Anterior cingulate cortex and conflict detection: an update of theory and data. *Cognitive, Affective, & Behavioral Neuroscience*, *7*(4), 367–379.

Chao, L., & Knight, R. (1995). Human prefrontal lesions increase distractability to irrelevant sensory inputs. *Neuroreport*.

Chen, Q., Wei, P., & Zhou, X. (2006). Distinct neural correlates for resolving stroop conflict at inhibited and noninhibited locations in inhibition of return. *Journal of Cognitive Neuroscience*, *18*(11), 1937–46.

Chen, S., & Melara, R. D. (2009). Sequential effects in the Simon task: conflict adaptation or feature integration? *Brain research*, *1297*, 89–100.

Christoff, K., & Keramatian, K. (n.d.). Abstraction of mental representations: Theoretical consideration and Neuroscientific evidence.

Christoff, K., Prabhakaran, V., Dorfman, J., Zhao, Z., Kroger, J. K., Holyoak, K. J., & Gabrieli, J. D. (2001). Rostrolateral prefrontal cortex involvement in relational integration during reasoning. *NeuroImage*, *14*(5), 1136–49.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychological review*, *97*(3), 332–61.

Cohen, M. X., Ridderinkhof, K. R., Haupt, S., Elger, C. E., & Fell, J. (2008). Medial frontal cortex and response conflict: evidence from human intracranial EEG and medial frontal cortex lesion. *Brain research*, *1238*, 127–42.

Cohen, R. a., Kaplan, R. F., Moser, D. J., Jenkins, M. a., & Wilkinson, H. (1999). Impairments of attention after cingulotomy. *Neurology*, *53*(4), 819–819.

Cole, M. W., Yeung, N., Freiwald, W., & Botvinick, M. (2010). Conflict over cingulate cortex: Between-species differences in cingulate may support enhanced cognitive flexibility in humans. *Brain, behavior and evolution*, *75*(4), 239–40.

Cole, M. W., Yeung, N., Freiwald, W. A., & Botvinick, M. (2009). Cingulate cortex: diverging data from humans and monkeys. *Trends in neurosciences*, *32*(11), 566–74.

Coles, M. G., Scheffers, M. K., & Holroyd, C. B. (2001). Why is there an ERN/Ne on correct trials? Response representations, stimulus-related components, and the theory of error-processing. *Biological psychology*, *56*(3), 173–89.

Coles, M., & Gratton, G. (1985). A Psychophysiological Investigation of the Continuous Flow Model of Human Information Processing. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(5), 529–553.

Davelaar, E. J. (2008). A computational study of conflict-monitoring at two levels of processing: reaction time distributional analyses and hemodynamic responses. *Brain research*, *1202*, 109–19.

Davis, K. D., Hutchison, W. D., Lozano, A. M., Tasker, R. R., & Dostrovsky, J. O. (2000). Human anterior cingulate cortex neurons modulated by attention-demanding tasks. *Journal of neurophysiology*, *83*(6), 3575–7.

Davis, K., & Taylor, K. (2005). Human anterior cingulate cortex neurons encode cognitive and emotional demands. *The Journal of Neuroscience*, *25*, 8402–8406.

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature neuroscience*, *8*(12), 1704–11.

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*(7095), 876–9.

De Pisapia, N., Sandrini, M., Braver, T. S., & Cattaneo, L. (2012). Integration in working memory: a magnetic stimulation study on the role of left anterior prefrontal cortex. *PloS one*, *7*(8).

Depisapia, N., & Braver, T. (2006). A model of dual control mechanisms through anterior cingulate and prefrontal cortex interactions. *Neurocomputing*, *69*(10-12), 1322–1326.

Di Pellegrino, G., Ciaramelli, E., & Làdavas, E. (2007). The regulation of cognitive control following rostral anterior cingulate cortex lesion in humans. *Journal of Cognitive Neuroscience*, *19*(2), 275–86.

Dias, R., Robbins, T., & Roberts, A. (1996). Primate Analogue of the Wisconsin Card Sorting Test : Effects of Excitotoxic Lesions of the Prefrontal Cortex in the Marmoset. *Behavioral neuroscience*, *110*(5), 872–886.

Dickinson, A. (1985). Actions and Habits: The Development of Behavioural Autonomy. *Philosophical Transactions of the Royal Society of London, 308*(1135), 67-78.

Donkin, C., Brown, S., & Heathcote, A. (2011). Drawing conclusions from choice response time models: A tutorial using the linear ballistic accumulator. *Journal of Mathematical Psychology*, *55*(2), 140–151.

Doyon, J., & Benali, H. (2005). Reorganization and plasticity in the adult brain during learning of motor skills. *Current opinion in neurobiology*, *15*(2), 161–7.

Doyon, J., Penhune, V., & Ungerleider, L. G. (2003). Distinct contribution of the cortico-striatal and cortico-cerebellar systems to motor skill learning. *Neuropsychologia*, *41*(3), 252–62.

Dreher, J.-C., Koechlin, E., Tierney, M., & Grafman, J. (2008). Damage to the fronto-polar cortex is associated with impaired multitasking. *PloS one*, *3*(9), e3227.

Dum, R. P., & Strick, P. L. (2002). Motor areas in the frontal lobe of the primate. *Physiology & behavior*, *77*(4-5), 677–82.

Duncan, J. (2001). An adaptive coding model of neural function in prefrontal cortex. *Nature reviews. Neuroscience*, *2*(11), 820–9.

Durston, S. (2003). Parametric manipulation of conflict and response competition using rapid mixed-trial event-related fMRI. *NeuroImage*, *20*(4), 2135–2141.

Egner, T. (2007). Congruency sequence effects and cognitive control. *Cognitive, affective & behavioral neuroscience*, *7*(4), 380–90.

Egner, T. (2008). Multiple conflict-driven control mechanisms in the human brain. *Trends in Cognitive Sciences*, *12*(10), 374-380.

Egner, T., Delano, M., & Hirsch, J. (2007). Separate conflict-specific cognitive control mechanisms in the human brain. *NeuroImage*, *35*(2), 940–8.

Egner, T., & Hirsch, J. (2005a). The neural correlates and functional integration of cognitive control in a Stroop task. *NeuroImage*, *24*(2), 539–47.

Egner, T., & Hirsch, J. (2005b). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience*, *8*, 1784–1790.

Elston, G. N. (2000). Pyramidal cells of the frontal lobe: all the more spinous to think with. *The Journal of Neuroscience*, *20*(18).

Emeric, E. E., Brown, J. W., Leslie, M., Pouget, P., Stuphorn, V., & Schall, J. D. (2008). Performance monitoring local field potentials in the medial frontal cortex of primates: anterior cingulate cortex. *Journal of neurophysiology*, *99*(2), 759–72.

Erickson, K. I., Milham, M. P., Colcombe, S. J., Kramer, A. F., Banich, M. T., Webb, A., & Cohen, N. J. (2004). Behavioral conflict, anterior cingulate cortex, and experiment duration: implications of diverging data. *Human brain mapping*, *21*(2), 98–107.

Eriksen, C. W., & Schultz, D. W. (1979). Information processing in visual search: a continuous flow conception and experimental results. *Perception & psychophysics*, *25*(4), 249–63.

Evrard, H. C., Forro, T., & Logothetis, N. K. (2012). Von Economo neurons in the anterior insula of the macaque monkey. *Neuron*, *74*(3), 482–9.

Fajardo, C., Escobar, M. I., Buriticá, E., Arteaga, G., Umbarila, J., Casanova, M. F., & Pimienta, H. (2008). Von Economo neurons are present in the dorsolateral (dysgranular) prefrontal cortex of humans. *Neuroscience letters*, *435*(3), 215–8.

Fan, J. (2003). Cognitive and Brain Consequences of Conflict. *NeuroImage*, *18*(1), 42–57.

Fellows, L. K., & Farah, M. J. (2005). Is anterior cingulate cortex necessary for cognitive control. *Brain*, *128*, 788–796.

Fernandez-Duque, D., & Knight, M. (2008). Cognitive control: dynamic, sustained, and voluntary influences. *Journal of Experimental Psychology. Human Perception and Performance*, *34*(2), 340–55.

Forster, S. E., Carter, C. S., Cohen, J. D., & Cho, R. Y. (2011). Parametric manipulation of the conflict signal and control-state adaptation. *Journal of cognitive neuroscience*, *23*(4), 923–35.

Forstmann, B. U., Anwander, A., Schäfer, A., Neumann, J., Brown, S., Wagenmakers, E.-J., Turner, R. (2010). Cortico-striatal connections predict control over speed and

accuracy in perceptual decision making. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(36), 15916–20.

Forstmann, B. U., Brown, S., Dutilh, G., Neumann, J., & Wagenmakers, E.-J. (2010). The neural substrate of prior information in perceptual decision making: a model-based analysis. *Frontiers in human neuroscience*, *4*, 40.

Forstmann, B. U., Dutilh, G., Brown, S., Neumann, J., von Cramon, D. Y., Ridderinkhof, K. R., & Wagenmakers, E.-J. (2008). Striatum and pre-SMA facilitate decision-making under time pressure. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(45), 17538–42.

Fuster, J., & Alexander, G. (1971). Neuron activity related to short-term memory. *Science, 173*(3977), 652-654.

Fuster, J. M. (1990). Prefrontal cortex and the bridging of temporal gaps in the perception-action cycle. *Annals of the New York Academy of Sciences*, *608*, 318–29.

Fuster, J.M. (2008). *The Prefrontal Cortex.* London: Academic Press.

Galashan, D., Wittfoth, M., Fehr, T., & Herrmann, M. (2008). Two Simon tasks with different sources of conflict: an ERP study of motion- and location-based compatibility effects. *Biological Psychology*, *78*(3), 246–52.

Gardner, R.W. (1959). Cognitive control principles and perceptual behavior. *Bulletin of the Menninger Clinic, 23*, 241-248.

Gardner, R.W., Holzman, P. S., Klein, G. S., Linton, H. B., & Spence, D. P. (1959). *Cognitive control: A study of individual consistencies in cognitive behavior*. *Psychological issues.* 1(4). Madison, CT, US: International Universities Press.

Gardner, R.W., Lohrenz, L.J., & Schoen, R.A. (1968). Cognitive Control of Differentiation in the Perception of Persons and Objects. *Perceptual and Motor Skills, 26,* 311-330.

Gasquoine, P. G. (2013). Localization of function in anterior cingulate cortex: from psychosurgery to functional neuroimaging. *Neuroscience and biobehavioral reviews*, *37*(3), 340–8.

Gilbert, S. J., Spengler, S., Simons, J. S., Frith, C. D., & Burgess, P. W. (2005). Differential Functions of Lateral and Medial Rostral Prefrontal Cortex (Area 10) Revealed by Brain-Behavior Associations. *Cerebral Cortex*, *16*(12), 1783–1789.

