# Controlling Memory and Resolving Interference: Prefrontal Contributions to Flexible Behavior

by

David Badre

B.S., Biopsychology and Cognitive Science
University of Michigan, 2000

Submitted to the Department of Brain and Cognitive Sciences in Partial Fulfillment
of the Requirements for the Degree of

Doctor of Philosophy In Cognitive Neuroscience

at the

Massachusetts Institute of Technology

June 2005

Signature of Author: ______________________________
Department of Brain and Cognitive Sciences
7th, June 2005

Certified by: ______________________________
Anthony D. Wagner
Assistant Professor of Psychology and Neuroscience
Thesis Supervisor

Accepted by: ______________________________
Earl K. Miller
Picower Professor of Neuroscience
Chairman, Department Graduate Committee

# Controlling Memory and Resolving Interference: Prefrontal Contributions to Flexible Behavior

by

David Badre

Submitted to the Department of Brain and Cognitive Sciences
on April 15, 2005 in partial fulfillment of the
requirements for the Degree of Doctor of Philosophy in
Cognitive Neuroscience

## Abstract

How does the brain meet shifting task demands? The experiments and formal theoretical framework presented in this dissertation characterize the cognitive and neural processes by which flexible performance is enabled during task switching. Chapter 1 reviews major findings and controversies in the task switching literature, highlighting (1) evidence that behavioral switch costs may derive from proactive interference due to the facilitated retrieval of irrelevant competitors from long-term memory and (2) the consistent finding of activation in ventrolateral prefrontal cortex (VLPFC) during task switching. These observations motivate the hypothesis that left VLPFC may resolve proactive interference arising from long-term memory during a task switch. Chapters 2 and 3 describe three fMRI experiments conducted in experimental contexts independent of task switching that directly link left mid-VLPFC (Brodmann's Area 45; inferior frontal gyrus *pars triangularis*) to a post-retrieval selection process that resolves proactive interference from irrelevant representations retrieved from long-term memory. Chapter 4 introduces a computational model that derives its task switch cost from interference due to performance-dependent changes in its associative structure, and that resolves this interference through a control process that biases retrieved conceptual representations. Critically, a conflict signal, derived from retrieved conceptual representations in the model, is shown to be characteristic of the pattern of response in left mid-VLPFC during an fMRI experiment that manipulates preparation and interference in task switching. Furthermore, this pattern dissociates left mid-VLPFC from other regions active during a task switch. These data strongly support the hypothesis that task switch costs derive from proactive interference due to facilitated retrieval of irrelevant representations and left mid-VLPFC serves to overcome this proactive interference. Chapter 5 provides further details of the model, demonstrates its power to explain a number of common task switching phenomena, and explores its relationship with three other prominent formal models of task switching. The experiments and associated theory presented in this thesis provide evidence that instances of flexible behavior, like task switching, may be understood as acts of memory, and are enabled by prefrontal cortex mechanisms that control memory to overcome interference.

**Thesis Advisor:** Anthony D. Wagner
**Title:** Assistant Professor of Psychology and Neuroscience

# Acknowledgements

In taking the time out of a busy research schedule to pursue the stressful business of writing that short book that gains you a degree, there are a number of people whose support has been indispensable and to whom I owe a lasting debt of gratitude. I would like to thank my advisor, Dr. Anthony Wagner, for his guidance these last five years. I am truly a better scientist today for his mentorship, and I will always benefit from having a little version of Anthony in my head looking on critically as I write a paper or design an experiment. I also would like to extend my thanks to my committee for their guidance through this process and for their insightful comments that have strengthened and clarified this work. I am grateful to all those faculty, students, and lab mates who have contributed to my training over the years, particularly my dear friend and teacher Dr. John Jonides and my compatriot and verbal sparring partner Itamar Kahn. I also thank Itamar, Russ Poldrack, and Joe Sala for their helpful comments on early versions of this thesis. Finally, I would like to convey my deepest appreciation to my family. To my parents, whose love and confidence in me have always been a tremendous support. And of course, to my wife, Sejal, who endured the day-to-day insanity of dissertation writing with her usual grace and humor and who was always available for wise council or even a quick (or not-so-quick) proofread. It is to my wife, my parents, and also to my grandparents, whose scholarly lineage I inherited and to which I also aspire, that I dedicate this thesis. Many thanks to all.

# Contents

## CHAPTER 1

## CHAPTER 2

## CHAPTER 3

## CHAPTER 4

## CHAPTER 5

## CHAPTER 6

# Chapter 1

## Introduction

In his classic profile of Ramón y Cajal, Wilder Penfield (Penfield, 1926) recounts finding the scientist brooding in his laboratory library, quietly contemplating an important scientific problem; a portrait consistent with the predominating conception of the successful scientist closeted away over a microscope for long uninterrupted hours. Of course, most of us only wish for such a professional life. Rather, the course of modern life is often interrupted by demands that do not await our disposition but must be addressed immediately. Though these bids on our cognitive resources are irritating and, indeed, are often obstacles to achieving goals, the capability of our cognitive system to reconfigure itself to meet shifting task demands is evident and remarkable. A fundamental problem in the study of cognitive control is specification of the psychological and neural processes by which we achieve such flexible behavior.

Consider the following demonstration. We can readily recite the letters of the alphabet. And likewise, we can rapidly count from one to twenty-six. However, attempt to recite the alphabet and count to twenty-six simultaneously, alternating between the sequences after each letter or number as fast as you can (e.g., A – 1 – B – 2 – C – 3...). The difficulty we experience, relative to reciting either simple well-learned sequence by itself, is illustrative of a simple fact of human performance; it is harder to do two things and alternate between them than to do one thing repeatedly. The source of this difficulty and those processes recruited to overcome it are the core problem in the study of flexible performance operationalized in task switching.

The difficult we experience in switching tasks may be partially attributable to the demand to activate a new set of task-relevant representations each time we engage in a new task. This task-specific configuration of the system is often referred to collectively as the *task set* (Anderson et al., 2004; Logan & Gordon, 2001; Mayr & Keele, 2000; Meyer & Kieras, 1997). For example, a production system model (e.g., Anderson et al., 2004; Meyer & Kieras, 1997) might require updating goals (e.g., recite alphabet), specifying relevant cues ('a' was the last letter), retrieving useful knowledge ('b' follows 'a'), and selecting response mappings (e.g., b -> "[biː]"). Furthermore, many theorists maintain that the system cannot reconfigure itself automatically, via simple bottom-up inputs, but rather requires top-down, cognitive control processes (e.g., Monsell, 2003).

Cognitive control enables flexible cognition by biasing processing in favor of task-relevant representations over prepotent competitors (Cohen, Dunbar, & McClelland, 1990; Desimone & Duncan, 1995; Miller & Cohen, 2001; Norman & Shallice, 1986). It follows, then, that the loss of cognitive control will be marked by an overdependence on external stimuli or strong associations to guide responses and so a loss of adaptive, goal-directed behavior. Such a pattern of behavior is evident following damage to PFC (e.g., utilization behavior), and has resulted in a predominant focus on PFC as central to the neural system supporting cognitive control. Hence, the extent and manner by which PFC is required for flexible alternation between tasks is a central concern in the study of task switching.

In the following review, I first consider the major behavioral phenomena in task switching and discuss the three major theoretical perspectives that find their support in these results and that frame the debate over task switching. In addition to grounding further discussion in the central phenomena, debates, and ambiguities of the task switching literature, a central theme to emerge from this review will be the role played by memory during task switching. In particular, performance-induced changes in long-term memory may produce interference that gives rise to task switch costs and incurs a control demand. With this background as context, I will then review the recent neuroimaging and neuropsychological literature on task switching, considering what insights these studies have offered with respect to the major controversies to emerge from the behavioral literature. Furthermore, I will note that the overwhelming results from these experiments implicate left ventrolateral prefrontal cortex (left VLFPC) as playing a critical, if as yet unspecified, role in task switching. Thus, two themes emerge from this introductory review: (1) Interference arising from performance-dependent changes in long-term memory may produce control demands during task switching. (2) VLPFC, a region strongly associated with the control of memory, is commonly among those prefrontal regions sensitive to task switching. These observations motivate the investigative approach taken by this thesis, which seeks to specify the mechanisms by which VLPFC controls memory and further to determine whether such mechanisms have a role to play in task switching.

### *The Task Switch Cost – Task Set Reconfiguration versus Task Set Inertia*

Task switching can be studied by comparing episodes in which subjects *switch* between two simple tasks to those in which they *repeat* the same task. In such comparisons, task switching is associated with a behavioral cost, typically a slowing in response time (RT) and/or a decline in accuracy (Jersild, 1927; Logan, 2003; Monsell, 2003). Switch costs can be obtained within a knowledge domain (e.g., addition versus subtraction; Jersild, 1927; Rubinstein, Meyer, & Evans, 2001; Spector & Biederman, 1976), across knowledge domains (e.g., vowel/consonant letters versus odd/even numbers; Rogers & Monsell, 1995; Sohn & Anderson, 2001), within perceptual domains (e.g., vertical versus horizontal spatial location; Meiran, Chorev, & Sapir, 2000), and across perceptual domains (e.g., spatial frequency versus face recognition decisions; Wylie, Javitt, & Foxe, 2004). Furthermore, switch costs seem fairly impervious to experience, persisting after considerable practice with task switching episodes (Rogers & Monsell, 1995; Yeung & Monsell, 2003). Indeed, it appears that switch costs are eliminated only to the extent that a target stimulus cannot be used to perform a competing task or there is no change in the relevant stimulus dimension, and the tasks are unambiguously cued (Jersild, 1927; Spector & Biederman, 1976).

Despite the unanimity with which switch costs have been obtained in the literature, their theoretical interpretation has been far more controversial. Perhaps following the logic used to understand the psychological refractory period phenomenon (Broadbent & Gregory, 1967; Meyer & Kieras, 1997; Pashler, 1994), early theoretical interpretations of the task switch cost considered it directly reflective of an intentional reconfiguration stage in a serial information processing chain that would shunt the system from one task state to the next on switch trials. As noted by Allport (2000), many of the assumptions of this model are consistent with those of a finite state machine. During performance of a given task, the cognitive system enters one task state. It will remain in this state until another task state is required, at which time an intentional control process shunts the system into the new task state, where it will remain until the intention to perform another task shunts it to the next state, and so on. From this perspective, the time required for this mental "gear shift" is directly evident in the switch RT cost.

Allport, Styles, & Hsieh (1994; Allport & Wylie, 2000; Wylie & Allport, 2000) provided an initial empirical challenge to the classic mental gear-shifting model and proposed an important alternative interpretation of switch costs. Using a variant of the Stroop task, subjects switched between a word reading task, wherein the subject read a color word printed in a different color ink, and a color naming task, wherein the subject named the ink color a color word was printed in. Consistent with classic work on this effect (MacLeod, 1991; Stroop, 1935), performance of the non-dominant color naming task resulted in slowing when the word named a color incongruent with the ink color, but performance of the dominant word reading task was unaffected by the congruency of the ink color in which the word was printed. Critically, switching from the dominant word reading task to the non-dominant color naming task produced little switch cost, whereas, switching from the non-dominant color naming task to the dominant word reading task produced a substantial switch cost.

That the switch costs were *asymmetrical*, and further that the cost was *greater* when switching to the *better*-learned task, was striking and counterintuitive from the perspective of a unitary switching mechanism that shunts the system from one task to another. It was difficult to imagine a gear shift mechanism which had a harder time shunting the system toward a better-learned task. Rather these results suggested that task switch costs might arise from *task set inertia* (TSI) effects that carry-over transiently from a previously performed task.

TSI conceptualizes the operation of cognitive control during the Stroop task in a way similar to prominent models of this task (e.g., Cohen et al., 1990; Cohen & Servan-Schreiber, 1992; Gilbert & Shallice, 2002) and so assumes that successful performance of the non-dominant color naming task requires biasing the relevant color naming pathway and inhibition of competing representations from the dominant word reading pathway. Consequently, when switching to the dominant word reading task, these representations are more difficult to access. Furthermore, the compensatory biasing of the non-dominant color naming task facilitates or primes its task set, and this competitor set carries over and further interferes with access to the word reading task set. Hence, TSI argues that performance of a given task results in effects that carry-over and interfere with performance of a subsequent competing task, and it is this interference that gives rise to task switch costs. Indeed, in its strongest form (Allport et al., 1994), TSI suggests that behavioral task switch costs are reflective of little control at all and emerge entirely from carry-over interference.

Closely after the proposal of TSI, the traditional "mental gear shift" model of task switching faced a further theoretical challenge from the *residual switch cost* phenomenon (Rogers & Monsell, 1995). In their classic experiment, Rogers and Monsell (1995) instructed subjects to alternate between simple vowel/consonant letter decisions and odd/even number decisions. Their procedure, called the alternating runs procedure, introduced two novel methodological features. First, stimuli were letter-number pairs, like the pair a1, and were presented clockwise around a 2 × 2 grid. Critically, the relevant task was cued by a stimulus' position in the top or bottom two cells of the grid (e.g., Top = Letter task; Bottom = Number task). Prominent previous studies had relied primarily on list completion times (e.g., Jersild, 1927; Spector & Biederman, 1976), in which the time to complete lists or blocks of alternating task stimuli was compared to the time to complete lists or blocks of single task stimuli. The alternating runs procedure, however, permitted switch and repeat trials to be inter-mixed and compared at the trial level. Furthermore, because of the consistent clockwise pattern to presentation, the subject could always anticipate the task for the upcoming trial before the stimulus was presented. Hence, varying the interval between a subject's response and the subsequent stimulus (response-to-stimulus interval [RSI]) provided a manipulation of preparation time. To the extent that task switching reflects intentional mental transmission time, as in the classical model, an increase in preparation time should result in a corresponding reduction in the switch cost. Indeed, the switch cost should eventually be eliminated whenever preparation time is greater than or equal to switch time. However, contrary to these predictions, though switch costs did decline with preparation time, they were not eliminated. Even at long preparation intervals, a residual switch cost remained. Such residual switch costs have been reported repeatedly in the literature and will persist even at very long preparation intervals (Meiran et al., 2000; Rogers & Monsell, 1995; Sohn, Ursu, Anderson, Stenger, & Carter, 2000).

Rogers and Monsell (1995) interpreted the residual switch cost as reflective of a set of processes, termed *exogenous reconfiguration processes*, which contribute to reconfiguring a new task and require presentation of a task-relevant stimulus to be completed. These processes were considered distinct from a separable set of intentional control process, *endogenous reconfiguration processes*, which partially reconfigure the system prior to the presentation of the task stimulus, resulting in the initial preparation-related decline in switch costs.