Gilbert, S. J, Spengler, S., Simons, J. S., Steele, J. D., Lawrie, S. M., Frith, C. D., & Burgess, P. W. (2006). Functional specialization within rostral prefrontal cortex (area 10): a meta-analysis. *Journal of cognitive neuroscience*, *18*(6), 932–48.

Gläscher, J., Adolphs, R., Damasio, H., Bechara, A., Rudrauf, D., Calamia, M., Tranel, D. (2012). Lesion mapping of cognitive control and value-based decision making in the prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(36), 14681–6.

Godefroy, O., & Rousseaux, M. (1996). Divided and focused attention in patients with lesion of the prefrontal cortex. *Brain and cognition*, *30*(2), 155–74.

Gold, J. I., & Shadlen, M. N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in cognitive sciences*, *5*(1), 10–16.

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

Goldman-Rakic, P. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. *Comprehensive Physiology*.

Grant, D., & Berg, E. (1948). A behavioural analysis of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of experimental psychology*, 404–411.

Gratton, G., Coles, M. G. H., & Donchin, E. (1992). Optimizing the use of information: strategic control of activation of responses. *Journal of Experimental Psychology: General, 121*(4), 480-506.

Gross, J., Schmitz, F., Schnitzler, I., Kessler, K., Shapiro, K., Hommel, B., & Schnitzler, A. (2006). Anticipatory control of long-range phase synchronization. *The European Journal of Neuroscience*, *24*(7), 2057–60.

Hallett, M. (2007). Transcranial magnetic stimulation: a primer. *Neuron*, *55*(2), 187–99.

Hanes, D. P., & Carpenter, R. H. (1999). Countermanding saccades in humans. *Vision research*, *39*(16), 2777–91.

Hayden, B. Y., Heilbronner, S. R., Pearson, J. M., & Platt, M. L. (2011). Surprise signals in anterior cingulate cortex: neuronal encoding of unsigned reward prediction errors driving adjustment in behavior. *The Journal of Neuroscience*, *31*(11), 4178–87.

Hayden, B. Y., Pearson, J. M., & Platt, M. L. (2009). Fictive reward signals in the anterior cingulate cortex. *Science, 324*(5929), 948–50.

Hayward, G., Goodwin, G. M., & Harmer, C. J. (2004). The role of the anterior cingulate cortex in the counting Stroop task. *Experimental brain research*, *154*(3), 355–8.

Hazeltine, E. (2003). Material-dependent and material-independent selection processes in the frontal and parietal lobes: an event-related fMRI investigation of response competition. *Neuropsychologia*, *41*(9), 1208–1217. d

Hazeltine, E., Poldrack, R., & Gabrieli, J. D. E. (2000). Neural activation during response competition. *Journal of Cognitive Neuroscience*, *12*(2), 118–129.

Hedge, A., & Marsh, N. (1975). The effect of irrelevant spatial correspondences on two-choice response-time. *Acta Psychologica*, *39*(6), 427–439.

Henik, A., Bibi, U., Yanai, M. & Tzelgov, J. (1997). The Stroop effect is largest during first trials. *Abstracts of the Psychonomic Society 2, 57*.

Ho, T., Brown, S., van Maanen, L., Forstmann, B. U., Wagenmakers, E.-J., & Serences, J. T. (2012). The optimality of sensory processing during the speed-accuracy tradeoff. *The Journal of Neuroscience*, *32*(23).

Ho, T. C., Brown, S., Abuyo, N. A., Ku, E. J., & Serences, J. T. (2012). Perceptual consequences of feature-based attentional enhancement and suppression. *Journal of Vision, 12*(8), 1–17.

Holroyd, C. B., Nieuwenhuis, S., Yeung, N., Nystrom, L., Mars, R. B., Coles, M. G. H., & Cohen, J. D. (2004). Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nature neuroscience*, *7*(5), 497–8.

Holroyd, C. B., Yeung, N., Coles, M. G. H., & Cohen, J. D. (2005). A mechanism for error detection in speeded response time tasks. *Journal of Experimental Psychology: General*, *134*(2).

Hommel, B., Proctor, R., & Vu, K. (2004). A feature-integration account of sequential effects in the Simon task. *Psychological research*, *68*, 1–17.

Hommel, B. (1994). Spontaneous decay of response-code activation. *Psychological Research*, *56*(4), 261–268.

Horga, G., Maia, T. V, Wang, P., Wang, Z., Marsh, R., & Peterson, B. S. (2011). Adaptation to conflict via context-driven anticipatory signals in the dorsomedial prefrontal cortex. *The Journal of Neuroscience*, *31*(45), 16208-16216.

Isaias, I.U., & Tagliati, M. (2008). Deep Brain Stimulation Programming for Motor Disorders. In D. Tarsy, J.L. Vitek, P.A. Starr, & M.S. Okun (Eds.). *Deep Brain Stimulation in Neurological and Pychiatric Disorders.* Totowa, NJ: Humana Press.

Ito, S., Stuphorn, V., Brown, J. W., & Schall, J. D. (2003). Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science*, *302*(5642), 120–2.

Jacobs, B., Schall, M., Prather, M., Kapler, E., Driscoll, L., Baca, S., Treml, M. (2001). Regional dendritic and spine variation in human cerebral cortex: a quantitative golgi study. *Cerebral cortex*, *11*(6), 558–71.

Jacoby, L. L., Lindsay, D. S., & Hessels, S. (2003). Item-specific control of automatic processes: stroop process dissociations. *Psychonomic bulletin & review*, *10*(3), 638–44.

Jahanshahi, M., Profice, P., Brown, R. G., Ridding, M. C., Dirnberger, G., & Rothwell, J. C. (1998). The effects of transcranial magnetic stimulation over the dorsolateral prefrontal cortex on suppression of habitual counting during random number generation. *Brain: a journal of neurology*, *121*, 1533–44.

James, W. (1980). *Principles of Psychology.* New York: Holt.

Jenkins, I. H., Brooks, D. J., Frackowiak, R. S. J., & Passingham, F. E. (1994). Motor Sequence Tomography Learning : A Study with Positron Emission Topography. *The Journal of Neuroscience, 14*(6), 3775-3790.

Jia, S., Li, H., Luo, Y., Chen, A., Wang, B., & Zhou, X. (2007). Detecting perceptual conflict by the feedback-related negativity in brain potentials. *Neuroreport*, *18*(13), 1385–8.

Jin, Y., Olk, B., & Hilgetag, C. C. (2010). Contributions of human parietal and frontal cortices to attentional control during conflict resolution: a 1-Hz offline rTMS study. *Experimental brain research*, *205*(1), 131–8.

Kadota, H., Sekiguchi, H., Takeuchi, S., Miyazaki, M., Kohno, Y., & Nakajima, Y. (2010). The role of the dorsolateral prefrontal cortex in the inhibition of stereotyped responses. *Experimental brain research*, *203*(3), 593–600.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: an individual-differences perspective. *Psychonomic bulletin & review*, *9*(4), 637–71.

Keil, J., Weisz, N., Paul-Jordanov, I., & Wienbruch, C. (2010). Localization of the magnetic equivalent of the ERN and induced oscillatory brain activity. *NeuroImage*, *51*(1), 404–11.

Kennerley, S. W., Behrens, T. E. J., & Wallis, J. D. (2011). Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nature Neuroscience, 14*(12), 1581-9.

Kennerley, S. W., & Wallis, J. D. (2009). Evaluating choices by single neurons in the frontal lobe: outcome value encoded across multiple decision variables. *The European Journal of Neuroscience, 29*(10), 2061-73.

Kennerley, S. W., Walton, M. E., Behrens, T. E. J., Buckley, M. J., & Rushworth, M. F. S. (2006). Optimal decision making and the anterior cingulate cortex. *Nature Neuroscience*, *9*(7), 940–7.

Kerns, J. G. (2006). Anterior cingulate and prefrontal cortex activity in an FMRI study of trial-to-trial adjustments on the Simon task. *NeuroImage*, *33*(1), 399–405.

Kerns, J. G., Cohen, J. D., MacDonald, A. W., Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science*, *303*(5660), 1023–6.

Kim, C., Chung, C., & Kim, J. (2012). Conflict adjustment through domain-specific multiple cognitive control mechanisms. *Brain research*, *1444*, 55–64.

Kim, C., Kroger, J. K., & Kim, J. (2011). A functional dissociation of conflict processing within anterior cingulate cortex. *Human brain mapping*, *32*(2), 304–12.

King, J. a, Donkin, C., Korb, F. M., & Egner, T. (2012). Model-based analysis of context-specific cognitive control. *Frontiers in psychology*, *3*.

Knight, R. T., Staines, W. R., Swick, D., & Chao, L. (1999). Prefrontal cortex regulates inhibition and excitation in distributed neural networks. *Acta psychologica*, *101*, 159–178.

Knight, R. T., Grabowecky, M. F., & Scabini, D. (1995, January). Role of human prefrontal cortex in attention control. *Advances in neurology*.

Knoch, D., Brugger, P., & Regard, M. (2005). Suppressing versus releasing a habit: frequency-dependent effects of prefrontal transcranial magnetic stimulation. *Cerebral cortex*, *15*(7), 885–7.

Ko, J. H., Monchi, O., Ptito, A., Petrides, M., & Strafella, A. P. (2008). Repetitive transcranial magnetic stimulation of dorsolateral prefrontal cortex affects performance of the wisconsin card sorting task during provision of feedback. *International Journal of Biomedical Imaging*, *2008*, 143238.

Koechlin, E., Basso, G., Pietrini, P., Panzer, S., & Grafman, J. (1999). The role of the anterior prefrontal cortex in human cognition. *Nature*, *399*(6732), 148–51.

Koechlin, E., Corrado, G., Pietrini, P., & Grafman, J. (2000). Dissociating the role of the medial and lateral anterior prefrontal cortex in human planning. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(13), 7651–6.

Koechlin, E., Danek, A., Burnod, Y., & Grafman, J. (2002). Medial prefrontal and subcortical mechanisms underlying the acquisition of motor and cognitive action sequences in humans. *Neuron*, *35*(2), 371–81.

Koechlin, E., & Hyafil, A. (2007). Anterior prefrontal function and the limits of human decision-making. *Science, 318*(5850), 594–8.

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science*, *302*(5648), 1181–5.

Koechlin, Etienne, & Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends in cognitive sciences*, *11*(6), 229–35.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus-response compatibility–A model and taxonomy. *Psychological Review*, *97*(2), 253.

Kringelbach, M., & Jenkinson, N. (2007). Translational principles of deep brain stimulation. *Nature Reviews Neuroscience*, 623–635.

Kroger, J. K., Sabb, F. W., Fales, C. L., Bookheimer, S. Y., Cohen, M. S., & Holyoak, K. J. (2002). Recruitment of anterior dorsolateral prefrontal cortex in human reasoning: a parametric study of relational complexity. *Cerebral cortex*, *12*(5), 477–85.

Kunde, W. (2003). Sequential modulations of stimulus-response correspondence effects depend on awareness of response conflict. *Psychonomic bulletin & review*, *10*(1), 198–205.

Laming, D. (1979). Choice Reaction Performnce Following and Error. *Acta Psychologica*, *43*, 199–224.

Lapish, C. C., Durstewitz, D., Chandler, L. J., & Seamans, J. K. (2008). Successful choice behavior is associated with distinct and coherent network states in anterior cingulate cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(33), 11963–8.

Larson, M. J., Kaufman, D. S., & Perlstein, W. M. (2009). Neural time course of conflict adaptation effects on the Stroop task. *Neuropsychologia*, *47*(3), 663–70.

Lie, C.-H., Specht, K., Marshall, J. C., & Fink, G. R. (2006). Using fMRI to decompose the neural processes underlying the Wisconsin Card Sorting Test. *NeuroImage*, *30*(3), 1038–49.

Liu, Y., & Postupna, N. (2008). High frequency deep brain stimulation: what are the therapeutic mechanisms? *Neuroscience & Biobehavioral Reviews*, *32*, 343–351.

Logan, G. (1979). When it helps to be misled: Facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Memory & Cognition*, *7*(3), 166–174.

Lu, C., & Proctor, R. W. (1995). The influence of irrelevant location information on performance : A review of the Simon and spatial Stroop effects. *Psychonomic Bulletin & Review*, *2*(2), 174–207.

Luks, T. L., Simpson, G. V, Dale, C. L., & Hough, M. G. (2007). Preparatory allocation of attention and adjustments in conflict processing. *NeuroImage*, *35*(2), 949–58.

Lungu, O. V, Binenstock, M. M., Pline, M., Yeaton, J. R., & Carey, J. R. (2007). Neural changes in control implementation of a continuous task. *The Journal of Neuroscience*, *27*(11), 3010–6.

MacDonald, W. (2000). Dissociating the Role of the Dorsolateral Prefrontal and Anterior Cingulate Cortex in Cognitive Control. *Science*, *288*(5472), 1835–1838.

Maier, E. M., diPellegrino, G. (2012). Impaired Conflict Adaptation in an Emotional Context Following Rostral Anterior Cingulate Cortex Lesions in Humans. *Journal of Cognitive Neuroscience, 24*(10), 2070-2079.

Mansouri, F., Buckley, M. J., & Tanaka, K. (2007a). Mnemonic function of the dorsolateral prefrontal cortex in conflict-induced behavioral adjustment. *Science*, *318*(5852), 987–90.

Mansouri, F., Buckley, M. J., & Tanaka, K. (2007b). Methods: Mnemonic function of the dorsolateral prefrontal cortex in conflict-induced behavioral adjustment. *Science, 318*(5852), 987–90.

Mansouri, F., Tanaka, K., & Buckley, M. J. (2009). Conflict-induced behavioural adjustment: a clue to the executive functions of the prefrontal cortex. *Nature Reviews Neuroscience*, *10*(2), 141–52.

Marchand, S., Kupers, R. C., Bushnell, M. C., & Duncan, G. H. (2003). Analgesic and placebo effects of thalamic stimulation. *Pain*, *105*(3), 481–488.

Margulies, D. S., Kelly, M. C., Uddin, L. Q., Biswal, B. B., Castellanos, F. X., & Milham, M. P. (2007). Mapping the functional connectivity of anterior cingulate cortex. *NeuroImage*, *37*(2), 579–88.

Mars, R. B., Klein, M. C., Neubert, F.-X., Olivier, E., Buch, E. R., Boorman, E. D., & Rushworth, M. F. S. (2009). Short-latency influence of medial frontal cortex on primary motor cortex during action selection under conflict. *The Journal of Neuroscience*, *29*(21), 6926–31.

Masterman, D. L., & Cummings, J. L. (1997). Frontal-subcortical circuits: the anatomic basis of executive, social and motivated behaviors. *Journal of Psychopharmacology*, *11*(2), 107–114.

Matsumoto, M., Matsumoto, K., Abe, H., & Tanaka, K. (2007). Medial prefrontal cell activity signaling prediction errors of action values. *Nature neuroscience*, *10*(5), 647–56.

Mayberg, H. S., Lozano, A. M., Voon, V., McNeely, H. E., Seminowicz, D., Hamani, C., Kennedy, S. H. (2005). Deep brain stimulation for treatment-resistant depression. *Neuron*, *45*(5), 651–60.

Mayr, U., Awh, E., & Laurey, P. (2003). Conflict adaptation effects in the absence of executive control. *Nature neuroscience*, *6*(5), 450–2.

Mead, L., Mayer, A. R., Bobholz, J., Woodley, S. J., Cunningham, J. M., Hammeke, T., & Rao, S. M. (2002). Neural basis of the Stroop interference task: response competition or selective attention? *Journal of the International Neuropsychological Society*, *8*(6), 735–42.

Mian, M. K., Sheth, S., Patel, S. R., Spiliopoulos, K., Eskandar, E. N., & Williams, Z. M. (2012). Encoding of Rules by Neurons in the Human Dorsolateral Prefrontal Cortex. *Cerebral cortex*, *6*, 1–10.

Milham, M. P., Banich, M. T., Claus, E., & Cohen, N. J. (2003). Practice-related effects demonstrate complementary roles of anterior cingulate and prefrontal cortices in attentional control. *NeuroImage*, *18*(2), 483–493.

Milham, M. P., Banich, M. T., Webb, A., Barad, V., Cohen, N. J., Wszalek, T., & Kramer, A. F. (2001). The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on nature of conflict. *Brain research. Cognitive brain research*, *12*(3), 467–73.

Miller, E. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *The Journal of Neuroscience*, *16*, 5154–5167.

Miller, E., & Cohen, J. (2001). An Integrative Theory of Prefrontal Cortex Function. *Annual review of neuroscience*, 167–202.

Miller, E., & Desimone, R. (1994). Parallel neuronal mechanisms for short-term memory. *Science*, *263*, 520–522.

Milner, B. (1963). Effects of different brain lesions on card sorting: The role of the frontal lobes. *Archives of Neurology*.

Milner, B. (1982). Some cognitive effects of frontal-lobe lesions in man. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *298*(1089), 211–26.

Mitchell, D. G. V., Rhodes, R., Pine, D. S., & Blair, R. J. R. (2008). The contribution of ventrolateral and dorsolateral prefrontal cortex to response reversal. *Behavioural brain research*, *187*(1), 80–7.

Monchi, O., Petrides, M., Petre, V., Worsley, K., & Dagher, A. (2001). Wisconsin Card Sorting revisited: distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. *The Journal of Neuroscience*, *21*(19), 7733–41.

Nachev, P., Rees, G., & Parton, A. (2005). Volition and conflict in human medial frontal cortex. *Current biology*, *15*, 122–128.

Nakamura, K., Roesch, M. R., & Olson, C. R. (2005). Neuronal activity in macaque SEF and ACC during performance of tasks involving conflict. *Journal of Neurophysiology*, *93*, 884–908.

Nee, D. E., Kastner, S., & Brown, J. W. (2010). Functional heterogeneity of conflict, error, task-switching, and unexpectedness effects within medial prefrontal cortex. *NeuroImage*, 1–13.

Neisser, U. (1976). *Cognition and Reality*. San Francisco, CA: W.H. Freeman.

Newman, L., & McGaughy, J. (2011). Attentional effects of lesions to the anterior cingulate cortex: how prior reinforcement influences distractibility. *Behavioral neuroscience*, *125*(3), 360–71.

Nieuwenhuis, S., & Stins, J. (2006). Accounting for sequential trial effects in the flanker task: Conflict adaptation or associative priming? *Memory & Cognition*, *34*, 1260–1272.

Nieuwenhuis, S., Yeung, N., van den Wildenberg, W., & Ridderinkhof, K. R. (2003). Electrophysiological correlates of anterior cingulate function in a go/no-go task: effects of response conflict and trial type frequency. *Cognitive, affective & behavioral neuroscience*, *3*(1), 17–26.

Nimchinsky, E., Gilissen, E., Allman, J. M., Perl, D. P., Erwin, J. M., & Hof, P. R. (1999). A neuronal morphologic type unique to humans and great apes. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(9), 5268–73.

Nishijo, H., Yamamoto, Y., Ono, T., Uwano, T., Yamashita, J., & Yamashima, T. (1997). Single neuron responses in the monkey anterior cingulate cortex during visual discrimination. *Neuroscience letters*, *227*(2), 79–82.

Niv, Y., Duff, M. O., & Dayan, P. (2005). Dopamine, uncertainty and TD learning. *Behavioral and brain functions*, *1*(6).

Nolte, J. (2009). *The Human Brain: An Introduction to its Functional Neuroanatomy.* Philadelphia, PA: Mosby, Elsevier.

Norman, D.A., & Shallice, T. (1980). *Attention to Action: Willed and Automatic Control of Behavior.* University of California San Diego, Center for Human Information Processing Report, 99.

Notebaert, W., & Verguts, T. (2007). Dissociating conflict adaptation from feature integration: a multiple regression approach. *Journal of Experimental Psychology. Human Perception and Performance*, *33*(5), 1256–60.

Ongür, D., Ferry, A. T., & Price, J. L. (2003). Architectonic subdivision of the human orbital and medial prefrontal cortex. *The Journal of comparative neurology*, *460*(3), 425–49.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational intelligence and neuroscience*, *2011*, 156869.

Owen, A., Downes, J., & Sahakian, B. (1990). Planning and Spatial Working Memory following Frontal Lobe Lesion in Man. *Neuropsychologia*, *28*(1), 1021–1034.

Owen, S. L. F., Green, A. L., Stein, J. F., & Aziz, T. Z. (2006). Deep brain stimulation for the alleviation of post-stroke neuropathic pain. *Pain*, *120*(1-2), 202–6.

Papez, J. (1937). A Proposed Mechanism of Emotion. *Archives of neurology and psychiatry*, *258*(March 5), 589–598.

Pascual-Leone, a, Walsh, V., & Rothwell, J. (2000). Transcranial magnetic stimulation in cognitive neuroscience--virtual lesion, chronometry, and functional connectivity. *Current opinion in neurobiology*, *10*(2), 232–7.

Passingham, R. E. (2002). The frontal cortex: does size matter? *Nature Neuroscience*, *5*(3), 190–2.