Furthermore, Rogers and Monsell (1995) noted that for a run of any given task, only the first trial in the run showed an elevated RT. The subsequent trials in a run (repeat trials) were all equally fast relative to the first switch trial. Hence, in contrast to what one might anticipate if the switch cost were due to passive, transient carry-over, there was no additional decline in RT after the first trial. This *first trial effect* was considered consistent with the task set reconfiguration (TSR) hypothesis and inconsistent with TSI. It should be noted, however, that at least one connectionist treatment of task switching has demonstrated that a first trial effect is not necessarily incommensurate with a TSI model (Gilbert & Shallice, 2002), in that interference can be effectively resolved on the first trial of a run with little subsequent carry-over. Beyond the first trial effect, preparation-related declines were not apparent when the RSI was randomized or the subjects were unable to anticipate when the next target stimulus would appear (e.g. a warning cue 500ms before target onset). Again, it is difficult to understand how a passive effect of transient carry-over would depend on a subject's ability to anticipate how much time is available before presentation of the next stimulus. Critically, however, the TSI hypothesis with its focus on carry-over and interference and the TSR hypothesis, which highlighted reconfiguration control processes, reframed the debate over task switching and served as a context for most future studies.

### *Task Set Priming and Long-Term Associative Memory in Task Switching*

In their initial formulations, the TSR and TSI hypotheses differed fundamentally in their interpretation of RSI-related declines in task switching RT costs. Specifically, TSI attributed the decay to the declining activation of the previous task set and so the diminishing impact of *transient carry-over.* By contrast, TSR attributed the decay in switch costs to the opportunity given an endogenous control process to prepare for an upcoming task. Critically, however, the implicit task cueing obligated by the alternating runs procedure, and similarly the procedure used by Allport et al. (1994), did not rest direct control over precisely when, prior to an upcoming trial, the subject initiated a task switch. Consequently, preparation time and decay time were fully confounded in the RSI manipulation. Indeed, some have noted that as the time at which a subject successfully engages a switch is not directly observable, residual switch costs may simply emerge from the likelihood that a subject fails to engage a reconfiguration mechanism during the preparation interval, potentially reconciling the residual cost with the classical gear shifting model (Dejong, 2000). Perhaps consistent with this perspective, motivational manipulations intended to vary subjects' engagement during preparation have been shown to produce shifts in this estimated probability of engagement (Nieuwenhuis & Monsell, 2002).

To address the confound of preparation time with decay time more directly, the explicit cueing procedure (Mayr & Kliegl, 2000; Meiran, 2000; Meiran et al., 2000; Sohn & Anderson, 2001) permitted separate manipulation of the intervals for decay and preparation (Figure 1). In the explicit cueing procedure, the subject receives a random sequence of bivalent task stimuli and is cued which task to perform prior to the presentation of the target stimulus. The cue is an explicit instruction about which task to perform next, and until the cue is presented, the subject does not know which task to perform with the upcoming stimulus and so whether to prepare for a task switch. Thus, independently varying the response-to-cue interval (RCI) and the cue to stimulus interval (CSI) permits separate manipulation of the interval for decay and the interval for preparation. Interestingly, switch costs will decay over a fixed CSI as RCI increased, reflective of transient carry-over, and will also decline sharply as CSI increases over a fixed RCI. Of course, increasing the CSI also increases the RSI (RCI + CSI) and so the opportunity for decay. However, CSI based declines are evident even at a very long constant RCI, a duration at which RCI-based declines are minimal and near asymptote (Meiran et al., 2000). Moreover, differences in the deceleration of the cost functions arising from the RCI and CSI manipulations also suggest that they arise from different sources. Hence, though preparation related declines are evident using the explicit cueing procedure, potentially consistent with a preparatory control process such as that

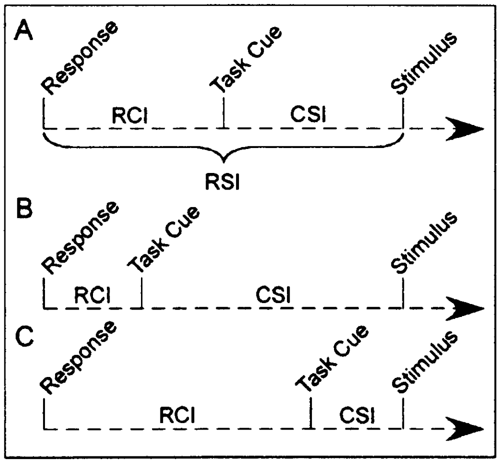

**Figure 1.** (A) Separation of the RSI into RCI and CSI based on the placement of the task cue in the explicit cueing task. A larger CSI (B) results in greater opportunity for preparation than a shorter CSI (C) over a constant RSI.

hypothesized by TSR, there is also evidence of transient carry-over, potentially consistent with TSI.

Recent results, however, have raised doubt over the conclusion that CSI-based declines in switch costs must be due to preparation related reconfiguration processes. Logan and Bundesen (2003; 2004; Arrington & Logan, 2004) have argued that, under certain conditions, cue encoding may entirely account for preparation effects during the explicit cueing procedure. Using a combination of empirical and theoretical arguments, they claim that the co-encoding of a task cue and target stimulus is sufficient to specify the appropriate response rule during a task switch and so no control process is differentially necessary during a switch in this procedure. From this perspective, switch costs during the explicit cueing procedure, after task or goal specification (see below), are simply due to repetition benefits in access to memory due to the combined task-cue–response rule.

Along similar lines, a number of recent results have suggested that, though some interference in task switching may be due to short-term transient carry-over of the type proposed by TSI, a much more substantial portion of task set interference may arise from interactions within long-term memory, termed task set priming (TSP) and this interference is evident in long-term carry-over effects (Allport & Wylie, 2000; Koch, Prinz, & Allport, 2005; Mayr & Kliegl, 2000; Sohn & Anderson, 2001, 2003; Waszak, Hommel, & Allport, 2003). Encountering a target that was previously encountered in the context of a given task, even after many intervening trials, results in slower RT to perform a different task upon second presentation (Allport & Wylie, 2000). This finding motivated the hypothesis that a target can act as a cue for retrieval of a task set with which it was previously associated. More specifically, TSP proposes that performance of a given task will result in strengthening associations among cues and stimuli encountered during performance of that task (and possibly weakening of associations among irrelevant stimuli). Upon subsequent encounter, these stimuli can act as cues that automatically retrieve irrelevant and competitive information. Hence, TSP hypothesizes that greater task set interference and greater switch costs should be evident to the extent that a target has been previously encountered in the context of another task. Consistent with this hypothesis, during a task switching regime, arranging stimuli such that some appear in both tasks and others appear in only one of the tasks results in greater switch costs for the former stimulus set (Waszak et al., 2003).

Critically, such stimulus-cued long-term carry-over is not necessarily related to exclusively stimulus-level processing, such as task cue encoding or stimulus recognition (Sohn & Anderson, 2003). If subjects are only required to identify the task and stimulus identity (e.g., "the instruction cues the number task" and "the number is an '8'") but not carry through with an actual categorization task (e.g., the odd or even number judgment), this partial overlap produces a switch effect only at very brief intervals when there is stimulus-level carry-over, consistent with some transient carry-over for these operations (Sohn & Anderson, 2003). But at longer intervals any cost related to stimulus-level priming during partial overlap is extinguished. By contrast, during full overlap, when subjects are also required to complete the categorization, the interaction of a stimulus prime with switch costs is substantial even at a long interval (Sohn & Anderson, 2003). Hence, long-term carry-over in task switching seems to impact the performance of the task during a switch rather than other aspects of cue specification or stimulus identification. These effects strongly suggest that task performance produces changes in long-term memory that impact access to relevant stored representations and contribute substantially to the behavioral effects of task switching.

In addition to long-term carry-over effects, a number of additional behavioral phenomena have been cited as evidence that changes in long-term associative memory may be central to the behavioral effects in task switching (Altmann, 2004; Arrington & Logan, 2004; Mayr & Kliegl, 2000). For instance, restart costs are evident in an elevated RT on the first trial of a run of trials following a brief delay but without a shift of task (Allport & Wylie, 2000; Altmann & Gray, 2002; Gopher, Armony, & Greenshpan, 2000). Furthermore, these restart costs can be modulated by exposure to a competing task set (Allport et al., 1994). To the extent that restart costs arise from a similar source as switch costs (Wylie & Allport, 2000), such effects suggest that task performance is impacted to the extent that stimuli cue competing task representations. Also potentially consistent with the associative memory perspective, a number of results suggest that switch costs may be enhanced to the extent that retrieval of response-relevant information or response rules is more difficult (Jersild, 1927; Mayr & Kliegl, 2000; Rubinstein et al., 2001). Hence, a large portion of task switch costs may be attributable to the demand to access long-term memory. During a task switch, the difficulty in achieving this access may be due to interference in memory, such as proactive interference (Allport & Wylie, 2000; Logan, 2003).

It is critical to note that long-term carry-over effects do not preclude the operation of a cognitive control process engaged during a switch of task. For example, to the extent that interference arises, due to short- or long-term carry-over, cognitive control mechanisms may be important to overcome this interference. Potentially consistent with this perspective, if subjects alternate among three tasks (Tasks A, B, and C), a switch cost is incurred upon each task switch. However, this cost is larger when switching back to a task that had recently been performed (e.g., the final A in the sequence A – B – A) than to a third task (e.g., C in the sequence A – B – C; Dreher & Berman, 2002; Mayr & Keele, 2000). This task-level backward inhibition effect has been interpreted as indicative of a control process operating during a task switch (Mayr & Keele, 2000). To the extent that such a control process is engaged to overcome interference during an initial task switch, this may make the competing representations relatively less accessible and so returning to the original task may also be more difficult. Hence, though interference in memory may be a source of task switch costs, cognitive control mechanisms may be critical in resolving this interference and permitting the relevant task to proceed.

### *Sources of Proactive Interference during Task Switching*

Central to the TSP hypothesis is the proposition that elements of the stimulus display may act as cues that give rise to interference. Two commonly cited task switching proactive interference effects, *cross-talk* and *response repetition*, may be understood from such a perspective, though no data yet directly confirms such an account. However, considering these proactive interference effects in some depth is of potential importance for evaluating the theoretical perspectives on task switching.

A common feature of task switching paradigms is a bivalent stimulus that does not, by itself, cue performance of either task but contains dimensions relevant to both. For example, in the alternating runs procedure, the stimuli were letter-number pairs (‘a1’) on which either task (odd/even or vowel/consonant) could be performed. Critically, the presentation of a stimulus relevant to a competing task should result in interference with task switching to the extent that it cues some aspect of the primed and irrelevant task set. Consistent with this hypothesis, trials in which an irrelevant flanker cues a competing task (cross-talk condition) produce enhanced switch costs relative to trials in which the flanker is a neutral symbol (no cross-talk condition), not relevant to the other task (Rogers & Monsell, 1995).

Interestingly, the increased cost during cross-talk conditions is not necessarily due to competition at the response level, as with the classic Erickson flanker task (Eriksen & Eriksen, 1974; Yeung, Cohen, & Botvinick, 2004) in that there is no consistent effect of the congruency of the response cued by the flanking stimulus on the switch cost (Meiran, 2000). In other words, switch costs are enhanced during cross-talk relative no cross-talk conditions regardless of whether the response cued by the flanker is congruent or incongruent with the response cued by the task-relevant stimulus (Rogers & Monsell, 1995). Furthermore, in contrast to the standard task switch cost, the effects of cross-talk do not interact with preparation time, as expressed in RSI (Meiran et al., 2000; Rogers & Monsell, 1995) and are present even at a long RSI. Though no account of cross-talk effects has been broadly accepted, these results are consistent with the view that a stimulus associated with a competing task will cue retrieval of this task and may enhance interference during task switching.

As a flanking stimulus may provide a cue for retrieval of a competing task set, generation of a response itself may also cue a competing task. Often responses from different tasks during a task switching regime are mapped to an overlapping set of responses (e.g., the same two fingers for two key stroke responses). Whereas typically, repetition of a given response will result in facilitation, the opposite occurs during a task switch. Task switch costs are greater during response repetition than during a switch of response (Meiran et al., 2000; Rogers & Monsell, 1995). Furthermore, a larger switch cost is not only due to facilitation from a repeated response on a task repeat trial but also to a slower response repetition on switch trials (Meiran et al., 2000). This latter effect indicates the presence of interference, even though the physical response repeats. Similar to the cross-talk effects discussed above, response repetition does not interact with preparation time (CSI) and larger response repetition switch costs are evident even at long preparation and decay intervals (Meiran et al., 2000).

As already noted, both cross-talk and response repetition effects are important as they are potential manipulations of stimuli encountered during performance of a given task that cue retrieval of a competing task set. However, advocates of TSR note the lack of decay in cross-talk and response repetition effects and suggest that these may contribute to the residual cost thought to be reflective of exogenous reconfiguration processes (Rogers & Monsell, 1995). Additional data is required to understand the relationship of these important proactive interference effects to the major perspectives on task switching.

### *Goal Setting and Task Specification*

An additional critical aspect of task switching, and one potentially independent of demands to access memory over interference, is determining what task to perform next, often termed goal setting or task specification. Indeed, goal setting was a pervasive confound in classical list completion studies in that a mixed list always required subjects to monitor where they were in sequence, whereas this was unnecessary during a pure list. Alternating runs was developed, in part, to alleviate this confound and require subjects to maintain both task goals throughout all trials. However, the requirement to determine what task to perform next was still just as important, but now it was simply externally cued by a position on a grid and was required equally between switch and repeat trials.

Underscoring these demand differences, a number of studies have noted that switching costs that emerge from the difference between switch and repeat trials during a given run, as in alternating runs, are distinguishable from so-called *mixing costs* (Meiran et al., 2000) that derive from the difference in average trial completion times from pure and mixed lists (Mayr, 2001; Meiran et al., 2000; Meiran, Gotler, & Perlman, 2001). Repeat trial RTs from a mixed block are longer than repeat trial RTs during a pure block. This advantage for pure blocks is susceptible to a form of carry-over, by which the facilitation for pure block repeat trials increases over the course of a pure block run. Furthermore, this change may be due to the operation of an active process rather than a diminishing effect of carry-over. In older adults, for whom cognitive control is compromised and who otherwise show enhanced switch costs, the slope of this pure repeat trial facilitation change is shallower than controls (Mayr, 2001). Hence, mixing costs may be indicative of an additional demand present during even Repeat trials of alternating blocks and not during pure blocks; namely the requirement to determine what task to do next.