Pasupathy, A., & Miller, E. (2005). Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature*, *433*, 873–876.

Paus, T. (2001). Primate anterior cingulate cortex: where motor control, drive and cognition interface. *Nature Reviews Neuroscience*, *2*(6), 417–24.

Paus, T., Castro-Alamancos, M., & Petrides, M. (2000). Cortico-cortical connectivity of the human mid-dorsolateral frontal cortex and its modulation by repetitive transcranial magnetic stimulation: a combined TMS/PET study. *NeuroImage*, *11*(5).

Paus, T., Tomaiuolo, F., & Otaky, N. (1996). Human cingulate and paracingulate sulci: pattern, variability, asymmetry, and probabilistic map. *Cerebral Cortex*, *6*(2), 207–214.

Perini, I., Bergstrand, S., & Morrison, I. (2013). Where pain meets action in the human brain. *The Journal of Neuroscience, 33*(40), 15930–9.

Perlmutter, J., & Mink, J. (2006). Deep brain stimulation. *Annual Review Neuroscience, 29,* 229-257.

Peterson, B., Kane, M., & Alexander, G. (2002). An event-related functional MRI study comparing interference effects in the Simon and Stroop tasks. *Cognitive Brain Research, 13*, 427–440.

Petrides, M. (2000). The role of the mid-dorsolateral prefrontal cortex in working memory. *Experimental brain research*, *133*(1), 44–54.

Petrides, M. (2005). Lateral prefrontal cortex: architectonic and functional organization. *Philosophical transactions of the Royal Society of London. Biological sciences*, *360*(1456), 781–95.

Petrides, M. (2007). Selection between Competing Responses Based on Conditional Rules. In S. A. Bunge & J. D. Wallis (Eds.), *Neuroscience of Rule-Guided Behaviour*. Oxford: Oxford University Press.

Petrides, M., & Pandya, D. N. (1994). Comparative architectonic analysis of the human and the macaque frontal cortex. *Handbook of Neuropsychology, 9*, 17-58.

Petrides, M., & Pandya, D. N. (1999). Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. *The European journal of neuroscience*, *11*(3), 1011–36.

Petrides, M., & Pandya, D. N. (2002). Association Pathways of the Prefrontal Cortex and Functional Observations. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function*, 1–34. Oxford: Oxford University Press.

Petrides, M., & Pandya, D. N. (2007). Efferent association pathways from the rostral prefrontal cortex in the macaque monkey. *The Journal of Neuroscience*, *27*(43), 11573–86.

Petrides, M., Tomaiuolo, F., Yeterian, E. H., & Pandya, D. N. (2012). The prefrontal cortex: comparative architectonic organization in the human and the macaque monkey brains. *Cortex*, *48*(1), 46–57.

Pew, R. (1969). The speed-accuracy operating characteristic. *Acta Psychologica*, 16–26.

Picard, N., & Strick, P. L. (1996). Motor areas of the medial wall: a review of their location and functional activation. *Cerebral cortex*, *6*(3), 342–53.

Poldrack, R., Sabb, F. W., Foerde, K., Tom, S. M., Asarnow, R. F., Bookheimer, S. Y., & Knowlton, B. J. (2005). The neural correlates of motor skill automaticity. *The Journal of Neuroscience*, *25*(22), 5356–64.

Posner, M.I. (1975). Attention and Cognitive Control. In R.L. Solso (Ed.), *Information Processing and Cognition: The Loyola Symposium* (pp. 55-85). Hillsdale, NJ: Erlbaum.

Praamstra, P., Kleine, B. U., & Schnitzler, A. (1999). Magnetic stimulation of the dorsal premotor cortex modulates the Simon effect. *Neuroreport*, *10*(17), 3671–4.

Prinzmetal, W., Taylor, J., Myers, L. B., & Nguyen-Espino, J. (2011). Contingent capture and inhibition of return: a comparison of mechanisms. *Experimental brain research*, *214*(1), 47–60.

Purmann, S., Badde, S., & Wendt, M. (2009). Adjustments to recent and frequent conflict reflect two distinguishable mechanisms. *Psychonomic bulletin & review*, *16*(2), 350–5.

Rabbitt, P. (1966). Errors and Error Correction in Choice-Response Tasks. *Journal of experimental psychology*, *71*(2), 264–272.

Rajkowska, G., & Goldman-Rakic, P. S. (1995). Cytoarchitectonic definition of prefrontal areas in the normal human cortex: Remapping of areas 9 and 46 using quantitative criteria. *Cerebral cortex, 5*(4), 307–22.

Ramnani, N., & Miall, R. C. (2003). Instructed delay activity in the human prefrontal cortex is modulated by monetary reward expectation. *Cerebral cortex, 13*(3), 318–27.

Ramnani, N., & Passingham, R. E. (2001). Changes in the human brain during rhythm learning. *Journal of cognitive neuroscience*, *13*(7), 952–66.

Ramnani, N., & Owen, A. M. (2004). Anterior prefrontal cortex: insights into function from anatomy and neuroimaging. *Nature Reviews Neuroscience*, *5*(3), 184–94.

Ratcliff, R., & Rouder, J. N. (1998). Modeling Response Times for Two-Choice Decisions. *Psychological Science*, *9*(5), 347–356.

Ratcliff, R. (1978). A Theory of Memory Retrieval. *Psychological review, 85*(2), 59-108.

Ratcliff, R., & Tuerlinck, F. (2002). Estimating parameters of the diffusion model: approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic bulletin & review*, *9*(3), 438–81.

Raye, C. L., Johnson, M. K., Mitchell, K. J., Reeder, J., & Greene, E. J. (2002). Neuroimaging a single thought: dorsolateral PFC activity associated with refreshing just-activated information. *NeuroImage*, *15*(2), 447–53.

Reddi, B. (2001). Decision making: the two stages of neuronal judgement. *Current biology*, *11*(15), R603–6.

Reddi, B., & Carpenter, R. H. (2000). The influence of urgency on decision time. *Nature Neuroscience*, *3*(8), 827–30.

Remondes, M., & Wilson, M. A. (2013). Cingulate-Hippocampus Coherence and Trajectory Coding in a Sequential Choice Task. *Neuron*, *80*(5), 1277–1289.

Reynolds, J. R., McDermott, K. B., & Braver, T. S. (2006). A direct comparison of anterior prefrontal cortex involvement in episodic retrieval and integration. *Cerebral cortex*, *16*(4), 519–28.

Ridderinkhof, K., & Ullsperger, M. (2004). The role of the medial frontal cortex in cognitive control. *Science, 306*, 443-447.

Roth, J. K., Johnson, M. K., Raye, C. L., & Constable, R. T. (2009). Similar and dissociable mechanisms for attention to internal versus external information. *NeuroImage*, *48*(3), 601–8.

Rudebeck, P. H., Behrens, T. E., Kennerley, S. W., Baxter, M. G., Buckley, M. J., Walton, M. E., & Rushworth, M. F. S. (2008). Frontal cortex subregions play distinct roles in choices between actions and stimuli. *The Journal of neuroscience*, *28*(51), 13775–85.

Rugg, M. D., Fletcher, P. C., Allan, K., Frith, C. D., Frackowiak, R. S., & Dolan, R. J. (1998). Neural correlates of memory retrieval during recognition memory and cued recall. *NeuroImage*, *8*(3), 262–73.

Rushworth, M. F. S., Walton, M. E., Kennerley, S. W., & Bannerman, D. M. (2004). Action sets and decisions in the medial frontal cortex. *Trends in cognitive sciences*, *8*(9), 410–7.

Rushworth, M. F. S., & Behrens, T. E. J. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature neuroscience*, *11*(4), 389–397.

Sasaki, K., & Gemba, H. (1986). Electrical activity in the prefrontal cortex specific to no-go reaction of conditioned hand movement with colour discrimination in the monkey. *Experimental Brain Research*, 603–606.

Schall, J. D., & Emeric, E. E. (2010). Conflict in cingulate cortex function between humans and macaque monkeys: More apparent than real. *Brain, behavior and evolution*, *75*(4), 237–8.

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron*, *36*(2), 241–63.

Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annual review of psychology*, *57*, 87–115.

Semendeferi, K., Armstrong, E., Schleicher, A., Zilles, K., & Van Hoesen, G. W. (2001). Prefrontal cortex in humans and apes: a comparative study of area 10. *American journal of physical anthropology*, *114*(3), 224–41.

Semendeferi, K., Lu, a, Schenker, N., & Damasio, H. (2002). Humans and great apes share a large frontal cortex. *Nature neuroscience*, *5*(3), 272–6.

Shallice, T., & Burgess, P. (1991). Deficits in Strategy Application following Prefrontal Damage in Man. *Brain, 114*, 727–741.

Sheth, S., Mian, M. K., Patel, S. R., Asaad, W. F., Williams, Z. M., Dougherty, D. D., Eskandar, E. N. (2012). Human dorsal anterior cingulate cortex neurons mediate ongoing behavioural adaptation. *Nature*, *488*(7410), 218–21.

Shiffrin, R., & Schneider, W. (1977). Controlled and Automatic Human Information Processing: II. Perceptual Learning, Automatic Attending, and a General Theory. *Psychological review, 84*(2), 127-190.

Silton, R. L., Heller, W., Towers, D. N., Engels, A. S., Spielberg, J. M., Edgar, J. C., Miller, G. a. (2010). The time course of activity in dorsolateral prefrontal cortex and anterior cingulate cortex during top-down attentional control. *NeuroImage*, *50*(3), 1292–302.

Simon, J. R. (1969). Reactions toward the source of stimulation. *Journal of experimental psychology*, *81*, 174–176.

Simon, J. R. (1990). Effect of conflicting cues on information processing: The “Stroop effect” vs. the “Simon effect.” *Acta Psychologica*, *73*, 159–170.

Simon, J. R., & LeMay, R. P. (1969). Temporal and symbolic S-R compatibility in a sequential information-processing task. *Journal of Experimental Psychology*, *80*(3), 558–560.

Simon, J. R., & Small, A. M. (1969). Processing auditory information: interference from an irrelevant cue. *The Journal of applied psychology*, *53*(5), 433–5.

Simon, J. R., Sly, P., & Vilapakkam, S. (1981). Effect of compatibility of SR mapping on reactions toward the stimulus source. *Acta Psychologica*, *47*, 63–81.

Smith, A., Li, M., Becker, S., & Kapur, S. (2006). Dopamine, prediction error and associative learning: a model-based account. *Network*, *17*(1), 61–84.

Smith, P. L., & Ratcliff, R. (2009). An integrated theory of attention and decision making in visual signal detection. *Psychological review*, *116*(2), 283–317.

Sohn, M.-H., Albert, M. V, Jung, K., Carter, C. S., & Anderson, J. R. (2007). Anticipation of conflict monitoring in the anterior cingulate cortex and the prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(25), 10330–4.

Soutschek, A., Taylor, P. C. J., Müller, H. J., & Schubert, T. (2013). Dissociable networks control conflict during perception and response selection: a transcranial magnetic stimulation Study. *The Journal of neuroscience*, *33*(13), 5647–54.

Specht, K., Lie, C.-H., Shah, N. J., & Fink, G. R. (2009). Disentangling the prefrontal network for rule selection by means of a non-verbal variant of the Wisconsin Card Sorting Test. *Human brain mapping*, *30*(5), 1734–43.

Stoffels, E. J. (1996). On stage robustness and response selection routes: Further evidence. *Acta Psychologica*, *91*, 67–88.

Stokes, M. G., Barker, A. T., Dervinis, M., Verbruggen, F., Maizey, L., Adams, R. C., & Chambers, C. D. (2013). Biophysical determinants of transcranial magnetic stimulation: effects of excitability and depth of targeted area. *Journal of neurophysiology*, *109*(2), 437–44.

Stokes, M. G., Chambers, C. D., Gould, I. C., English, T., McNaught, E., McDonald, O., & Mattingley, J. B. (2007). Distance-adjusted motor threshold for transcranial magnetic stimulation. *Clinical neurophysiology*, *118*(7), 1617–25.

Stokes, M. G., Chambers, C. D., Gould, I. C., Henderson, T. R., Janko, N. E., Allen, N. B., & Mattingley, J. B. (2005). Simple metric for scaling motor threshold based on scalp-cortex distance: application to studies using transcranial magnetic stimulation. *Journal of neurophysiology*, *94*(6), 4520–7.

Stone, M. (1960). Models for Choice-Reaction Time. *Psychometrika*, *25*(3), 251–260.

Strange, B., Henson, R. N., Friston, K. J., & Dolan, R. J. (2001). Anterior prefrontal cortex mediates rule learning in humans. *Cerebral cortex*, *11*(11), 1040–6.

Stroop, J.R. (1935). Studies of Interference in Serial Verbal Actions. *Journal of Experimental Psychology, 18*, 643-662.

Stürmer, B., & Leuthold, H. (2003). Control over response priming in visuomotor processing: a lateralized event-related potential study. *Experimental brain research*, *153*(1), 35–44.

Sturmer, B., Leuthold, H., Soetens, E., Schro, H., & Sommer, W. (2002). Control over location-based response activation in the Simon task: Behavioral and

electrophysiological evidence. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(6), 1345–1363.

Stürmer, B., Redlich, M., Irlbacher, K., & Brandt, S. (2007). Executive control over response priming and conflict: a transcranial magnetic stimulation study. *Experimental brain research*, *183*(3), 329–39.

Stürmer, B., Siggelkow, S., Dengler, R., & Leuthold, H. (2000). Response priming in the Simon paradigm. *Experimental Brain Research*, *135*(3), 353–359.

Stuss, D. T., Floden, D., Alexander, M. P., Levine, B., & Katz, D. (2001). Stroop performance in focal lesion patients: dissociation of processes and frontal lobe lesion location. *Neuropsychologia*, *39*(8), 771–86.

Stuss, D. T., Kaplan, E. F., Benson, D. F., Weir, W. S., Chiulli, S., & Sarazin, F. F. (1982). Evidence for the involvement of orbitofrontal cortex in memory functions: an interference effect. *Journal of comparative and physiological psychology*, *96*(6), 913–25.

Suri, R., & Schultz, W. (1999). A neural network model with dopamine-like reinforcement signal that learns a spatial delayed response task. *Neuroscience*, *91*, 871.

Swick, D., & Jovanovic, J. (2002). Anterior cingulate cortex and the Stroop task: Neuropsychological evidence for topographic specificity. *Neuropsychologia*, *40*, 1240–1253.

Swick, D., & Turken, A. U. (2002). Dissociation between conflict detection and error monitoring in the human anterior cingulate cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(25), 16354–9.

Tagliabue, M., Zorzi, M., Umiltà, C., & Bassignani, F. (2000). The role of long-term-memory and short-term-memory links in the Simon effect. *Journal of experimental psychology. Human perception and performance*, *26*(2), 648–70.

Tanaka, K., Matsumoto, K., Mansouri, F., & Buckley, M. J. (2012). Functional division among monkey prefrontal areas in goal-directed behavior, 249–258.

Taylor, P. C. J., Nobre, A. C., & Rushworth, M. F. S. (2007). Subsecond changes in top down control exerted by human medial frontal cortex during conflict and action selection: a combined transcranial magnetic stimulation electroencephalography study. *The Journal of neuroscience*, *27*(42), 11343–53.

Thompson-Schill, S. L., Jonides, J., Marshuetz, C., Smith, E. E., D'Esposito, M., Kan, I. P., Swick, D. (2002). Effects of frontal lobe damage on interference effects in working memory. *Cognitive, affective & behavioral neuroscience*, *2*(2), 109–20.

Totah, N. K. B., Jackson, M. E., & Moghaddam, B. (2013). Preparatory Attention Relies on Dynamic Interactions between Prelimbic Cortex and Anterior Cingulate Cortex, *Cerebral Cortex, 23*(3), 729–738.

Treichler, F. (1973). Effects of extensive training on object reversal in frontal monkeys. *Neuropsychologia*, *11,* 57-65.

Tsujimoto, S., Genovesio, A., & Wise, S. P. (2010). Evaluating self-generated decisions in frontal pole cortex of monkeys. *Nature neuroscience*, *13*(1), 120–6.

Ullsperger, M., Bylsma, L. M., & Botvinick, M. M. (2005). The conflict adaptation effect: It's not just priming. *Cognitive, Affective, & Behavioral Neuroscience*, *5*(4), 467-472.

Usher, M., & McClelland, J. (2001). The Time-Course of Perceptual Choice: The Leaky Competing Accumulator Model. *Psychological Review, 108*(3), 550-592.

Van Maanen, L., Brown, S. D., Eichele, T., Wagenmakers, E.-J., Ho, T., Serences, J., & Forstmann, B. U. (2011). Neural correlates of trial-to-trial fluctuations in response caution. *The Journal of neuroscience*, *31*(48), 17488–95.

Van Veen, V., Cohen, J. D., Botvinick, M. M., Stenger, V. A., & Carter, C. S. (2001). Anterior cingulate cortex, conflict monitoring, and levels of processing. *NeuroImage*, *14*(6), 1302–8.

Van Veen, V., & Carter, C. S. (2002). The timing of action-monitoring processes in the anterior cingulate cortex. *Journal of cognitive neuroscience*, *14*(4), 593–602.

Van Veen, V., & Carter, C. S. (2005). Separating semantic conflict and response conflict in the Stroop task: a functional MRI study. *NeuroImage*, *27*(3), 497–504.

Vanderhasselt, M.-A., De Raedt, R., Baeken, C., Leyman, L., Clerinx, P., & D'haenen, H. (2007). The influence of rTMS over the right dorsolateral prefrontal cortex on top-down attentional processes. *Brain research*, *1137*(1), 111–6.

Vanderhasselt, M.-A., De Raedt, R., Baeken, C., Leyman, L., & D'haenen, H. (2006). The influence of rTMS over the left dorsolateral prefrontal cortex on Stroop task performance. *Experimental brain research*, *169*(2), 279–82.

Vendrell, P., Junqué, C., Pujol, J., & Jurado, M. (1995). The role of prefrontal regions in the Stroop task. *Neuropsychologia*, *33*(3), 341–352.

Verbruggen, F., Notebaert, W., Liefooghe, B., & Vandierendonck, A. (2006). Stimulus- and response-conflict-induced cognitive control in the flanker task. *Psychonomic bulletin & review*, *13*(2), 328–33.

Vogt, B., Nimchinsky, E. a, Vogt, L. J., & Hof, P. R. (1995). Human cingulate cortex: surface features, flat maps, and cytoarchitecture. *The Journal of comparative neurology*, *359*(3), 490–506.

Wagenmakers, E.-J., Krypotos, A.-M., Criss, A. H., & Iverson, G. (2012). On the interpretation of removable interactions: a survey of the field 33 years after Loftus. *Memory & cognition*, *40*(2), 145–60.

Wagenmakers, E.-J., van der Maas, H. L. J., & Grasman, R. P. P. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic bulletin & review*, *14*(1), 3–22.

Wagner, M., Rihs, T., Mosimann, U. P., Fisch, H. U., & Schlaepfer, T. E. (2006). Repetitive transcranial magnetic stimulation of the dorsolateral prefrontal cortex affects divided attention immediately after cessation of stimulation. *Journal of psychiatric research*, *40*(4), 315–21.

Wallis, J. D., Anderson, K. C., & Miller, E. K. (2001). Single neurons in prefrontal cortex encode abstract rules. *Nature*, *411*(6840), 953–6.

Wallis, J. D, & Miller, E. K. (2003). From rule to response: neuronal processes in the premotor and prefrontal cortex. *Journal of neurophysiology*, *90*(3), 1790–806.

Walsh, K.W. (1978). *Neuropsychology: A Clinical Approach.* Edinburgh: Churchill Livingstone.

Walsh, V., & Cowey, A. (2000). Transcranial magnetic stimulation and cognitive neuroscience. *Nature reviews. Neuroscience*, *1*(1), 73–9.

Walsh, V., & Rushworth, M. (1999). A primer of magnetic stimulation as a tool for neuropsychology. *Neuropsychologia*, *37*(2), 125–35.

Walton, M. E., Croxson, P. L., Behrens, T. E. J., Kennerley, S. W., & Rushworth, M. F. S. (2007). Adaptive decision making and value in the anterior cingulate cortex. *Neuroimage, 36,* 142-54.

Wang, C., Ulbert, I., Schomer, D. L., Marinkovic, K., & Halgren, E. (2005). Responses of human anterior cingulate cortex microdomains to error detection, conflict monitoring, stimulus-response mapping, familiarity, and orienting. *The Journal of Neuroscience*, *25*(3), 604–13.

Wang, X. (2006). A microcircuit model of prefrontal functions: ying and yang of reverberatory neurodynamics in cognition. In J. Risberg & J. Grafman (Eds.), *The frontal lobes*. Cambridge: Cambridge University Press.

Wang, Y., Tian, S., Wang, H., Cui, L., Zhang, Y., & Zhang, X. (2003). Event-related potentials evoked by multi-feature conflict under different attentive conditions. *Experimental brain research*, *148*(4), 451–7.

Ward, R., & Ward, R. (2006). Cognitive conflict without explicit conflict monitoring in a dynamical agent. *Neural networks : the official journal of the International Neural Network Society*, *19*(9), 1430–6.