Determining what task to perform next may also be a difference between alternating runs and other procedures, such as the explicit cueing procedure. During alternating runs, the next task to perform must be determined from the position of a target on a display, and so may be relatively more difficult than during explicit cueing. Furthermore, processes that determine the next task to perform, when this information is not made explicit, may be distinct from those processes that access task relevant knowledge, and so may not be sensitive to task set interference or ease of access to long-term memory. As already noted, switch costs are enhanced to the extent that access to memory is made more difficult (Jersild, 1927; Rubinstein et al., 2001; Spector & Biederman, 1976). Furthermore, the difficulty with which a subject can determine what task to perform next, termed goal setting, will also increase switch costs (Rubinstein et al., 2001; Spector & Biederman, 1976). Interestingly, making both access to memory and goal setting difficult results in an additive enhancement to switch costs, suggesting that these changes may arise from independent processes (Sternberg, 1969).

Additional evidence for separable goal setting or task specification processes comes from studies using explicit cueing procedures. During explicit cueing, a switch of task always necessitates a switch of task cue as well, confounding task cue encoding with task switching. By arranging more than one task cue for a single task, a task cue switch can occur independent of a task switch. Using this procedure, a considerable portion of task switch costs during the explicit cueing procedure have been shown to be attributable to cue switching rather than task switching (Logan & Bundesen, 2003, 2004; Mayr & Kliegl, 2003). Hence, encoding a task cue may depend upon or inform a process of goal specification.

In their analysis of the explicit task cueing procedure, Logan and Bundesen (2003; 2004) cite the task cue switching effect as the primary component in the switch cost measured during explicit cueing, and suggest that repetition facilitated during a preparation interval accounts entirely for the switch cost and preparation effect in the explicit cueing procedure. Consequently, they argue that there is little need for recourse to a reconfiguration process or even cognitive control to account for the behavioral effects associated with this procedure. This is an important challenge to the TSR hypothesis, though it is initially unclear how such a model accounts for the findings showing enhanced switch costs due to enhanced difficulty with rule retrieval as additively independent from that associated with task specification (Rubinstein et al., 2001). To the extent that effective task cue encoding forms one part of the cue for retrieval of a response rule, in their model, one might expect an interaction between goal specification and rule activation manipulations. Furthermore, there is some evidence showing that the effectiveness of preparation impacts behavioral measures of task switching (Nieuwenhuis & Monsell, 2002; Yeung & Monsell, 2003), even during explicit cueing. Finally, it is unclear how the Logan and Bundesen (2003) model makes contact with the neuroimaging and neuropsychological data implicating prefrontal cortex as critical for task switching and mostly using the explicit cueing procedure. However, it does seem clear that a consensus has not yet emerged from the behavioral literature regarding the role of interference in task switching, and whether multiple cognitive control mechanisms, or even no cognitive control mechanisms, are required to flexibly alternate between tasks.

### *Summary of the Cognitive Perspectives on Task Switching*

The source of the RT cost incurred during a switch of task has generated considerable controversy. TSR suggests that the switch cost may be attributed to a time-consuming control process that reconfigures the system for an upcoming task. TSI suggests that a task set, activated on the previous trial, carries over transiently to interfere with a subsequent task. Finally, TSP suggests that the task switch cost may be due to interference arising from retrieval of associated competitive information in long-term memory. To some extent, memory plays a central role in all of these theories, with the task switch cost alternatively reflecting the time required to retrieve relevant information or the additional time caused by proactive interference from activated long-term memory representations. Hence, investigation of task switching as an act of memory and in the context of the processes that access and operate on retrieved representations may shed considerable light on this controversy. In this final section, I consider the neurobiological literature on task switching, and focus on the extent to which the neural substrates in these studies, and particularly ventrolateral prefrontal cortex, may provide insights into the contribution of mnemonic processes in task switching.

### *The Neurobiology of Task Switching*

Neurobiological investigation in humans has closely linked task switching with neural computations in PFC. Patients with broad lesions in lateral PFC demonstrate deficits in task switching performance (Aron, Monsell, Sahakian, & Robbins, 2004; Mecklinger, von Cramon, Springer, & Matthes-von Cramon, 1999; Rogers et al., 1998). And, evidence from patients with striatal diseases and studies using pharmacological interventions implicate the dopaminergic system as playing a part in shifting/maintaining task set (Cools, Barker, Sahakian, & Robbins, 2003). In addition, lesion studies in the non-human primate (Dias, Robbins, & Roberts, 1997) and single unit recording studies (Miller & Cohen, 2001; Wallis & Miller, 2003) have demonstrated that PFC is critical for selection for action when it is necessary to encode and maintain context information to bias a response, particularly when that contextual information changes, instances potentially analogous to a task switch.

Neuroimaging methodologies have contributed additional high-spatial and temporal resolution data to the study of task control, and the basic comparison of switch with repeat trials has revealed a fairly replicable network of switch-related regions (Figure 2). An overwhelmingly consistent finding is of ventro-lateral prefrontal cortex (VLPFC), including the posterior portion of the inferior frontal gyrus (BA44, pars opercularis) and extending several centimeters anterior to BA 45 (pars triangularis), often bordering superiorly in the inferior frontal sulcus (Brass & von Cramon, 2002, 2004a, 2004b; DiGirolamo et al., 2001; Dove, Pollmann, Schubert, Wiggins, & von Cramon, 2000; Dreher & Grafman, 2003; Dreher & Berman, 2002; Dreher, Koechlin, Ali, & Grafman, 2002; Konishi et al., 2002; Luks, Simpson, Feiwell, & Miller, 2002; Meyer et al., 1998; Meyer et al., 1997; Reynolds, Donaldson, Wagner, & Braver, 2004; Ruge et al.,

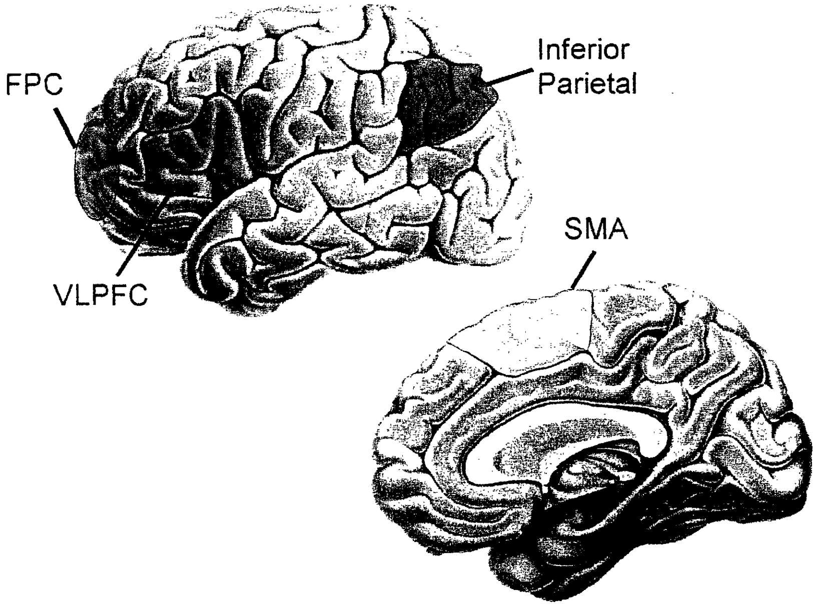

**Figure 2.** Schematic of common sites of activation reported during studies of task switching. VLPFC, particularly on the left, in addition to SMA and inferior (and sometimes superior) parietal cortices are commonly observed. In addition, under specific experimental conditions, FPC is also activated during task switching.

2005; Shulman, d'Avossa, Tansy, & Corbetta, 2002; Smith et al., 2001; Sohn et al., 2000; Swainson et al., 2003; Wylie et al., 2004). Though, it is notable that studies are not entirely consistent with respect to which hemisphere they obtain switch-related activation in VLPFC (Aron et al., 2004; Sohn et al., 2000; Wylie et al., 2004), activation is also typically left lateralized or bilateral. Beyond VLPFC, switch-related activation has also been observed in lateral frontal polar cortex (FPC), DLPFC, and SMA/pre-SMA (Braver, Reynolds, & Donaldson, 2003; DiGirolamo et al., 2001; Dreher et al., 2002; Gurd et al., 2002; Kimberg, Aguirre, & D'Esposito, 2000; Konishi, Jimura, Asari, & Miyashita, 2003; Luks et al., 2002; Meyer et al., 1998; Meyer et al., 1997; Nagahama et al., 2001; Rushworth, Passingham, & Nobre, 2002; Sakai & Passingham, 2003). Correlations between behavioral switch costs and neural markers of switching have been obtained in multiple regions such as VLPFC, pre-SMA (Sohn et al., 2000), and frontal pole (Dreher et al., 2002).

There is, however, no evidence for a "switch-exclusive" region of PFC (Brass & von Cramon, 2002; Dove et al., 2000; Dreher et al., 2002; Sohn et al., 2000). Indeed, all of the regions cited above are also active during repeat trials compared to a baseline condition. Hence, whatever control processes may be necessary during switch trials must also be involved, albeit to a lesser extent, during repeat trials.

Outside of PFC, activation is commonly observed in superior and inferior parietal cortex as well as the cerebellum (Dove et al., 2000; Dreher & Grafman, 2002; Dreher et al., 2002; Gurd et al., 2002; Kimberg et al., 2000; Meyer et al., 1998; Sohn et al., 2000). Consistent with this fronto-parietal network, ERP studies of task switching have found the global timecourse for task switching to be characterized by two separable temporal components: an early frontal component (~300-500ms) followed by a subsequent parietal component (~500-1000ms; Karayanidis, Coltheart, Michie, & Murphy, 2003; Lorist et al., 2000; Rushworth, Passingham et al., 2002; Swainson et al., 2003).

Despite the relative consistency of neuroimaging reports, several theorists (e.g., Logan, 2003; Monsell, 2003) have correctly pointed out that, as with behavioral RT costs, simple comparison of switch to repeat trials, even allowing for a preparation interval, can conflate multiple processing differences and do not distinguish between the agent and patient of control. But, unlike standard behavioral measures, neuroimaging data represent a multidimensional response measure. Hence, to the extent that different component processes are sensitive to different aspects of task control, one can potentially distinguish between the different processes that combine to produce a single RT.

### *State versus Item Effects*

Within-block explicit cueing procedures used in most imaging studies may mask important effects, particularly those that are sustained throughout a block (Donaldson, Petersen, Ollinger, & Buckner, 2001). For example, goal setting/task specification may not necessarily be captured in trial-by-trial variation. Indeed, as already noted, mixing costs (separable list-derived switch costs) are a well-established phenomenon (e.g., Meiran et al., 2000), and may correspond to a differential demand to set or maintain goals during a mixed block. Braver et al. (2003) compared state- and item-based switching effects using fMRI by testing mixed versus pure block effects and within-block switch versus repeat effects. As with the studies reviewed above, trial-based switching produced activation in left VLPFC. However, a sustained periodic component was associated with activity in lateral frontal polar cortex (FPC; BA 10), as it alternated between mixed and pure blocks.

In other experiments, in which subjects switch between abstract response rules or working memory manipulation tasks, FPC has been associated with the need to instantiate processing of abstract task representations already maintained by other PFC regions (Bunge, Kahn, Wallis, Miller, & Wagner, 2003; Sakai & Passingham, 2003). Indeed, this view of FPC in higher-level integration or maintenance/processing of subgoals is in line with other empirical perspectives on FPC function (Badre & Wagner, 2004; Braver & Bongiolatti, 2002; Bunge, Wendelken, Badre, & Wagner, 2005; Christoff & Gabrieli, 2000; Christoff et al., 2001; Dreher et al., 2002; Koechlin, Basso, Pietrini, Panzer, & Grafman, 1999). Hence, at least one interpretation of this effect is that activation in FPC reflects increased demands to set the goal or specify the task during a task switch, a demand that is more constant in trial-by-trial variation required by the explicit cueing procedure.

### *Neural Evidence for the Endogenous versus Exogenous Distinction*

As already reviewed, behavioral RT costs are quite durable, and will persist even after long preparatory intervals, and the residual switch cost has been interpreted by some to reflect a set of exogenous processes that come on-line upon presentation of a target and are distinct from intentional, endogenous processes (Meiran, 2000; Rogers & Monsell, 1995). Notably, however, the two-process view is not necessarily compelled given only the RT data. Two components may not reflect two processes but rather one process (or several processes) operating twice and at two points in time. For instance, a high field electrical recording study revealed no component present differentially before or after preparation (Wylie, Javitt, & Foxe, 2003b). Statistical modeling has also shown that one can account for residual costs probabilistically in terms of an increased likelihood of being prepared (Dejong, 2000; Nieuwenhuis & Monsell, 2002). Variables affecting this residual cost as a separable component have been difficult to obtain (Hubner, Kluwe, Luna-Rodriguez, & Peters, 2004). And even the classic data from Rogers and Monsell (1995) can be fit by a simple monotonic decay function.

Neuroimaging data can, however, provide some leverage on this controversy. To the extent that there are separable endogenous and exogenous control processes at temporally distinct epochs of a task switch, one might predict separable regions of PFC operating during task switching, one in advance of a target stimulus when a subject has foreknowledge and can endogenously prepare and the other at target onset when final adjustments must be performed. However, the fMRI studies directly testing such a prediction have produced only equivocal results (Brass & von Cramon, 2002; Kimberg et al., 2000; Luks et al., 2002; Sohn et al., 2000).

On the one hand, evidence for endogenous preparation processes prior to presentation of a target seem highly consistent. For example, a parietal discrimination potential that occurs almost immediately (140-150ms) after the onset of a target stimulus is beneficially impacted by preparation (Wylie, Javitt, & Foxe, 2003a), potentially reflective of a state of task-specific readiness. Furthermore, when a subject has foreknowledge of an upcoming task, prior to presentation of a target, there is preparation-related activation in VLPFC, pre-SMA, and parietal cortex (Brass & von Cramon, 2002; Konishi et al., 2002; Luks et al., 2002; Shulman et al., 2002; Sohn et al., 2000).

Critically, the preparation-related activation in VLPFC is greater on switch than repeat trials (Brass & von Cramon, 2002; Luks et al., 2002; Sohn et al., 2000), and this effect in left VLPFC occurs even when cues switch independent of task (Brass & von Cramon, 2004a). Beyond VLPFC, selective disruption of pre-SMA using transcranial magnetic stimulation has its greatest impact on switching performance during the pre-target preparation epoch (Rushworth, Hadland, Paus, & Sipila, 2002).

Unfortunately, the evidence for a separable exogenous process is less clear. Sohn et al. (2000) identified target-related (as temporally distinct from preparation-related) activation in posterior middle frontal gyrus (BA 8), a region clearly distinct from the VLPFC region associated with preparation for a switch. However, this activation was only present when subjects had no foreknowledge of a switch and could not have prepared. When a subject could prepare in advance, and hence only a truly residual process would be active, there were no reliable switch versus repeat differences in PFC at the presentation of the target stimulus. Similarly, other studies have either found preparation-related regions to be the same as target-related regions or found differential target-related activation in DLPFC only when the subject is unprepared (Brass & von Cramon, 2002; Dreher & Grafman, 2003; Dreher et al., 2002; Luks et al., 2002).