Weissman, D. ., Giesbrecht, B., Song, A. ., Mangun, G. ., & Woldorff, M. . (2003). Conflict monitoring in the human anterior cingulate cortex during selective attention to global and local object features. *NeuroImage*, *19*(4), 1361–1368.

Wendt, M., Heldmann, M., Münte, T. F., & Kluwe, R. H. (2007). Disentangling sequential effects of stimulus- and response-related conflict and stimulus-response repetition using brain potentials. *Journal of cognitive neuroscience*, *19*(7), 1104–12.

Wendt, M., Kluwe, R. H., & Peters, A. (2006). Sequential modulations of interference evoked by processing task-irrelevant stimulus features. *Journal of experimental psychology. Human perception and performance*, *32*(3), 644–67.

West, R., Jakubek, K., Wymbs, N., Perry, M., & Moore, K. (2005). Neural correlates of conflict processing. *Experimental brain research*, *167*(1), 38–48.

West, R., & Travers, S. (2008). Tracking the temporal dynamics of updating cognitive control: an examination of error processing. *Cerebral cortex*, *18*(5), 1112–24.

White, I. M., & Wise, S. P. (1999). Rule-dependent neuronal activity in the prefrontal cortex. *Experimental brain research*, *126*(3), 315–35.

Wickelgren, W. (1977). Speed-Accuracy Tradeoff and Information Processing Dynamics. *Acta psychologica, 41,* 67-85.

Wilk, H., Ezekiel, F., & Morton, J. B. (2012). Brain regions associated with moment-to-moment adjustments in control and stable task-set maintenance. *NeuroImage*, *59*(2), 1960–7.

Wittfoth, M., Schardt, D. M., Fahle, M., & Herrmann, M. (2009). How the brain resolves high conflict situations: double conflict involvement of dorsolateral prefrontal cortex. *NeuroImage*, *44*(3), 1201–9.

Wood, J., & Grafman, J. (2003). Human prefrontal cortex: processing and representational perspectives. *Nature Reviews Neuroscience*, *4*, 139–147.

Wühr, P. (2005). Evidence for gating of direct response activation in the Simon task. *Psychonomic bulletin & review*, *12*(2), 282–8.

Wühr, P., & Ansorge, U. (2005). Exploring trial-by-trial modulation of the Simon effect. *Psychonomic bulletin & review*, *12*(2), 282–8.

Yeterian, E. H., Pandya, D. N., Tomaiuolo, F., & Petrides, M. (2012). The cortical connectivity of the prefrontal cortex in the monkey brain. *Cortex*, *48*(1), 58–81.

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychological Review*, *111*, 931–959.

Yeung, N., & Cohen, J. D. (2006). The impact of cognitive deficits on conflict monitoring. Predictable dissociations between the error-related negativity and N2. *Psychological science*, *17*(2), 164–71.

Yoshida, W., & Ishii, S. (2006). Resolution of uncertainty in prefrontal cortex. *Neuron*, *50*(5), 781–9.

APPENDIX

APPENDIX 1

Appendix 1a: Location of lesions in the ACC group illustrated on standard MNI brain in sagittal views.



Figure A1. ***ACC group*** *– Patient 1.*

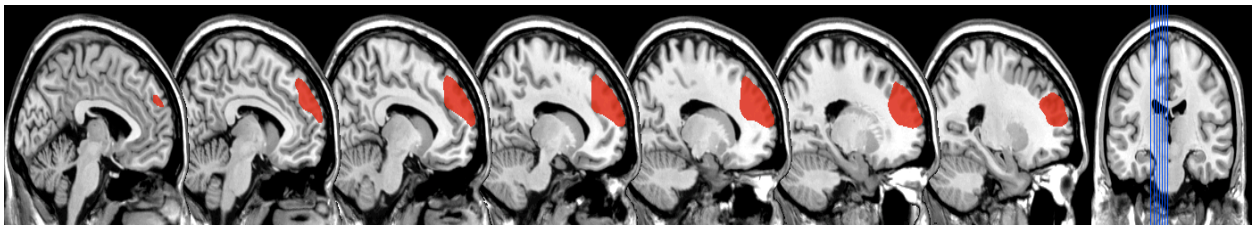


Figure A2. ***ACC group*** *– Patient 2.*

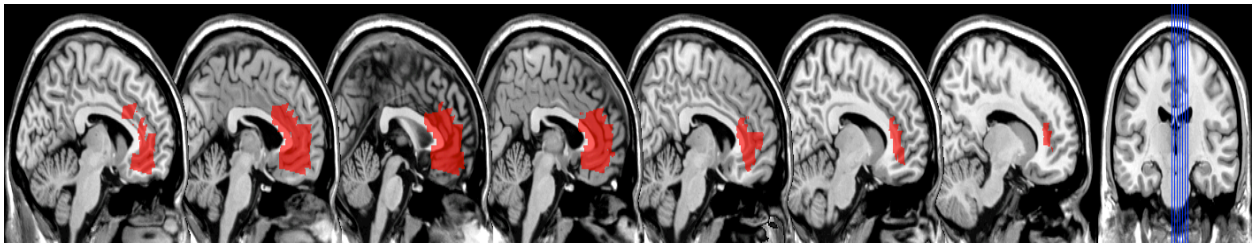


Figure A3. *ACC group – Patient 3.*

Appendix 1b: Location of lesions in the dlPFC group illustrated on standard MNI brain in horizontal views.

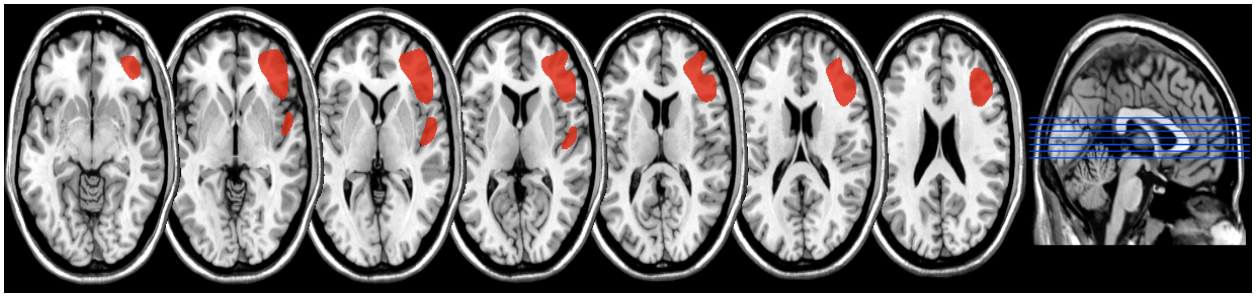


Figure A4. ***dlPFC group*** – *Patient 4.*

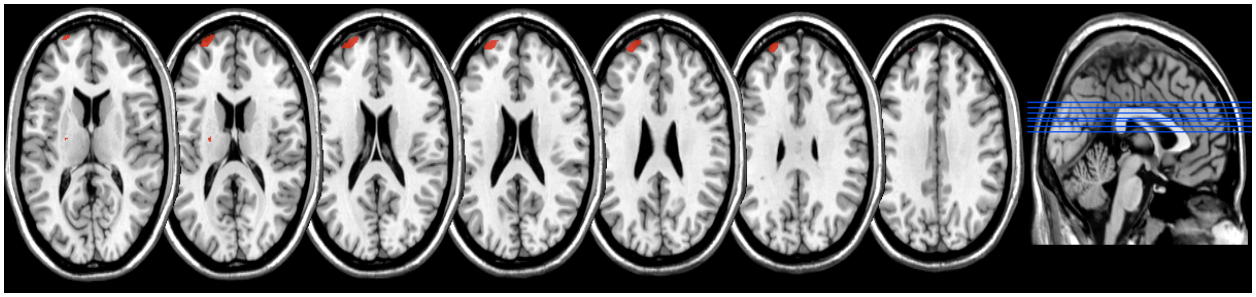


Figure A5. ***dlPFC group*** – *Patient 5.*

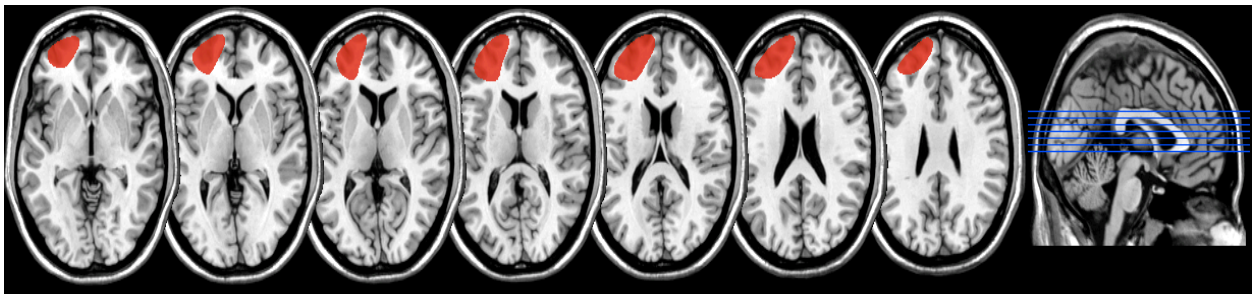


Figure A6. ***dlPFC group*** – *Patient 6.*

Appendix 1c: Location of lesions in the OFC group illustrated on standard MNI brain in coronal views.

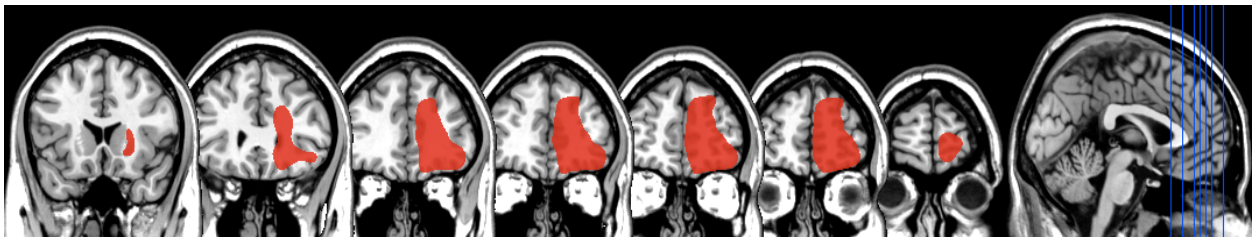


Figure A7. ***OFC group*** – *Patient 7.*

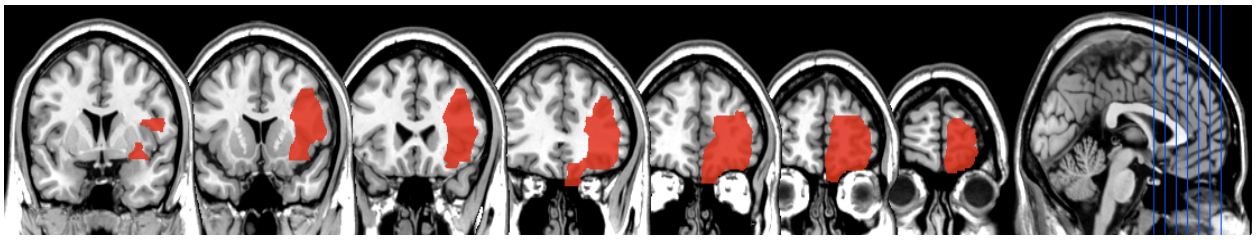


Figure A8. ***OFC group*** – *Patient 8.*

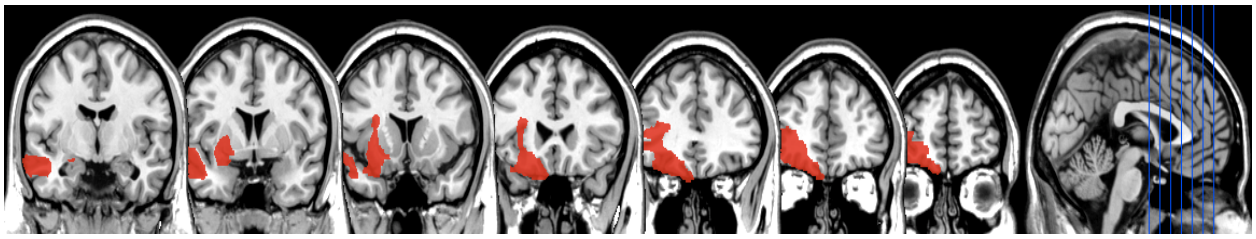


Figure A9. ***OFC group*** – *Patient 9.*

APPENDIX 2

Appendix 2a: Lesion information. Volumes (in mm^3) for OFC, dlPFC and ACC patients.

REGIONS OF INTEREST (ROI)																				
	DLPFC							OFC							ACC					
	Brodmann areas																			
ID	8L	8R	9L	9R	46L	46R	TOTAL	10L	10R	11L	11R	47L	47R	TOTAL	24L	24R	32L	32R	TOTAL	TOTAL ROI	TOTAL VOL
1	0	6072	0	2880	0	104	9056	0	0	0	0	0	0	0	0	1272	0	5048	6320	15376	30192
2	96	0	5416	0	1520	0	7032	5256	0	32	0	72	0	5360	0	0	4128	0	4128	16520	20864
3	0	0	0	0	0	0	0	1555	1530	3814	3178	1481	0	11558	2285	1540	1489	2001	7315	18873	39014
4	0	136	0	2432	0	5320	7888	0	904	0	1072	0	6576	8552	0	0	0	0	0	16440	49984
5	0	0	32	0	464	0	496	632	0	0	0	0	0	632	0	0	0	0	0	1128	1384
6	0	0	3144	0	6760	0	9904	9008	0	4216	0	3104	0	16328	0	0	1040	0	1040	27272	34960
7	0	0	0	2496	0	3296	5792	0	9736	0	16520	12024	0	38280	0	16	0	5824	5840	49912	75616
8	0	0	0	8	0	5072	5080	0	9472	0	19440	0	5072	33984	0	0	0	984	984	40048	103408
9	0	0	0	0	1880	0	1880	880	0	13088	0	10416	0	24384	0	0	0	0	0	26264	59928

Lesion volumes in cubic mm

Table A1. *For each patient (ACC group = green, dlPFC group = blue, OFC group = pink), lesion volumes in mm^3 are reported for each general region of interest (dlPFC, OFC and ACC) and for specific Brodmann areas within the ROIs. Total ROI volume in the second rightmost column is calculated as the sum of the volume of damage across all general ROIs. Total volume in the rightmost column refers to the total volume of damage across the whole brain.*

Appendix 2b: Lesion information. Proportions (%) for OFC, dlPFC and ACC patients.

Overall regions of interest

ID	OUT of ROIs VOLUME			OUT OF TOTAL VOLUME		
	DLPFC	OFC	ACC	DLPFC	OFC	ACC
1	59%	0%	41%	30%	0%	21%
2	43%	32%	25%	34%	26%	20%
3	0%	61%	39%	0%	30%	19%
4	48%	52%	0%	16%	17%	0%
5	44%	56%	0%	36%	46%	0%
6	36%	60%	4%	28%	47%	3%
7	12%	77%	12%	8%	51%	8%
8	13%	85%	2%	5%	33%	1%
9	7%	93%	0%	3%	41%	0%

Table A2. *Percentage of lesion localized to each specific ROIs out of the lesion volume across all general ROIs (left) and out of total brain damage (right) for each patient.*

Specific Brodmann areas of interest

ID	OUT OF ROIs VOLUME			OUT OF TOTAL VOLUME		
	46	11	24+32	46	11	24+32
1	1%	0%	41%	0%	0%	21%
2	9%	0%	25%	7%	0%	20%
3	0%	37%	39%	0%	18%	19%
4	32%	7%	0%	11%	2%	0%
5	41%	0%	0%	34%	0%	0%
6	25%	15%	4%	19%	12%	3%
7	7%	33%	12%	4%	22%	8%
8	13%	49%	2%	5%	19%	1%
9	7%	50%	0%	3%	22%	0%

Table A3. *Percentage of lesion localized to each specific Brodmann area of interest out of the lesion volume across all general ROIs (left) and out of total brain damage (right) for each patient.*

APPENDIX 3

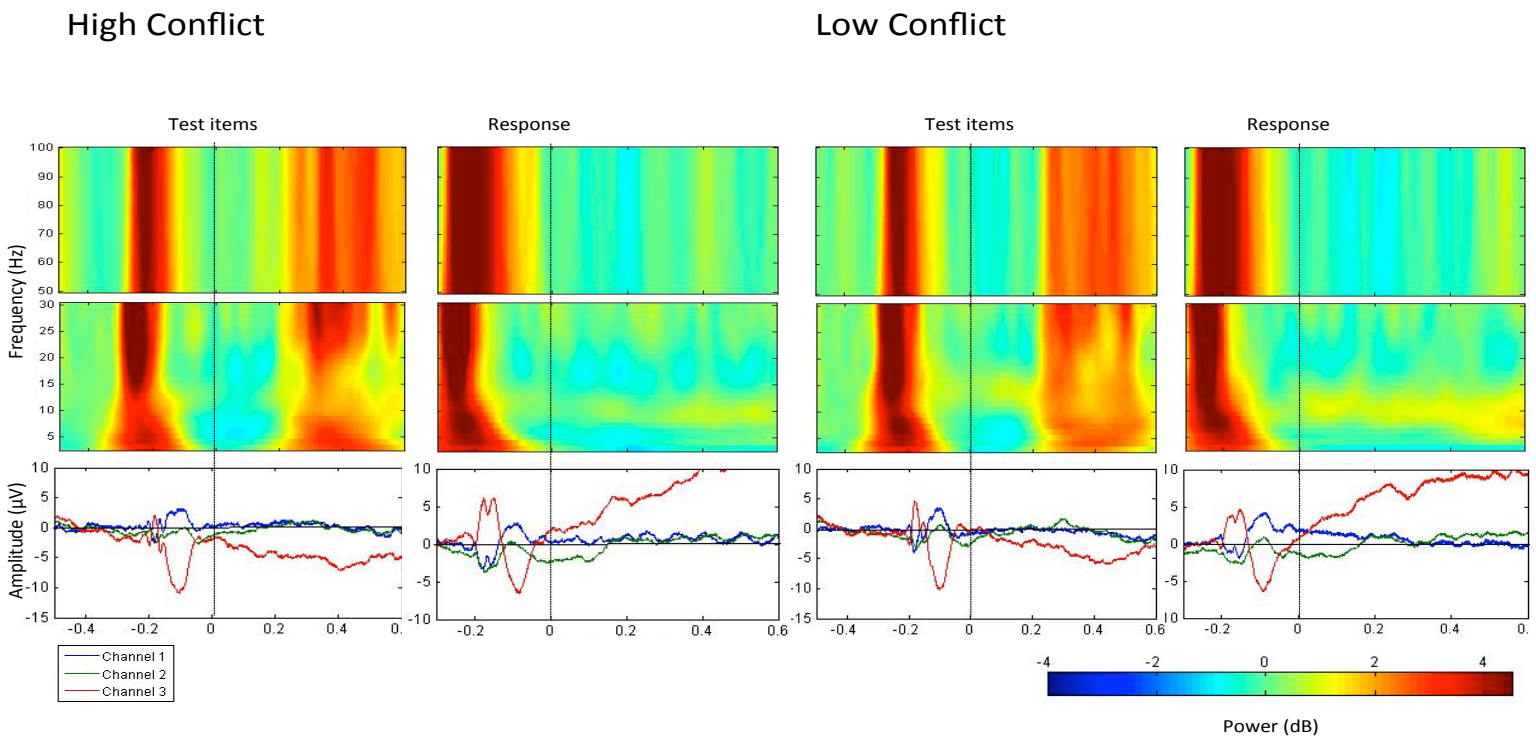


Figure A10. *Time-frequency decomposition and event-related potential activity in the right ACC for patient WS on high- and low-conflict trials in the conflict WCST analogue. Stimulus-locked and response-locked activation plotted on adjacent columns. As response time varied, data is plotted from -0.5s pre-test stimuli onset to 0.6s post-test stimuli onset for the stimulus-locked data and from -0.3s pre-response to 0.6s post-response for the response-locked data.*