To the extent that differential regions of PFC are only active under these unprepared conditions, this may reflect a strategy difference more than a residual, exogenous adjustment process. For example, unable to prepare for a particular task, subjects may maintain multiple target responses in working memory and thus selection of an appropriate response becomes more like a dual-task than a serial task switch. Indeed, direct comparison of these two scenarios reveals DLPFC more active during a dual-task and VLPFC during a task switch (Dreher & Grafman, 2003). Likewise, DLPFC is more active when task sequences are themselves unpredictable (Dreher et al., 2002). It certainly may be the case that exogenous processes are real; there exists at least some evidence from ERP for two temporally separable stages of processing (Karayanidis et al., 2003; Lorist et al., 2000; Rushworth, Passingham et al., 2002). However, there is little fMRI evidence for such distinct processes.

### *Interference and Memory Retrieval in VLPFC*

As reviewed above, one potentially critical factor to emerge from behavioral investigation of task switching is the need to overcome interference while accessing memory during a task switch. Hence, it is critical to determine whether regions implicated in task switching and preparation for a task switch, such as left VLPFC, are sensitive to the effects of interference and access to memory.

Though only a handful of neuroimaging studies have directly examined interference phenomena within the context of task switching, several studies have interpreted their findings of lateral PFC and pre-SMA/ACC activation in terms of interference resolution (Brass & von Cramon, 2004b; Dove et al., 2000; Dreher & Berman, 2002; Dreher & Grafman, 2002; Dreher et al., 2002; Konishi et al., 2003; Nagahama et al., 2001; Swainson et al., 2003; Wylie et al., 2004). These studies often cite the well-established association between lateral PFC, particularly DLPFC, and response selection within the context of response interference/selection tasks (e.g. Badre & Wagner, 2004; Jiang & Kanwisher, 2003a, 2003b; Schumacher & D'Esposito, 2002; Schumacher, Elston, & D'Esposito, 2003). More directly, Brass and von Cramon reported switch sensitivity in left VLPFC to be greater during cross talk conditions, potential evidence of sensitivity to interference. Dreher and Berman (2002) investigated the backward inhibition effect (Mayr & Keele, 2000) using fMRI. Activation in left VLPFC was consistent with the standard behavioral effect, being greater when switching back to a recently performed task—the second A in the sequence A-B-A—than to a novel task—C in the sequence A-B-C. This result may be interpreted as evidence for interference processing within left VLPFC, and this pattern was distinct from that in ACC.

It is interesting, given the central role interference is thought to play in task switching, that ACC is not a more consistent finding in studies of task switching (but see Swainson et al., 2003). Indeed, following removal of the cingulate cortex, macaques showed little post-surgery impact on task switching beyond a generalized slowing of performance (Rushworth, Hadland, Gaffan, & Passingham, 2003). Potentially, the level of interference in task switching may be more abstract than that produced by competing responses typically associated with ACC activation, a hypothesis also supported by divergence in the behavioral phenomena (e.g. the cross-talk effect). Alternatively, some authors have proposed other control processes for ACC, such as a general monitoring or initiation function (Dreher & Berman, 2002; Luks et al., 2002) that might dictate its participation or absence from the network engaged during task switching. Further study may better resolve this potential discrepancy.

Results, such as the backward inhibition effect described above, suggest that VLPFC responds to the presence of interference, but the sensitivity of VLPFC to task set interference may not characterize its essential processing per se. Outside of the context of task switching, a large number of neuropsychological and neuroimaging studies have associated VLPFC with control during retrieval from long-term memory (Badre & Wagner, 2002; Petersen, Fox, Posner, Mintun, & Raichle, 1988; Poldrack et al., 1999; Wagner, Paré-Blagoev, Clark, & Poldrack, 2001). And this region may be engaged in a voluntary form of controlled retrieval or selection that is particularly important in the presence of interference from activated but irrelevant information (Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997; Thompson-Schill, D'Esposito, & Kan, 1999). Indeed, activation in left VLPFC will increase to the extent that a previously relevant (primed) feature of a cue becomes irrelevant, such as when a task changes, and so must be selected against. Critically, this selection-related increase in left VLPFC contrasts with across feature priming reductions in left temporal cortex (Thompson-Schill et al., 1999). In the present context, these results may implicate left VLPFC in overcoming interference due to priming of an irrelevant task set. As summarized previously, several theorists have identified primed competitor task associations within long term memory as a potentially central contributor to task switching effects (Allport & Wylie, 2000; Mayr & Kliegl, 2000; Rubinstein et al., 2001; Sohn & Anderson, 2001, 2003; Waszak et al., 2003), and a variety behavioral phenomena such as long-term carry-over effects (Sohn & Anderson, 2003; Waszak et al., 2003), restart costs (Allport & Wylie, 2000; Wylie & Allport, 2000), and increased switching costs with increased retrieval effort (Jersild, 1927; Mayr & Kliegl, 2000; Rubinstein et al., 2001) have provided empirical support. Hence, controlled retrieval/selection processing in VLPFC could play a central role in accessing memory and overcoming interference during task switching.

Consistent with a common mnemonic and task switching process in left VLPFC, paradigms combining both factors produce potentially informative interactions in left VLPFC responses. Gurd et al. (2002) scanned subjects while they performed fluency tasks. Subjects either had to generate from an unfamiliar semantic category for which greater control at retrieval is required or a well-learned sequence (e.g., days of the week) wherein retrieval proceeds more automatically. On "no switch" blocks, subjects generated from only one semantic category or one well-learned sequence, whereas on "switch" blocks subjects had to switch between two semantic categories or well-learned sequences. Consistent with previous studies of verbal fluency (Fiez, 1997; Thompson-Schill et al., 1997), generating from a single semantic category produced greater activation in left VLPFC than generating from a well-learned sequence. Likewise, switching produced greater activation in left VLPFC than not switching. However, there was a cross-over interaction such that switching between semantic categories produced less activation than generating from a single semantic category or switching between well-learned categories. In light of the current discussion, one interpretation of this finding is that, during switches, VLPFC is engaged in retrieving the category/task information and so has fewer resources to direct to individual items for the fluency task.

A similar task switching interaction in left VLPFC has been reported in the episodic domain. Reynolds et al., (2004) required subjects to switch between two semantic categorization tasks during a scanned encoding phase. Based on a subsequent surprise memory test, encode trials were binned on the basis of whether an item was subsequently remembered. Consonant with findings from past studies of episodic memory (e.g., Wagner et al., 1998), greater activation in left VLPFC predicted superior subsequent memory. Also in line with expectations, switching between encoding tasks produced greater activation in left VLPFC than repeating. However, memory for items that were encoded during a switch trial was worse than those encoded during a repeat trial. Hence, in this case, the standard subsequent memory effect in left VLPFC is reversed, such that greater activation (due to task switching) predicts worse memory. As with the previous interaction, this can be accounted for if a common VLPFC processor is engaged in retrieval for both task switching and semantic categorization. To the extent that this mechanism is engaged in selecting abstract task-set representations during a task switch it may be unable to select the item-specific information required to produce a durable trace. Hence, similar to the fluency study described above, this result also suggests that left VLPFC may be involved in a memory retrieval or selection process during task switching. Though the relationship of such a process to interference effects, such as transient and long-term carry-over, to which it is also sensitive (e.g., Brass & von Cramon, 2004b; Dreher & Berman, 2002) remains unclear.

### *Summary and Outline of Thesis*

Though evolving, the TSI, TSP, and TSR hypotheses still frame the theoretical debate over mechanisms of task switching. The evidence reviewed here suggests there are likely multiple components to task switching, including a mechanism of goal setting and a mechanism that retrieves or selects relevant task rules or a task set over interference. To the extent that demands on these processes are manipulated by task switching, it remains controversial whether behavioral costs reflect an active control process, differentially engaged during preparation for an upcoming task switch, or if, under circumstances such as explicit cueing, no active control process is necessary at all and interference forces the system to simply take more time to converge on a result otherwise specified by external cues. However, from at least one perspective, task switching effects may be understood in terms of changes occurring in long-term associative memory and the impact of these changes on subsequent performance. Consequently, control processes, such as those engaged to regulate memory (e.g., Shimamura, 1995), may be particularly critical during task switching.

Evidence from neuroimaging has started to provide some insights toward addressing these fundamental questions, and may initially support the hypothesis that control of memory is central to task switching. Left VLPFC is of particular interest as there are indications that it is sensitive both to switch preparation and to the effects of task switch interference. Hence, specifying in concrete terms the role that left VLPFC plays during task switching may provide important insights into the capacities and limitations in task switching and multitask performance.

The present thesis seeks to specify the role of left VLPFC in task switching, particularly as it relates to control of memory. An advantage to functional neuroimaging is its capacity to test neuroanatomically defined processing constructs across multiple task contexts. To the extent that theoretical constructs are manipulated carefully and prospectively across experiments, such an approach can provide multiple sources of complementary evidence on which to constrain and ground theory. Furthermore, this approach provides a direct means of assessing convergent validity, in addition to construct validity; a goal that is difficult to achieve with strictly behavioral measures. Therefore, we can gain confidence regarding the role of left VLPFC during task switching when informed by its role across analogous contexts. It is this approach that is embraced in this thesis.

In line with this approach, Chapters 2 and 3 present experiments that test the role of left VLPFC in resolving interference arising from memory. Critically, the forms of interference highlighted by these experiments are analogous to those proposed for task switching, but here are tested outside the context of a task switching regime. In particular, Chapter 2 describes an experiment designed to test resolution of proactive interference due to a recently relevant but currently irrelevant target within the context of short-term/working memory. Under such circumstances, presentation of a recently encountered but irrelevant item could cue automatic retrieval of irrelevant information and so elicit proactive interference, a case similar to that hypothesized by TSP to occur during task switching. In Chapter 3, two experiments are reported that test the resolution of interference from dominant but irrelevant representations during semantic retrieval. Interference from these competitors is modulated, in part, by automatic retrieval due to their strong association with a cue; again, an instance of retrieval induced interference analogous to that proposed for task switching. Critically, the experiments in Chapters 2 and 3 implicate a specific region of left mid-VLPFC, corresponding to left inferior frontal gyrus pars triangularis (BA 45), and suggest that this region mediates a control process that selects relevant retrieved representations from amidst competitors. Hence, given the correspondence of these sources of interference with that proposed by TSP for task switching, it is reasonable to hypothesize that left mid-VLPFC plays a similar role in task switching as it does in these experiments.

Having established a specific region of interest in left mid-VLPFC associated with the resolution of interference from memory, Chapter 4 describes a behavioral and fMRI experiment designed to provide insights toward resolution of the controversy surrounding control and the source of the task switch cost. More specifically, these experiments test the hypothesis that (1) behavioral costs are due to interference arising from differential retrieval of competing long-term memory representations during a task switch and (2) left VLPFC is engaged in a mnemonic control mechanism sensitive to this conflict and that biases selection of relevant retrieved representations. In an effort to be theoretically explicit, this chapter introduces a computational model in which task switch costs emerge exclusively from performance-based changes in the model's long-term associative structure and for which control is engaged via a top-down bias signal in preparation for a task switch. Critically, after validation on an independent data set from an initial behavioral experiment, a signature of conflict arising from retrieved conceptual representations is defined on the model. With this signature as a theoretical context, an fMRI experiment defines the pattern of task switching preparation and interference response characteristic of left mid-VLPFC. Critically, this pattern dissociates left mid-VLPFC from other regions associated with task switching and implicates this region as sensitive to conflict arising from retrieved conceptual representations.

Chapter 5 expands upon the computational model introduced in Chapter 4 in greater depth. In addition to detailing the model's architecture and dynamics, this chapter describes a number of simulations that demonstrate the model's capacity to account for a wide range of behavioral phenomena in the task switching literature. This chapter concludes with a discussion of the relationship between this model and extant formal models in the task switching literature.

In conclusion, I will argue in this thesis that these experiments provide convergent evidence for a mnemonic control process in left mid-VLPFC that selects relevant retrieved representations in response to conflict from retrieved competitors. Furthermore, this control mechanism plays a critical role in reducing the behavioral switch costs incurred during task switching, as these costs arise from competition among automatically retrieved long-term memory representations.

## References

Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and Performance XV: Conscious and nonconscious information processing* (pp. 421-452). Cambridge, MA: MIT Press.

Allport, A., & Wylie, G. (2000). 'Task-switching', stimulus-response bindings and negative priming. In S. Monsell & J. Driver (Eds.), *Control of Cognitive Processes: Attention and Performance XVIII* (pp. 35-70). Cambridge, MA: MIT Press.

Altmann, E. M. (2004). Advance preparation in task switching: what work is being done? *Psychol Sci, 15*(9), 616-622.

Altmann, E. M., & Gray, W. D. (2002). Forgetting to remember: the functional relationship of decay and interference. *Psychol Sci, 13*(1), 27-33.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychol Rev, 111*(4), 1036-1060.

Aron, A. R., Monsell, S., Sahakian, B. J., & Robbins, T. W. (2004). A componential analysis of task-switching deficits associated with lesions of left and right frontal cortex. *Brain, 127*(Pt 7), 1561-1573.

Arrington, C. M., & Logan, G. D. (2004). Episodic and semantic components of the compound-stimulus strategy in the explicit task-cuing procedure. *Mem Cognit, 32*(6), 965-978.

Badre, D., & Wagner, A. D. (2002). Semantic retrieval, mnemonic control, and prefrontal cortex. *Behavioral and Cognitive Neuroscience Reviews, 1*(3), 206-218.

Badre, D., & Wagner, A. D. (2004). Selection, integration, and conflict monitoring; assessing the nature and generality of prefrontal cognitive control mechanisms. *Neuron, 41*(3), 473-487.

Brass, M., & von Cramon, D. Y. (2002). The role of the frontal cortex in task preparation. *Cereb Cortex, 12*(9), 908-914.

Brass, M., & von Cramon, D. Y. (2004a). Decomposing components of task preparation with functional magnetic resonance imaging. *J Cogn Neurosci, 16*(4), 609-620.

Brass, M., & von Cramon, D. Y. (2004b). Selection for cognitive control: a functional magnetic resonance imaging study on the selection of task-relevant information. *J Neurosci, 24*(40), 8847-8852.

Braver, T. S., & Bongiolatti, S. R. (2002). The role of frontopolar cortex in subgoal processing during working memory. *NeuroImage, 15*, 523-536.

Braver, T. S., Reynolds, J. R., & Donaldson, D. I. (2003). Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron, 39*(4), 713-726.

Broadbent, D. E., & Gregory, M. (1967). Psychological refractory period and the length of time required to make a decision. *Proceedings of the Royal Society, Series B, 158*, 222-231.

Bunge, S. A., 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*, 3419-3428.

Bunge, S. A., Wendelken, C., Badre, D., & Wagner, A. D. (2005). Analogical reasoning and prefrontal cortex: evidence for separable retrieval and integration mechanisms. *Cereb Cortex, 15*(3), 239-249.

Christoff, K., & Gabrieli, J. D. E. (2000). The frontopolar cortex and human cognition: Evidence for a rostrocaudal hierarchal organization within the human prefrontal cortex. *Psychobiology, 28*, 168-186.

Christoff, K., Prabhakaran, V., Dorfman, J., Zhao, Z., Kroger, J. K., Holyoak, K. J., et al. (2001). Rostrocaudal prefrontal cortex involvement in relational integration during reasoning. *NeuroImage, 14*, 1136-1149.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97*, 332-361.

Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: A connectionist approach to behavior and biology in schizophrenia. *Psychological Review, 99*, 45-77.

Cools, R., Barker, R. A., Sahakian, B. J., & Robbins, T. W. (2003). L-Dopa medication remediates cognitive inflexibility, but increases impulsivity in patients with Parkinson's disease. *Neuropsychologia, 41*(11), 1431-1441.

Dejong, R. (2000). An intention-activation account of residual switch costs. In S. Monsell & J. Driver (Eds.), *Control of Cognitive Processes: Attention and Performance XVIII* (pp. 357-376). Cambridge, MA: MIT Press.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective attention. *Annual Review of Neuroscience, 18*, 193-222.

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

DiGirolamo, G. J., Kramer, A. F., Barad, V., Cepeda, N. J., Weissman, D. H., Milham, M. P., et al. (2001). General and task-specific frontal lobe recruitment in older adults during executive processes: a fMRI investigation of task-switching. *Neuroreport, 12*(9), 2065-2071.

Donaldson, D. I., Petersen, S. E., Ollinger, J. M., & Buckner, R. L. (2001). Dissociating state and item components of recognition memory using fMRI. *Neuroimage, 13*(1), 129-142.

Dove, A., Pollmann, S., Schubert, T., Wiggins, C. J., & von Cramon, D. Y. (2000). Prefrontal cortex activation in task switching: an event-related fMRI study. *Brain Res Cogn Brain Res, 9*(1), 103-109.

Dreher, J.-C., & Grafman, J. (2003). Dissociating the roles of the rostral anterior cingulate and the lateral prefrontal cortices in performing two tasks simultaneously or successively. *Cerebral Cortex, 13*, 329-339.

Dreher, J.-C., & Berman, K. F. (2002). Fractionating the neural substrate of cognitive control processes. *Proc Natl Acad Sci U S A, 99*(22), 14595-14600.

Dreher, J.-C., & Grafman, J. (2002). The roles of the cerebellum and basal ganglia in timing and error prediction. *Eur J Neurosci, 16*(8), 1609-1619.

Dreher, J.-C., Koechlin, E., Ali, S. O., & Grafman, J. (2002). The roles of timing and task order during task switching. *Neuroimage, 17*(1), 95-109.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of target letters in a non-search task. *Perception and Psychophysics, 16*, 143-149.

Fiez, J. A. (1997). Phonology, semantics, and the role of the left inferior prefrontal cortex. *Hum Brain Mapp, 5*(2), 79-83.

Gilbert, S. J., & Shallice, T. (2002). Task switching: a PDP model. *Cognit Psychol, 44*(3), 297-337.

Gopher, D., Armony, L., & Greenshpan, Y. (2000). Switching tasks and attention policies. *J Exp Psychol Gen, 129*(3), 308-339.

Gurd, J. M., Amunts, K., Weiss, P. H., Zafiris, O., Zilles, K., Marshall, J. C., et al. (2002). Posterior parietal cortex is implicated in continuous switching between verbal fluency tasks: an fMRI study with clinical implications. *Brain, 125*(Pt 5), 1024-1038.

Hubner, M., Kluwe, R. H., Luna-Rodriguez, A., & Peters, A. (2004). Response selection difficulty and asymmetrical costs of switching between tasks and stimuli: no evidence for an exogenous component of task-set reconfiguration. *J Exp Psychol Hum Percept Perform, 30*(6), 1043-1063.

Jersild, A. T. (1927). Mental set and shift. *Archives of Psychology*(Whole No. 89).

Jiang, Y., & Kanwisher, N. (2003a). Common neural mechanisms for response selection and perceptual action. *Journal of Cognitive Neuroscience, 15*(1095-1110).

Jiang, Y., & Kanwisher, N. (2003b). Common neural substrates for response selection across modalities and mapping paradigms. *Journal of Cognitive Neuroscience, 15*, 1082-1094.

Karayanidis, F., Coltheart, M., Michie, P. T., & Murphy, K. (2003). Electrophysiological correlates of anticipatory and poststimulus components of task switching. *Psychophysiology, 40*(3), 329-348.

Kimberg, D. Y., Aguirre, G. K., & D'Esposito, M. (2000). Modulation of task-related neural activity in task-switching: an fMRI study. *Brain Res Cogn Brain Res, 10*(1-2), 189-196.

Koch, I., Prinz, W., & Allport, A. (2005). Involuntary retrieval in alphabet-arithmetic tasks: Task-mixing and task-switching costs. *Psychol Res, 69*(4), 252-261.

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

Konishi, S., Hayashi, T., Uchida, I., Kikyo, H., Takahashi, E., & Miyashita, Y. (2002). Hemispheric asymmetry in human lateral prefrontal cortex during cognitive set shifting. *Proceedings of the National Academy of Sciences, USA, 99*(11), 7803-7808.

Konishi, S., Jimura, K., Asari, T., & Miyashita, Y. (2003). Transient activation of superior prefrontal cortex during inhibition of cognitive set. *J Neurosci, 23*(21), 7776-7782.

Logan, G. D. (2003). Executive control of thought and action: In search of the wild homunculus. *Current Directions in Psychological Science, 12*(2), 45-48.

Logan, G. D., & Bundesen, C. (2003). Clever homunculus: is there an endogenous act of control in the explicit task-cuing procedure? *J Exp Psychol Hum Percept Perform, 29*(3), 575-599.

Logan, G. D., & Bundesen, C. (2004). Very clever homunculus: compound stimulus strategies for the explicit task-cuing procedure. *Psychon Bull Rev, 11*(5), 832-840.

Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychol Rev, 108*(2), 393-434.

Lorist, M. M., Klein, M., Nieuwenhuis, S., De Jong, R., Mulder, G., & Meijman, T. F. (2000). Mental fatigue and task control: planning and preparation. *Psychophysiology, 37*(5), 614-625.

Luks, T. L., Simpson, G. V., Feiwell, R. J., & Miller, W. L. (2002). Evidence for anterior cingulate cortex involvement in monitoring preparatory attentional set. *Neuroimage, 17*(2), 792-802.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychol Bull, 109*(2), 163-203.

Mayr, U. (2001). Age differences in the selection of mental sets: the role of inhibition, stimulus ambiguity, and response-set overlap. *Psychol Aging, 16*(1), 96-109.

Mayr, U., & Keele, S. W. (2000). Changing internal constraints on action: The role of backward inhibition. *Journal of Experimental Psychology: General, 129*(1), 4-26.

Mayr, U., & Kliegl, R. (2000). Task-set switching and long-term memory retrieval. *J Exp Psychol Learn Mem Cogn, 26*(5), 1124-1140.

Mayr, U., & Kliegl, R. (2003). Differential effects of cue changes and task changes on task-set selection costs. *J Exp Psychol Learn Mem Cogn, 29*(3), 362-372.

Mecklinger, A. D., von Cramon, D. Y., Springer, A., & Matthes-von Cramon, G. (1999). Executive control functions in task switching: evidence from brain injured patients. *J Clin Exp Neuropsychol, 21*(5), 606-619.

Meiran, N. (2000). Modeling cognitive control in task-switching. *Psychological Research, 63*, 234-249.

Meiran, N., Chorev, Z., & Sapir, A. (2000). Component processes in task switching. *Cognit Psychol, 41*(3), 211-253.

Meiran, N., Gotler, A., & Perlman, A. (2001). Old age is associated with a pattern of relatively intact and relatively impaired task-set switching abilities. *J Gerontol B Psychol Sci Soc Sci, 56*(2), P88-102.

Meyer, D. E., Evans, J. E., Lauber, E. J., Gmeindl, L., Rubinstein, J., Junck, L., et al. (1998, April). *The role of dorsolateral prefrontal cortex for executive cognitive processes in task switching.* Paper presented at the Cognitive Neuroscience Society, San Francisco, CA.

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanisms. *Psychol Rev, 104*(1), 3-65.

Meyer, D. E., Lauber, E. J., Rubinstein, J., Gmeindl, L., Junck, L., & Koeppe, R. A. (1997, March, 23-25). *Activation of brain mechanisms for executive mental processes in cognitive task switching.* Paper presented at the Cognitive Neuroscience Society, Boston, MA.

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

Monsell, S. (2003). Task switching. *Trends Cogn Sci, 7*(3), 134-140.

Nagahama, Y., Okada, T., Katsumi, Y., Hayashi, T., Yamauchi, H., Oyanagi, C., et al. (2001). Dissociable mechanisms of attentional control within the human prefrontal cortex. *Cereb Cortex, 11*(1), 85-92.

Nieuwenhuis, S., & Monsell, S. (2002). Residual costs in task switching: Testing the 'failure to engage' hypothesis. *Psychon Bull Rev, 9*, 86-92.

Norman, D. A., & Shallice, T. (1986). Attention to action: willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz & D. Shapiro (Eds.), *Consciousness and Self-regulation* (Vol. 4, pp. 1-18). New York: Plenum Press.

Pashler, H. (1994). Dual-task interference in simple tasks: data and theory. *Psychol Bull, 116*(2), 220-244.

Penfield, W. (1926). The career of Ramon y Cajal. *Archives of Neurology and Psychiatry, 16*, 213-220.

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single- word processing. *Nature, 331*(6157), 585-589.

Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage, 10*(1), 15-35.

Reynolds, J. R., Donaldson, D. I., Wagner, A. D., & Braver, T. S. (2004). Item- and task-level processes in the left inferior prefrontal cortex: positive and negative correlates of encoding. *Neuroimage, 21*(4), 1472-1483.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124*(2), 207-231.

Rogers, R. D., Sahakian, B. J., Hodges, J. R., Polkey, C. E., Kennard, C., & Robbins, T. W. (1998). Dissociating executive mechanisms of task control following frontal lobe damage and Parkinson's disease. *Brain, 121 ( Pt 5)*, 815-842.

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *J Exp Psychol Hum Percept Perform, 27*(4), 763-797.

Ruge, H., Brass, M., Koch, I., Rubin, O., Meiran, N., & von Cramon, D. Y. (2005). Advance preparation and stimulus-induced interference in cued task switching: further insights from BOLD fMRI. *Neuropsychologia, 43*(3), 340-355.

Rushworth, M. F. S., Hadland, K. A., Paus, T., & Sipila, P. K. (2002). Role of the human medial frontal cortex in task switching: a combined fMRI and TMS study. *J Neurophysiol, 87*(5), 2577-2592.

Rushworth, M. F. S., Passingham, R. E., & Nobre, A. C. (2002). Components of switching intentional set. *J Cogn Neurosci, 14*(8), 1139-1150.

Rushworth, M. F. S., Hadland, K. A., Gaffan, D., & Passingham, R. E. (2003). The effect of cingulate cortex lesions on task-switching and working memory. *Journal of Cognitive Neuroscience, 15*(3), 338-353.

Sakai, K., & Passingham, R. E. (2003). Prefrontal interactions reflect future task operations. *Nat Neurosci, 6*(1), 75-81.

Schumacher, E. H., & D'Esposito, M. (2002). Neural implementation of response selection in humans as revealed by localizing effects of stimulus-response compatibility on brain activation. *Human Brain Mapping, 17*, 193-201.

Schumacher, E. H., Elston, P. A., & D'Esposito, M. (2003). Neural evidence for representation-specific response selection. *Journal of Cognitive Neuroscience, 15*, 1111-1121.

Shimamura, A. P. (1995). Memory and frontal lobe function. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences* (pp. 803-813). Cambridge, MA: MIT Press.

Shulman, G. L., d'Avossa, G., Tansy, A. P., & Corbetta, M. (2002). Two attentional processes in the parietal lobe. *Cereb Cortex, 12*(11), 1124-1131.

Smith, E. E., Geva, A., Jonides, J., Miller, A., Reuter-Lorenz, P., & Koeppe, R. A. (2001). The neural basis of task-switching in working memory: effects of performance and aging. *Proc Natl Acad Sci U S A, 98*(4), 2095-2100.

Sohn, M.-H., & Anderson, J. R. (2001). Task preparation and task repetition: Two-component model of task switching. *Journal of Experimental Psychology: General, 130*(4), 764-778.

Sohn, M.-H., & Anderson, J. R. (2003). Stimulus-related priming during task-switching. *Memory and Cognition, 31*(5), 775-780.

Sohn, M.-H., Ursu, S., Anderson, J. R., Stenger, V. A., & Carter, C. S. (2000). Inaugural article: the role of prefrontal cortex and posterior parietal cortex in task switching. *Proc Natl Acad Sci U S A, 97*(24), 13448-13453.

Spector, A., & Biederman, I. (1976). Mental set and mental shift revisited. *American Journal of Psychology, 89*, 669-679.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30*, 276-315.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643-662.

Swainson, R., Cunnington, R., Jackson, G. M., Rorden, C., Peters, A. M., Morris, P. G., et al. (2003). Cognitive control mechanisms revealed by ERP and fMRI: evidence from repeated task-switching. *J Cogn Neurosci, 15*(6), 785-799.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proc Natl Acad Sci U S A, 94*(26), 14792-14797.

Thompson-Schill, S. L., D'Esposito, M., & Kan, I. P. (1999). Effects of repetition and competition on activity in left prefrontal cortex during word generation. *Neuron, 23*(3), 513-522.

Wagner, A. D., Paré-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning: Left prefrontal cortex guides controlled semantic retrieval. *Neuron, 31*, 329-338.

Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., et al. (1998). Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. *Science, 281*(5380), 1188-1191.

Wallis, J. D., & Miller, E. K. (2003). From rule to response: Neuronal processes in the premotor and prefrontal cortex. *Journal of Neurophysiology, 90*, 1790-1806.