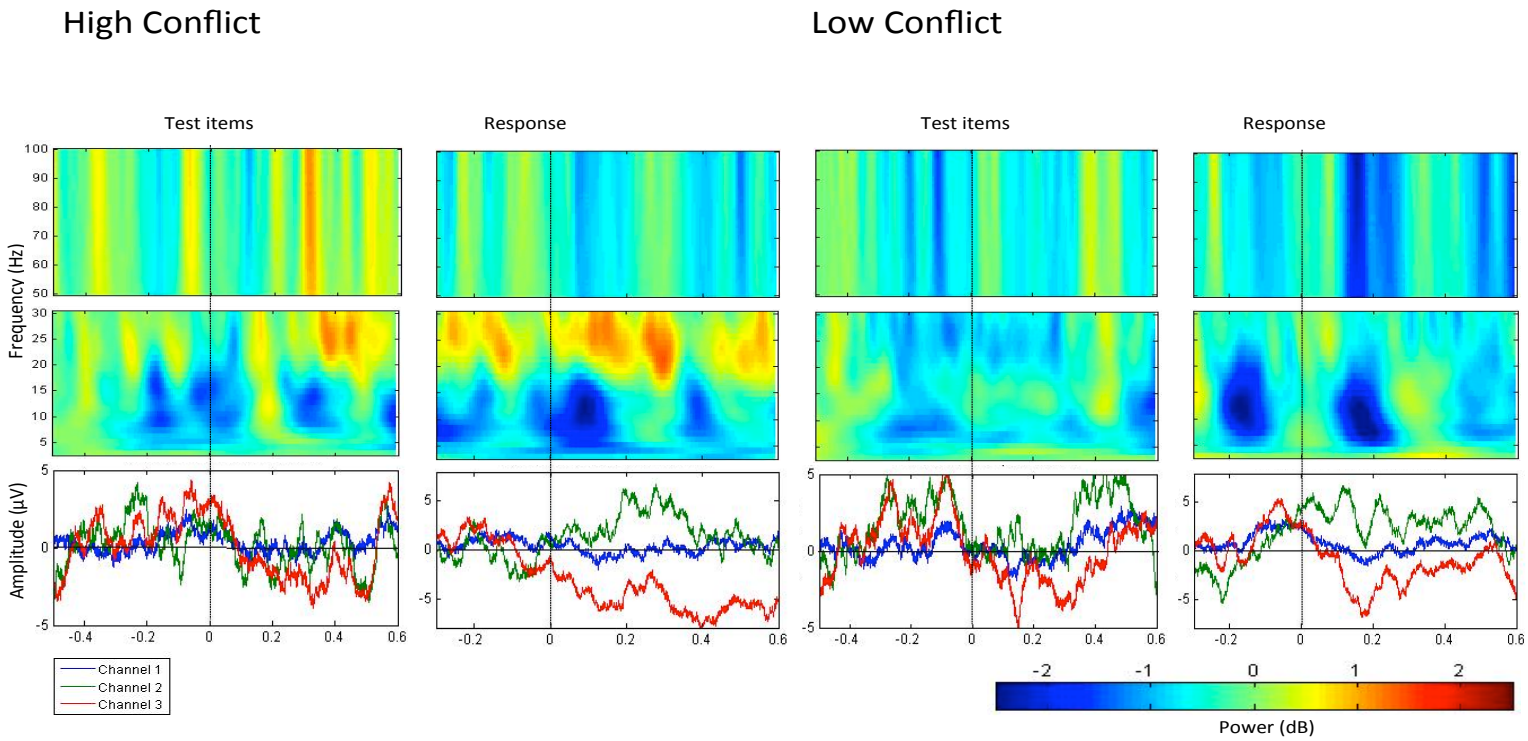


Figure A11. *Time-frequency decomposition and event-related potential activity in the left ACC for patient MG on high- and low-conflict trials in the conflict WCST analogue. Stimulus-locked and response-locked activation plotted on adjacent columns.*

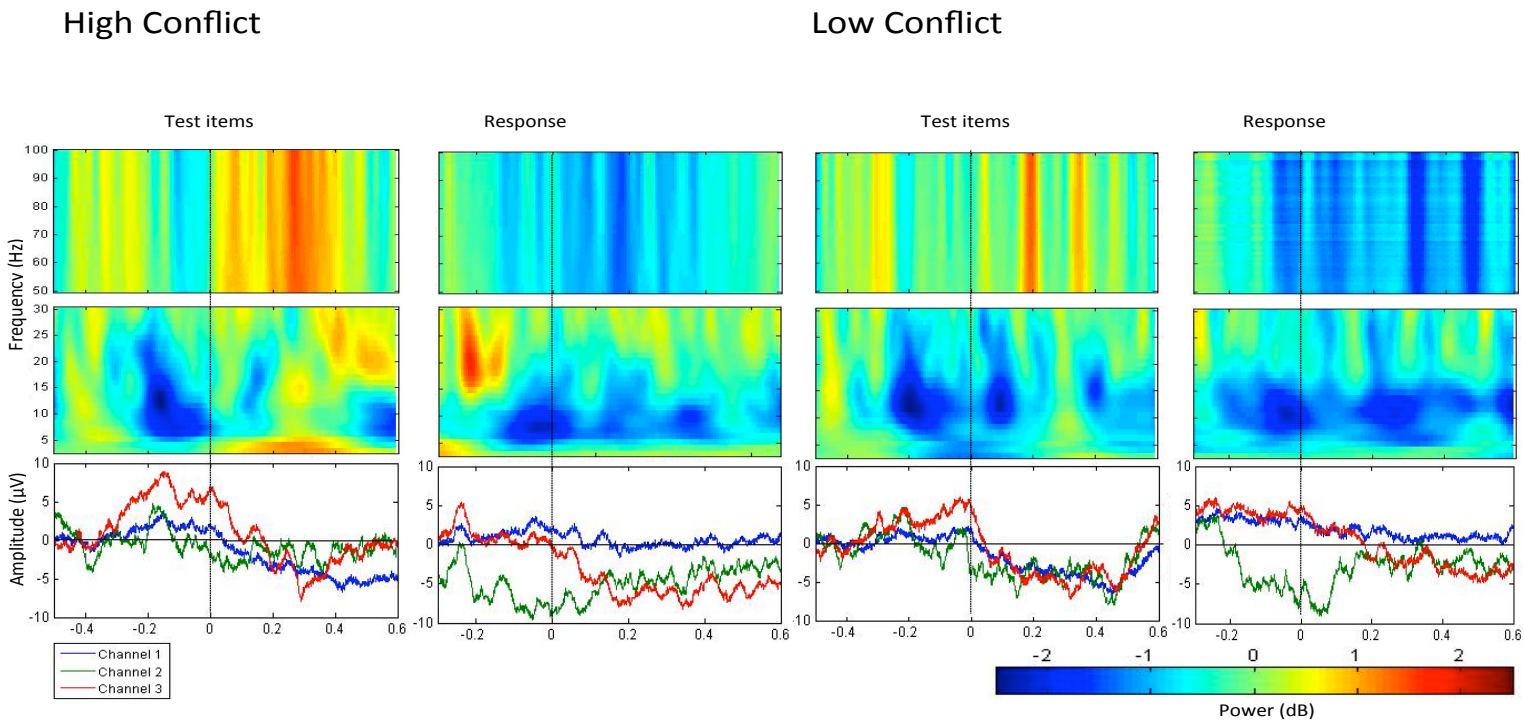


Figure A12. *Time-frequency decomposition and event-related potential activity in the left ACC for patient MG on high- and low-conflict trials in the 3-item Simon task. Stimulus-locked and response-locked activation plotted on adjacent columns.*

APPENDIX 4

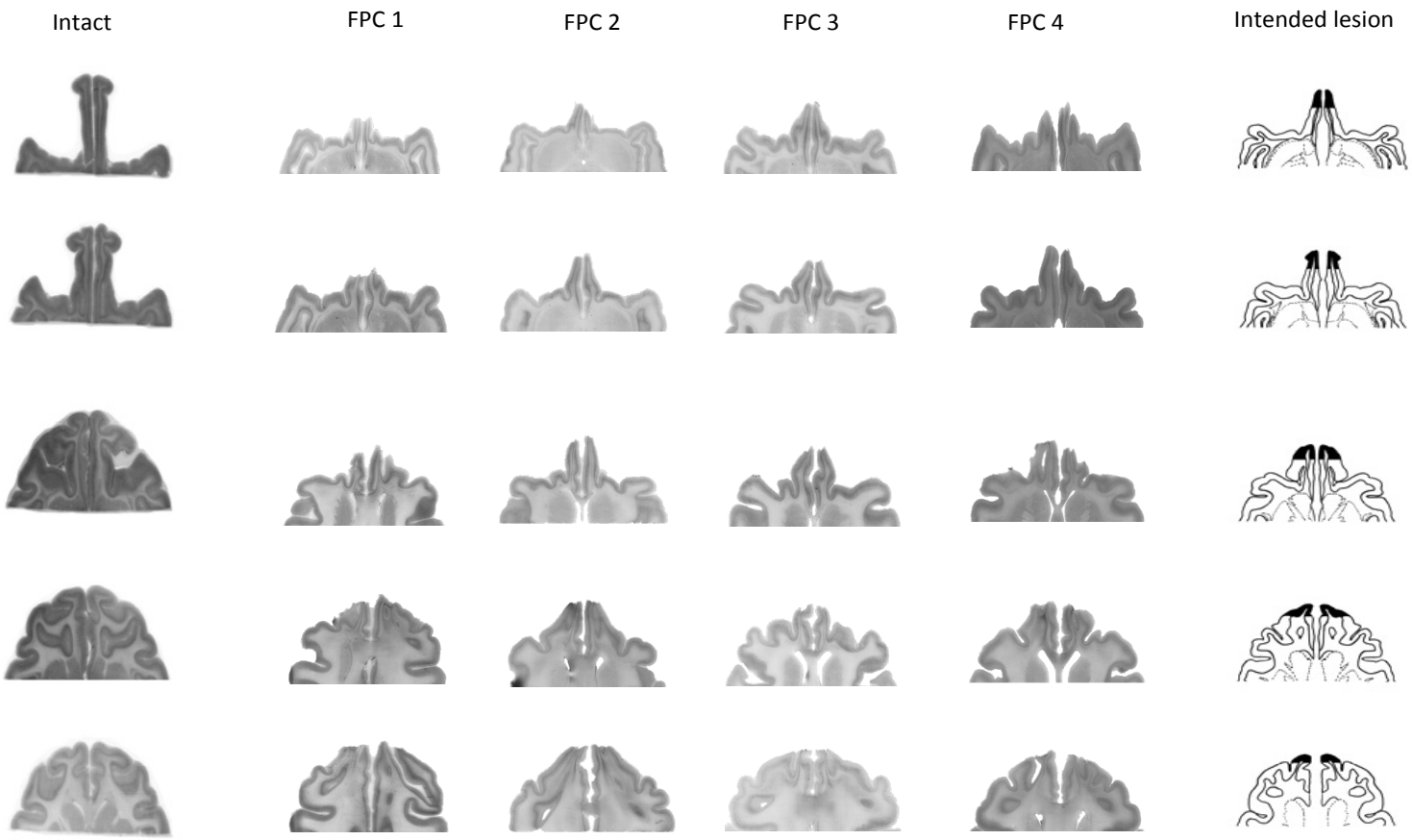


Figure A13. *Five dorso-ventral horizontal brain sections confirming the extent of FPC lesions. Leftmost columns shows the corresponding sections of an intact FPC. Rightmost column shows an illustration of the intended lesion in a standard brain. The middle columns (FPC 1, FPC 2, FPC 3 and FPC 4) show sections for the lesioned-animals included in the experiment.*