Waszak, F., Hommel, B., & Allport, A. (2003). Task-switching and long-term priming: role of episodic stimulus-task bindings in task-shift costs. *Cognit Psychol, 46*(4), 361-413.

Wylie, G., & Allport, A. (2000). Task switching and the measurement of "switch costs". *Psychol Res, 63*(3-4), 212-233.

Wylie, G. R., Javitt, D. C., & Foxe, J. J. (2003a). Cognitive control processes during an anticipated switch of task. *Eur J Neurosci, 17*(3), 667-672.

Wylie, G. R., Javitt, D. C., & Foxe, J. J. (2003b). Task switching: a high-density electrical mapping study. *Neuroimage, 20*(4), 2322-2342.

Wylie, G. R., Javitt, D. C., & Foxe, J. J. (2004). Don't think of a white bear: An fMRI investigation of the effects of sequential instructional sets on cortical activity in a task-switching paradigm. *Hum Brain Mapp, 21*, 279-297.

Yeung, N., Cohen, J. D., & Botvinick, M. M. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychol Rev, 111*(4), 931-959.

Yeung, N., & Monsell, S. (2003). The effects of recent practice on task switching. *J Exp Psychol Hum Percept Perform, 29*(5), 919-936.

# Chapter 2

## Frontal Lobe Mechanisms that Resolve Proactive Interference

As introduced in Chapter 1, a cue previously associated with one task can produce interference when it is subsequently encountered during performance of another task, giving rise to the behavioral switch cost. In terms of traditional memory theory, such interference can be thought of as a form of proactive interference (PI), in that a memory formed during a past experience can interfere with processing during a subsequent experience. In the case of task switching, this PI comes about when a stimulus associated with a competing task automatically cues retrieval of the competing task set. This retrieved information can compete with relevant target information for processing and so produce PI. The goal of the present chapter is to investigate the PFC cognitive control mechanisms that contribute to PI resolution. Results from this experiment will be important in (1) identifying a precise region of PFC (i.e., mid-VLPFC) specifically associated with resolution of PI, and so implicating this region as a candidate to perform a similar role during task switching, and (2) to provide some empirical constraints on the hypothesized control mechanisms supported by this region.

Recently, neuroimaging and neuropsychological evidence have implicated left mid-ventrolateral prefrontal cortex (mid-VLPFC) in PI resolution during short-term item recognition, though the precise mechanisms await specification. This chapter describes an fMRI experiment, currently in press at *Cerebral Cortex*, which seeks to advance theorizing regarding PI resolution. On each trial, subjects maintained a target set of words, and then decided if a subsequent probe was contained in the target set (positive) or not (negative). Importantly, for half of the negative and half of the positive trials, the probe had been contained in the previous target set (recent). Relative to non-recent trials, negative–recent trials produced an increase in response times and error rates, behavioral markers of PI. In fMRI measures, negative recency was associated with increased activation in left mid-VLPFC, as well as in bilateral fronto-polar cortex, providing evidence for multiple components in PI resolution. Furthermore, recency effects were evident during both negative and positive trials, with the magnitude of the recency effect in mid-VLPFC being greater on negative trials. Collectively, these results serve to specify and constrain proposed models of PI resolution, and further, serve as a basis for understanding the potential control processes brought on-line to resolve interference during task switching.

## Introduction

Memory for the past shapes processing in the present, though such influences sometimes prove detrimental, a phenomenon termed *proactive interference* (PI). For example, the frustrating experience of forgetting where your car is parked in a regularly used lot is attributable, at least in part, to interference from memories established during prior occasions of parking your car in the lot. The importance of PI as a fundamental processing constraint in memory and cognition is well recognized, being highlighted in classic work implicating interference as a cause of forgetting from long-term memory (McGeoch, 1942) as well as a source for age-related declines in cognitive function (Hasher & Zacks, 1988). Importantly, PI can constrain active memory processing, potentially contributing to short-term forgetting (Brown, 1958; Keppel & Underwood, 1962; Peterson & Peterson, 1959). It follows then that processes that resolve or resist PI may be critical for the flexible updating and maintenance of task-relevant goals, stimuli, and responses. Given the processing costs of interference, specifying the neural mechanisms that overcome or resolve PI is a fundamental objective (Shimamura, 1995).

An illustrative paradigm in which mechanisms that resolve PI have been examined is short-term item recognition (Figure 1), wherein subjects judge whether a probe stimulus matches (*positive*) or mismatches (*negative*) one of a set of maintained target stimuli (Jonides, Smith, Marshuetz, Koeppe, & Reuter-Lorenz, 1998; Monsell, 1978). During the critical condition, trials are arranged such that the current probe overlaps the target set on the previous trial (*recent*). By this arrangement, negative–recent probes, though not members of the currently maintained target set, nevertheless give rise to a sense of familiarity due to their presence in the previous target set, an attribute they share with positive probes (which are items in the currently maintained target set). Hence, the familiarity of negative–recent probes is thought to elicit conflict at the response and/or stimulus representation levels (D'Esposito, Postle, Jonides, & Smith, 1999; Jonides et al., 1998; Mecklinger, Weber, Gunter, & Engle, 2003; Nelson, Reuter-Lorenz, Sylvester, Jonides, & Smith, 2003). Behaviorally, this PI-derived conflict is reflected in elevated response times (RT) and errors to negative–recent probes relative to negative–non-recent probes. At the neural level, functional imaging studies have revealed increased activation in left mid-ventrolateral prefrontal cortex (VLPFC; ~Brodmann's area [BA] 45) during negative–recent relative to negative–non-recent trials (Bunge, Ochsner, Desmond, Glover, & Gabrieli, 2001; D'Esposito et al., 1999; Jonides et al., 2000; Jonides et al., 1998; Mecklinger et al., 2003; Nelson et al., 2003), and neuropsychological data indicate that lesions of this PFC region result in greatly enhanced PI-related errors and response slowing (Thompson-Schill et al., 2002). Thus, a left mid-VLPFC mechanism appears to contribute to resolving PI during short-term item recognition, though the nature of this mechanism remains underspecified.

Extant accounts of left mid-VLPFC involvement in PI resolution have focused on the potential role of this region in inhibiting or selecting against the irrelevant attribution of familiarity to the negative–recent probe (Jonides et al., 1998; Nelson et al., 2003) or familiarity-triggered inappropriate response tendencies (D'Esposito et al., 1999; Jonides et al., 1998). From these perspectives, interference on negative–recent trials arises directly or indirectly from conflict between the familiar nature of the probe (due to its presence in the previous target set) and the status of the probe as a non-member of the current target set. The resolution of this conflict may proceed through inhibition of the familiar representation, the inappropriate response, or the attribution of familiarity (D'Esposito et al., 1999; Jonides et al., 1998; Nelson et al., 2003). Hence, it is posited that conflict resolution in this task has the effect of bringing responses arising from mnemonic signals, including familiarity, in line with response decision criteria that depend on the temporal context of a given probe. To date, theorizing has largely considered this mechanism to be restricted in scope and interaction, being active principally on negative–recent trials and not necessarily requiring the simultaneous operation of additional cognitive control processes.

Further insight into the nature of the mechanism(s) supporting PI resolution may be gleaned from consideration of the neural correlates associated with performance of other memory paradigms that require determining the context in which a familiar item was previously encountered in order to guide a response. For example, in episodic retrieval tasks that require context or source recollection, subjects must determine the context in which a familiar memory probe was encountered (Johnson, Hashtroudi, & Lindsay, 1993; Johnson & Raye, 1981). Similarly, in the N-back working memory task, performance partially rests on determining whether a familiar probe occurred within a specific temporal context (N trials back) or a different temporal context (N+1 or N-1 trials back) (Braver et al., 2001; Nyberg et al., 2003). Intriguingly, in contrast to studies of PI resolution during short-term item recognition, these other mnemonic tasks have strongly favored a multi-component cognitive control system that both guides on-going processes as well as evaluates/monitors the results of those processes and integrates them with currently maintained goals or decision criteria (Buckner, 2003; Cabeza, Dolcos, Graham, & Nyberg, 2002; Dobbins, Foley, Schacter, & Wagner, 2002; Johnson, Raye, Mitchell, Greene, & Anderson, 2003; Wagner, 1999).

Fronto-polar cortex (FPC), in particular, has been implicated in evaluation, monitoring, or integration processes (Badre & Wagner, 2004; Braver & Bongiolatti, 2002; Bunge, Badre, Wendelken, & Wagner, in press; Christoff, Ream, Geddes, & Gabrieli, 2003; Dobbins et al., 2002; Koechlin, Basso, Pietrini, Panzer, & Grafman, 1999). In the domain of working memory, activation in FPC has been reported during N-back tasks relative to control conditions (Braver et al., 2001; Nyberg et al., 2003). Similarly, tasks that involve selecting responses or response rules from working memory based on maintained contextual rules also implicate FPC in integrating across time and goals (Badre & Wagner, 2004; Braver, Reynolds, & Donaldson, 2003; Bunge, Kahn, Wallis, Miller, & Wagner, 2003; Christoff et al., 2003; Koechlin et al., 1999; Koechlin, Ody, & Kouneiher, 2003; Sakai & Passingham, 2003). Finally, FPC activation has been consistently observed during studies of episodic retrieval (Buckner et al., 1995; Fletcher & Henson, 2001; Nyberg et al., 1996; Rugg, Fletcher, Frith, Frackowiak, & Dolan, 1996; Squire, 1992; Tulving, Kapur, Craik, Moscovitch, & Houle, 1994; Wagner, Desmond, Glover, & Gabrieli, 1998), with left FPC being particularly engaged during retrieval tasks that require assignment of a familiar item to the source in which it was previously encountered (Cansino, Maquet, Dolan, & Rugg, 2002; Dobbins, Rice, Wagner, & Schacter, 2003; Nolde, Johnson, & D'Esposito, 1998; Ranganath, Johnson, & D'Esposito, 2000). Such effects have been interpreted as reflecting sub-processes during retrieval, wherein the mnemonic products elicited by current retrieval cues are evaluated/monitored and integrated with decision criteria (Dobbins et al., 2002; Rugg & Wilding, 2000).

Given these observations, one might anticipate that PI resolution will also depend on FPC processes, given the importance of monitoring the relationship between target familiarity and the criterial context (in this case, membership in the current target set). Intriguingly, though no study to date has reported reliable activation in FPC during negative–recent vs. non-recent trials, the inaugural PET study of PI during short-term item recognition appears to have observed subthreshold activation in FPC (see Figure 1-[High-Recency minus Low-Recency] in Jonides et al., 1998), raising the possibility that a multi-component cognitive control network operates in the service of resolving PI. One objective of the present study was to directly consider the sensitivity of FPC to interference during short-term item recognition.

Beyond determining whether PI resolution depends on a unitary mechanism or multiple control mechanisms, the nature of the inhibitory/selection mechanism putatively supported by left mid-VLPFC awaits further specification. Leverage on the nature of left mid-VLPFC contributions to PI resolution may come from considering how this region is engaged by positive–recent trials, on which the test probe is a member of the current *and* the preceding target set, such that familiarity is in concert with current set membership and a positive response. Critically, all positive test probes are familiar, as they were recently encountered in the current target set. However, relative to positive–non-recent probes, positive–recent probes may possess enhanced familiarity, having been present in the previous target set as well the current target set. The impact of this enhanced familiarity on positive–recent relative to positive–non-recent trials has yet to be formally investigated.

Both congruency-based decreases and congruency-based increases in neural activation have precedent elsewhere in the cognitive control literature, and have provided important constraints on mechanistic models of cognitive control. For example, functional imaging studies of the Stroop task (Stroop, 1935) have revealed facilitative *decreases* in dorsolateral PFC (DLPFC) and anterior cingulate cortex activation on congruent trials (e.g., when the word names the color of ink in which it is printed) compared to neutral trials, suggesting that these activation reductions reflect decreased demands on response selection processes due to multiple converging sources of evidence in support of a response (Carter, Mintun, & Cohen, 1995; Milham et al., 2001; Milham et al., 2002). By contrast, *increases* in left VLPFC on congruent compared to neutral Stroop trials have also been reported, with these effects being similar to those observed for incongruent trials (Milham et al., 2001; Milham et al., 2002). In this case, congruent trial increases were interpreted as reflecting general sensitivity to multiple sources of response relevant information (e.g., a nameable color and word), even if these sources were not in conflict. As with Stroop, inspection of the neural responses during positive–recent trials during short-term item recognition may provide additional leverage on the interference-resolution mechanism putatively mediated by left mid-VLPFC.

The present functional MRI (fMRI) experiment was designed to further specify the PFC mechanisms that resolve short-term PI. In particular, the design and analyses emphasized two novel aspects of the functional imaging data, with the goal of providing constraints on theorizing regarding PI resolution. First, voxel-based, region-of-interest, and cross-experiment convergence analyses assessed the multi-component nature of PI resolution, focusing on the impact of recency on activation in FPC as well as in VLPFC.

Second, the impact of the enhanced familiarity during positive–recent trials was assessed in these regions. Finally, an attempt was made to assess the domain generality of the neural responses to PI by including verbal (words) and non-verbal (visual patterns) stimuli in separate blocks, though as will be described, issues with behavioral performance complicate interpretation of the data from the non-verbal condition.

## *Methods*

### *Subjects*

Seventeen right-handed, native English speakers (7 female; ages 18-31 yrs) gave informed consent in a manner approved by the institutional review boards at Massachusetts General Hospital and MIT. Data from three additional subjects were excluded due to fMRI spike artifacts. Subjects received $50 remuneration.

### *Design*

Subjects performed alternate blocks of a short-term item recognition task using Words and Patterns. Word stimuli consisted of 20 five-letter, one-syllable abstract nouns; word frequency (Mean = 114) was matched across experimental conditions. Pattern stimuli consisted of 20 abstract visual patterns. On each trial in the Words blocks (Figure 1a), subjects were presented with a memory set of four target words about a central fixation cross. Subjects had 2.5 s to encode the set, 3 s to maintain the set over a delay, and then were centrally presented a word probe. Upon probe presentation, subjects were to endorse the probe as matching one of the items in the currently maintained memory set or reject it as a non-match. Subjects had 2.5 s in which to respond, pressing one of two buttons under their left hand; failures to respond were scored as incorrect.

Across all Words trials, 50% of the probes matched one of the items in the current memory set (Positive probes) and 50% were non-matches (Negative probes).

Trial events in the Patterns blocks paralleled those in the Words blocks, except that stimuli were abstract visual patterns (Figure 1b). Abstract visual patterns (fractals) were selected with the goal of minimizing participants' ability to adopt a verbal naming strategy. Based on a pre-experimental questionnaire (N = 10), 20 difficult-to-name abstract visual patterns were selected from a set of 91, selecting those for which the fewest number of participants were able to generate a single name for the items (median and mode = 2/10 participants; max = 3/10 participants). On a post-scanning questionnaire, fMRI participants reported naming an average of 7 of the 20 images (SD = 4).

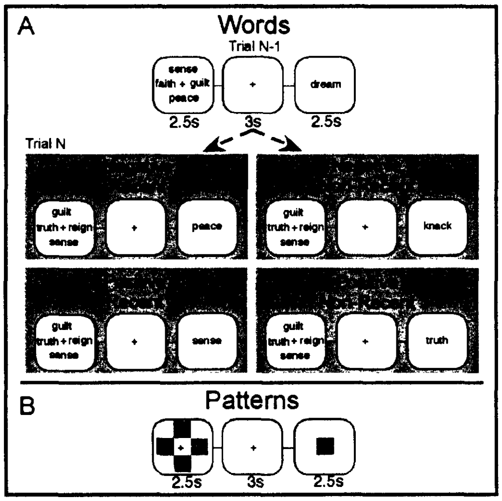

**Figure 1.** Schematics depicting the trial elements (upper) and the four critical conditions (grey) for the Words task (A) and a sample trial from the Patterns task (B). Probe type reflects the presence (Positive) or absence (Negative) of the probe in the current trial's memory set (Trial N). Recency reflects the presence (Recent) or absence (Non-Recent) of the probe in the previous trial's memory set (Trial N-1).

The relation between the probe on a given trial (*trial n*) and the items presented in the previous trial (*trial n-1*) was arranged to elicit proactive interference (PI) on a subset of events. Specifically, memory sets were constructed such that two items from the previous trial's memory set were repeated in the current trial's memory set, with the additional constraint that no single item was repeated more than three times in a row. This arrangement permitted 50% of all probes to match (Recent) one of the items in the *previously* encoded, but now irrelevant, memory set (i.e., *trial n-1*) and 50% to mismatch (Non-Recent) any of the previously (*trial n-1*) encoded items. Hence, on Recent trials, the recent exposure to a given stimulus on *trial n-1* should elicit a sense of familiarity when encountering the probe on *trial n*, and thus could potentially give rise to PI. Of the Recent trials, half entailed a Positive probe (i.e., the stimulus was in the *trial n* and *n-1* memory sets) and half a Negative probe (i.e., the stimulus was not in *trial n* but was in *n-1*). It is of further note that because all memory sets had an equal number of items repeating from the previous trial's memory set (n = 2) and of items that were not in the prior set (n = 2), any effects of Recency or Probe Type were restricted to the probe phase of a trial. In addition to isolating the effects of familiarity to that deriving from the probe, this design also permitted event-related analysis to focus specifically on the probe phase (see below).

Subjects encountered 240 Words trials and 240 Patterns trials, which were divided into 16 fifteen-trial blocks of each stimulus type. Within each of 4 scan runs, subjects alternated between 4 Words and 4 Patterns blocks in an ABBA fashion. Blocks were separated by 12-s periods during which a cue was presented that named the upcoming task (WORD or PATTERN). Although stimulus type was blocked, within each block, trial types (Positive–Recent, Negative–Recent, Positive–Non-Recent, and Negative–Non-Recent) were intermixed in an event-related manner, with variable duration fixation events (0-4 s) permitting signal deconvolution (Dale, 1999). The order of blocks and stimuli were counterbalanced across subjects.

### *fMRI Procedures*

Whole-brain imaging was performed on a 3T Siemens Allegra system. Functional data were acquired using a gradient-echo echo-planar sequence (TR=2 s, TE=30 ms, 21 axial slices, 3.125 x 3.125 x 5 mm, 1 mm skip, 692 volumes/run). High-resolution T1-weighted (MP-RAGE) anatomical images were collected for visualization. Head motion was restricted using firm padding that surrounded the head. Visual stimuli were projected onto a screen, and viewed through a mirror attached to a standard head coil.

Data were preprocessed using SPM99 (http://www.fil.ion.ucl.ac.uk/spm/). Functional images were corrected for differences in slice acquisition timing, followed by motion correction (using sinc interpolation). Structural and functional data were spatially normalized to the MNI template (Cocosco, Kollokian, Kwan, & Evans, 1997)—an approximation of canonical space (Talairach & Tournoux, 1988)—using a 12-parameter affine transformation along with a nonlinear transformation using cosine basis functions. Images were resampled into 3-mm cubic voxels and spatially smoothed with an 8-mm FWHM isotropic Gaussian kernel.

### *fMRI Analyses*

Statistical models were constructed using SPM99 under the assumptions of the general linear model. Trial events were modeled as two components—encoding/delay and probe. As described above, for each Content type (Words or Patterns), events from the initial portion of each trial—encoding of the memory set and maintenance of the set across the delay—were identical across all four probe conditions (i.e., Recent/Non-Recent × Negative/Positive). Accordingly, the encoding/delay period was modeled as a 5.5-s epoch according to Content, irrespective of probe condition. Hence, though the presentation of the memory set and the delay period always preceded presentation of the probe, the regressors for each phase (e.g., encoding/delay versus probe) were sufficiently uncorrelated to permit assessment of the unique contribution of each to the overall variance in MRI signal. The probe portion of the trial, which corresponds to the data of interest, was modeled as an event according to Content (Words/Patterns), Probe (Positive/Negative), and Recency (Recent/Non-Recent). Correct trials were modeled separately from incorrect trials.

Effects were estimated using a subject-specific fixed-effects model, with session effects treated as confounds. Estimates were entered into a second-level group analysis treating subjects as a random effect, using a one-sample *t*-test against a contrast value of zero at each voxel. Unless otherwise noted, regions were considered reliable to the extent that they consisted of at least 5 contiguous voxels that exceeded an uncorrected threshold of $p < .001$.

To reveal common effects of Recency (i.e., familiarity) across Positive and Negative trials in *a priori* predicted regions (i.e., left mid-VLPFC and FPC), a conjunction analysis was performed at a conjoint alpha level of $p < .001$ at each voxel, using the independent contrasts Negative–Recent > Non-Recent and Positive–Recent > Non-Recent (each thresholded at $p < .032$). This method may be interpreted as setting the probability of a Type I error occurring in both contrasts to be less than .001. That is, a significant conjunction does not indicate that both contrasts were individually significant at standard thesholds (Nichols, Brett, Andersson, Wager, & Poline, in press), but rather means that both were significant at more lenient thresholds (with a joint probability of a Type I being less than .001).

Voxel-based contrasts were supplemented with region-of-interest (ROI) analyses that provided quantitative characterization of the effects. ROIs included all significant voxels within a 6-mm radius of each *a priori* targeted maximum. Selective averaging permitted extraction of the peak percent signal change associated with each condition; ROI data were submitted to repeated-measures analyses of variance.

## Results

### *Behavioral Markers of PI*

Accuracy and reaction time (RT; restricted to correct trials) data revealed three central results that ground understanding of the imaging data (Figure 2). (A) Though subjects performed well, Patterns trials were more difficult than Words trials. (B) PI was observed, wherein encountering a Negative–Recent probe resulted in interference (increased errors and longer RTs) relative to encountering a Negative–Non-Recent probe. (C) Though within-Content Recency effects in RT and accuracy were restricted to Negative–Recent trials, Recency main effects were evident in accuracy for both Negative and Positive probes collapsed across Content type.

The Patterns condition was more difficult. Subjects responded more slowly ($\underline{F}(1,16) = 17.0$, $p < .001$) and considerably less accurately ($\underline{F}(1,16) = 70.4$, $p < .0001$) on Patterns than on Words trials (RT: 1138 vs. 1016 ms; Proportion Errors: .23 vs. .06, respectively). The increased error rate was partially attributable to a bias to respond 'non-match' on Patterns trials, evident in a reliable interaction of Content (Words/Patterns) × Probe (Positive/Negative) on accuracy (F(1,16) = 31.3, $p < .0001$).

Recency affected RT, such that overlap of the current probe with one of the memoranda in the preceding trial slowed RTs (F(1,16) = 81.4, $p < .0001$). Although subjects were slower on Recent (1108 ms) relative to Non-Recent trials (1046 ms), a Probe × Recency interaction (F(1,16) = 38.3, $p < .0001$) indicated that this Recency-induced RT slowing was reliable for Negative probes (1139 vs. 1023 ms; F(1,16) = 86.0, $p < .0001$) but not for Positive probes (1076 vs. 1070 ms; F < 1). Recency-induced slowing was present for both Words (F(1,16) = 37.3, $p < .0001$) and Patterns (F(1,16) = 47.7, $p < .0001$).

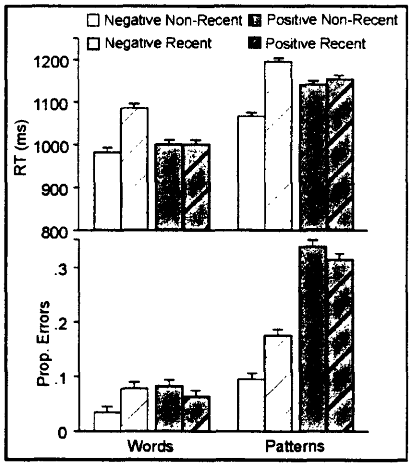

**Figure 2.** Behavioral reaction time (RT) and proportion of errors.

Though Recency did not reliably affect RT to endorse Positive probes, Recency impacted accuracy on Positive and Negative trials, but in different directions (Figure 2). Specifically, although overall accuracy was lower on Recent relative to Non-Recent trials (Proportion Errors: .16 vs. .14; F(1,16) = 17.3, $p < .001$), a Probe × Recency interaction (F(1,16) = 45.1, $p < .0001$) indicated that, consistent with PI, accuracy declined on

Negative–Recent relative to Negative–Non-Recent trials (Proportion Errors: .13 vs. .07; F(1,16) = 49.2, $p < .0001$), whereas accuracy modestly increased on Positive–Recent relative to Positive–Non-Recent trials (Proportion Errors: .19 vs. .21; F(1,16) = 6.2, $p < .05$). The Recency-induced decline in accuracy on Negative trials was present when collapsing across both Words and Patterns (Fs > 9.3, *p*s < .01). By contrast, the modest increase in accuracy on Positive trials was not reliable when considering Words or Patterns alone (Fs < 3.1, *p*s > .09).

### *Neural Responses to the Probe*

Initial fMRI analyses identified structures that were active when generating a response during the probe stage, collapsed across all conditions. Comparison of All Probes vs. the fixation baseline revealed broad activation inclusive of regions commonly associated with cognitive control and working memory, including bilateral VLPFC (~BA 44/45), bilateral anterior DLPFC (~BA 9/46/10) and left mid-DLPFC (~BA 9/46), and superior (~BA 7) and inferior parietal cortices (~BA 40). In addition, bilateral activation was observed in posterior hippocampus/parahippocampal gyrus, and in fusiform and lateral temporal cortices (a complete list of coordinates is available upon request).

We primarily consider the effects of Recency and Probe type on activation during Words trials, as these conditions are most analogous to previous reports that investigated PI using verbal stimuli (in particular, published reports have used letters as stimuli; Bunge et al., 2001; D'Esposito et al., 1999; Jonides et al., 2000; Jonides et al., 1998; Mecklinger et al., 2003; Nelson et al., 2003; Thompson-Schill et al., 2002). Moreover, and of greater importance, the data from the Words trials provide a sufficient basis on which to further specify the PFC mechanisms that resolve short-term PI. We conclude by briefly examining Recency and Probe effects during Patterns trials, which warrant interpretative caution given the low accuracy levels during the Patterns task.

*Neural Response to PI during Words Trials*

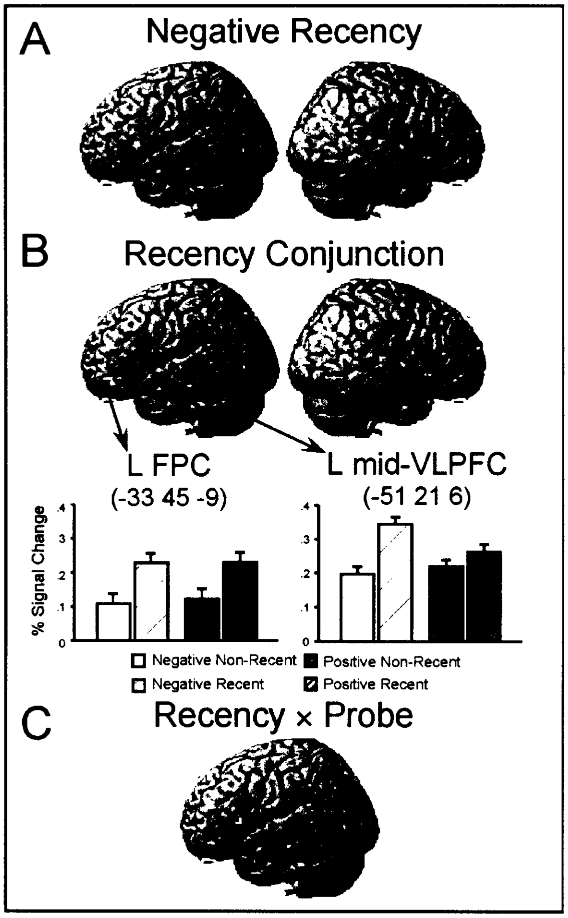

**Figure 3.** The effects of Recency on activation during Words trials, as depicted in surface renderings and graphs of peak percent signal change. (A) The contrast of Negative–Recent to Negative–Non-Recent trials ($p < .005$). (B) Common effects of Positive and Negative Recency (cojoint probability, $p < .001$). (C) Voxel-based Recency X Probe interaction ($p < .005$).

Activation associated with performance in the face of PI was assessed through the contrast of Negative–Recent to Negative–Non-Recent Words trials (Figure 3a). Consistent with prior reports, Negative-Recency produced activation in left mid-VLPFC (~BA 45; MNI coordinates of -51, 21, 6). However, in contrast to earlier reports, reliable activation was also observed in right FPC (~BA 10; 36, 57, -6) and right VLPFC/fronto-operculum (~BA 47/45; 42, 15, 6), a finding that bears on the potential multi-component nature of PI resolution. Given the *a priori* prediction that left PFC regions associated with mnemonic monitoring/integration processes—notably left FPC—may also be engaged by PI, the effect of negative recency was further assessed at a slightly more lenient statistical threshold ($p < .005$). Consistent with this *a priori* expectation, at this threshold negative recency also elicited activation in left FPC/anterior VLPFC (~BA 10; -33, 45, -9 and

~BA 47/10; -45, 42, -3) and in a more superior extent of left VLPFC (~BA 44/45; -42, 21, 24). In addition, the response in right fronto-operculum spread at its most anterior extent into right mid-VLPFC (~BA 45). Collectively, these data suggest that PI and its resolution engage multiple prefrontal cognitive control mechanisms, including bilateral FPC mechanisms that are often engaged during mnemonic monitoring/evaluation and integration processes.

### *Neural Overlap between PI during Words Trials and Episodic Recollection*

As introduced above, one mnemonic context in which FPC is thought to play a central role is in post-retrieval monitoring/decision processes during episodic retrieval. In particular, FPC is thought to contribute to evaluating the products of retrieval attempts, integrating emerging mnemonic information with decision criteria so as to guide action (Dobbins et al., 2002; Rugg & Wilding, 2000). As has been argued elsewhere, FPC activation during episodic recollection is not thought to reflect memory-specialized mechanisms, but rather the recruitment of basic cognitive control processes—monitoring/evaluation, integration, and/or subgoaling—in the service of guiding memory decisions (Buckner & Wheeler, 2001; Christoff et al., 2001; Dobbins et al., 2002; Wagner, 1999). Here, we sought to determine whether PI resolution elicits activation in PFC regions also engaged during episodic recollection. To do so, we explored the anatomic overlap between the presently observed regions showing a Negative–Recent > Negative–Non-Recent pattern and regions engaged during Source Recollection vs. Novelty Detection in a recent episodic retrieval study (Dobbins & Wagner, in press).

In particular, the episodic recollection data derive from a study that identified regions of left PFC—specifically, FPC, mid-VLPFC, and posterior DLPFC—that were more active during recollection of perceptual episodic details (specifically, details about the perceptual size of an object's prior presentation) and during recollection of conceptual episodic details (specifically, details about the semantic task performed during prior object presentation) compared to during assessment of relative stimulus novelty/familiarity (Dobbins & Wagner, in press). This pattern of activation suggests that specific left PFC subregions are generally engaged during episodic recollection regardless of the domain of the to-be-recollected details (i.e., perceptual or conceptual), and thus may subserve basic cognitive control processes recruited during attempts to remember details about the past (see also, Buckner, 2003). Furthermore, Dobbins and Wagner (in press) report strong within-experiment evidence of a domain-general selection process in left mid-VLPFC that may be distinguished from an FPC monitoring mechanism operating in the service of episodic recollection. As such, this study provides a highly appropriate basis of comparison with the current results.

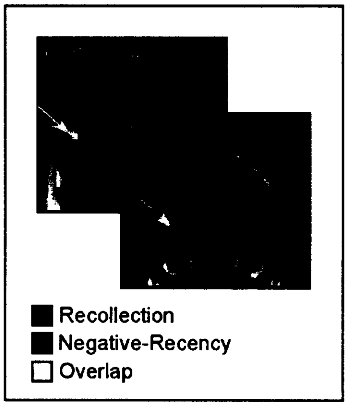

**Figure 4.** Overlap of regions showing greater activation during Negative-Recent vs. Non-Recent trials with regions associated with episodic contextual recollection. Episodic recollection effects were defined as demonstrating greater activation during both perceptual recollection and conceptual recollection compared to a novelty detection task (Dobbins and Wagner, in press).

Comparison of the present left FPC and mid-VLPFC regions elicited by PI resolution during short-term item recognition to the left PFC regions engaged during Recollection was performed through superposition of the domain-general recollection effect of Dobbins and Wagner with our contrast of Negative–Recent > Negative–Non-Recent. This analysis revealed high overlap within left mid-VLPFC and FPC, though the posterior DLPFC region observed during Recollection was not observed to be sensitive to PI in the present experiment (Figure 4). Importantly, this high degree of convergence was not simply a by-product of the data chosen for comparison. Indeed, the focus in left FPC arising from the contrast of Negative–Recent > Negative–Non-Recent in the present study fell in close proximity to findings of left FPC activation reported in a number of previously published studies of episodic retrieval (e.g., Dobbins et al., 2003 [-45 45 –6]; Kahn, Davachi, & Wagner, 2004 [-48 42 –6]; Ranganath et al., 2000 [-53 41 0]; Rugg et al., 1996 [-30 48 –2]; Wagner et al., 1998 [-43 50 4]).

In contrast to this correspondence between the left mid-VLPFC and FPC regions engaged during PI resolution and those associated with recollection of episodic detail, there was remarkably little overlap between the presently observed PI effects and a left anterior VLPFC region (~BA 47) that Dobbins and Wagner (in press) observed to be selectively engaged during controlled retrieval of semantic information. This outcome provides an important control, demonstrating that PI resolution does not overlap with PFC regions engaged simply during any retrieval condition requiring cognitive control, but rather seems relatively specific to left FPC control processes associated with monitoring and integration and left mid-VLPFC processes associated with selecting target representations in the face of interference/competition (or, alternatively, actively inhibiting competing item or response representations).

### *Neural Response to Recency during Words Trials*

Assessment of Recency effects on Positive trials may provide important constraints on theoretical accounts of the mechanisms of PI resolution. In particular, it is critical to assess whether regions that show an increase in response to Negative–Recency are also sensitive to the enhanced familiarity of Positive–Recent probes and whether this sensitivity is reflected in a signal increase or decrease relative to Positive–Non-Recent probes.

Accordingly, we performed a conjunction analysis to determine whether a convergent effect of Recency was present during Negative and Positive trials, and subsequently performed an interaction analysis to determine whether the magnitude of the Recency effect differed across Probe type. The conjunction analysis, targeting the independent effects of Recent > Non-Recent on Negative and on Positive trials (conjoint alpha-level = .001), revealed Recency-induced activation increases in left mid-VLPFC (~BA 45), left FPC (~BA 10), and right VLPFC (~BA 45) (Figure 3b). The Recency × Probe analysis ($p < .005$) revealed a reliable interaction only in left mid-VLPFC — the Recency-induced activation increase in this region was greater on Negative than on Positive trials (Figure 3c, see also 3b), whereas the magnitude of the Recency effects in left FPC and right VLPFC were comparable on Negative and Positive trials (Figure 3b). These results were confirmed using independent $t$-tests contrasting the average beta values within each region of interest against zero. Specifically, the extracted beta values from the contrast of Negative–Recent > Negative–Non-Recent in left mid-VLPFC ($t(16) = 4.5, p < .0005$) and left FPC ($t(16) = 3.6, p < .005$) were reliable, as were those from the contrast of Positive–Recency > Positive–Non-Recent in left mid-VLPFC ($t(16) = 2.4, p < .05$) and left FPC ($t(16) = 3.6, p < .005$).

Two central findings emerged from these analyses. First, despite the limited impact of Recency on behavior during Positive trials and PI during Negative trials, the effect of Recency on left mid-VLPFC, left FPC, and right VLPFC activation was similar on Positive and Negative trials; all regions showed a Recent > Non-Recent effect. Additional analyses revealed that no region showed a Non-Recent > Recent pattern. Second, the effect of Recency in left mid-VLPFC was modulated by Probe type, being greater during Negative than during Positive probes. This interaction is important and may distinguish left mid-VLPFC from the other PFC regions that were affected by Recency but were otherwise insensitive to Probe type. Indeed, the Recency × Probe × Region interaction was reliable when comparing left mid-VLPFC with right VLPFC ($F(1,16) = 6.0$, $p < .05$), though this interaction did not reach significance when comparing left mid-VLPFC with left FPC ($F(1,16) = 2.7$, $p = .12$).

*Patterns Task: Domain Generality of PI Resolution*

Evidence for the domain generality of PI resolution mechanisms would provide an important additional theoretical constraint. To the extent that the mechanisms recruited to resolve PI are domain general, the regions showing a Negative–Recency effect for Words should also show such an effect for the non-verbal Patterns task. However, voxel-based comparison of Negative–Recent to Negative–Non-Recent trials revealed no reliable activation in PFC at a standard ($p<.001$) or a more lenient threshold ($p<.005$) during the Patterns task. Furthermore, though peak signal differences in the left mid-VLPFC region that showed a Negative–Recency effect for Words also showed a qualitative pattern of Recent > Non-Recent for Patterns, this difference was not reliable ($F = 2.5$, $p = .13$). Hence, similar to previous reports (Mecklinger et al., 2003), the present experiment failed to provide evidence in favor of domain general PI resolution in mid-VLPFC. However, strong inferences cannot be drawn from this null finding because accuracy was markedly lower in the Patterns tasks.

## Discussion

The present experiment sought to advance understanding of the nature of PFC mechanisms that resolve PI during short-term item recognition. Four central findings emerged from consideration of PI during the Words task. First, an extended set of PFC regions were sensitive to Negative–Recency, including a left mid-VLPFC region previously associated with PI resolution as well as bilateral FPC. Second, convergence analyses revealed that these regions anatomically overlapped with left PFC regions engaged during domain-general episodic recollection. Third, probe Recency also elicited greater PFC activation during Positive trials. Finally, probe type modulated the magnitude of the Recency effect in left mid-VLPFC, but not in FPC or right VLPFC, suggesting that multiple cognitive control processes may be recruited to resolve PI. We consider each of these findings in turn, and then discuss how these observations provide important constraints on mechanistic accounts of PI resolution.

### *Frontopolar Cortex and Proactive Interference*

The observed sensitivity of FPC to Recency is broadly consistent with studies that have assessed the contributions of cognitive control to working memory and episodic memory. Such studies have reported increased FPC activation during the performance of tasks that involve minimal response conflict or selection, but require the generation of subgoals and the integration of representations deriving from different subgoal stages

(Badre & Wagner, 2004; Braver & Bongiolatti, 2002; Bunge et al., 2003; Koechlin et al., 1999). In the context of episodic retrieval, FPC may support post-retrieval monitoring, which also putatively requires integration of retrieved mnemonic information with decision criteria en route to a response (e.g., Buckner & Wheeler, 2001; Dobbins et al., 2002; Dobbins et al., 2003; Fletcher & Henson, 2001; Kahn et al., 2004; Rugg, Fletcher, Chua, & Dolan, 1999; Rugg & Wilding, 2000; Wagner, 1999).

During short-term item recognition, a similar integration process might be differentially required to guide a decision on how to respond when the probe was a member of the previous target set (i.e., a Recent probe). Specifically, multiple mnemonic signals, including familiarity and information about membership in the current and/or previous target set, must be evaluated with respect to decision criteria, such as with respect to the appropriate temporal context. Consistent with this perspective, the present study provides the first compelling evidence that FPC is engaged in the face of PI during short-term item recognition, extending a previously suggested subthreshold effect by Jonides et al. (1998). Reliable detection of this effect may derive from increased power due to inclusion of more experimental trials (~50% more on average) than previous reports (Bunge et al., 2001; D'Esposito et al., 1999; Jonides et al., 2000; Jonides et al., 1998; Mecklinger et al., 2003; Nelson et al., 2003; Thompson-Schill et al., 2002), thus increasing sensitivity to detect effects in FPC, a region that suffers modest susceptibility-induced signal loss.

To the extent that a similar FPC process supporting monitoring and/or integration is engaged during both episodic retrieval and PI resolution, then one might expect that the regions showing sensitivity to PI during short-term item recognition would overlap with those engaged during episodic recollection. Consistent with this prediction, the left FPC and mid-VLPFC regions observed to be sensitive to Recency converged with those implicated in domain-general retrieval of episodic details in an independent sample of subjects (Figure 4). This convergence analysis provides evidence for the hypothesized commonality in process, moving beyond a qualitative or general "regional" similarity to that of voxel-level overlap.

Of course, this convergence cannot provide irrefutable evidence of reliance on a common process, because, even at the resolution of fMRI, two distinct and independent neural processes may occupy co-local PFC voxels. Such additivity of processes within a single region of FPC is not without precedent in the cognitive control literature (e.g., Badre & Wagner, 2004). Moreover, it is also clear that the PFC regions observed during episodic recollection and PI resolution do not perfectly overlap, particularly with respect to posterior DLPFC. Hence, the convergence analysis suggests that some of the cognitive control processes engaged during episodic recollection—in particular, those subserved by left FPC and mid-VLPFC—may also be engaged in the face of PI.

### *Positive Recency Effects*

Analysis of the effects of Recency revealed increased activation in left mid-VLPFC and FPC on both Negative– and Positive–Recent trials. Interestingly, in the present experiment mild Recency-induced facilitation was evident behaviorally on Positive trials, consistent with quantitative patterns reported in at least two previous studies using this paradigm (D'Esposito et al., 1999; Thompson-Schill et al., 2002). Though the Positive trial effects on activation levels diverge somewhat from these corresponding behavioral effects, such divergence is consistent with prior reports of left

VLPFC increases during congruent trials in the Stroop paradigm (Milham et al., 2001), a task in which behavioral facilitation effects are more substantial than those reported here. Similarly, within the context of a conceptual repetition priming experiment, within-feature behavioral priming (RT speeding) has been associated with reduced left VLPFC activation relative to an unprimed baseline (e.g., Demb et al., 1995; Wagner, Koutstaal, Maril, Schacter, & Buckner, 2000), whereas across-feature behavioral priming can be accompanied by *increased* left VLPFC activation relative to baseline (Thompson-Schill, D'Esposito, & Kan, 1999). Thompson-Schill et al. (1999) hypothesized that this latter activation increase reflects greater demands on selection processes that resolve interference due to the priming of a task-irrelevant feature. Within the present context, the observed activation increases in FPC and mid-VLPFC during Positive–Recent trials are particularly important as they argue strongly against a mechanistic framework that predicts a uniform reduction in control demands due to the convergence of multiple sources of evidence favoring a particular response.

Of further theoretical significance, the pattern of activity in FPC and left mid-VLPFC across Positive and Negative trials is inconsistent with a hypothesis that these regions are globally sensitive to the presence or absence of familiarity. In particular, the magnitude of activation in these regions did not differ between Positive–Non-Recent trials and Negative–Non-Recent trials (see Figure 3b), even though a Positive probe's membership in the current target set ensures that it will be relatively more familiar than a Negative probe.

Analysis of Positive trials also suggested functional differences between FPC and left mid-VLPFC. Activity in FPC increased comparably in response to Recency, irrespective of whether the probe was Positive or Negative. That is, FPC was not sensitive to the congruency between familiarity and a response, as this would have produced a Probe × Recency interaction. By contrast, a Probe × Recency interaction was observed in left mid-VLPFC. This interaction may suggest that left mid-VLPFC is not exclusively sensitive to the history of a given probe (i.e., its presence or absence in the previous trial target set), but is also modulated by the response attribution of the probe. This characteristic may differentiate processing in this region from FPC, and further points to multi-component cognitive control contributions to PI resolution.

Supplemental exploratory correlation analyses further associate the control processes mediated by FPC and left mid-VLPFC with different components of behavior.

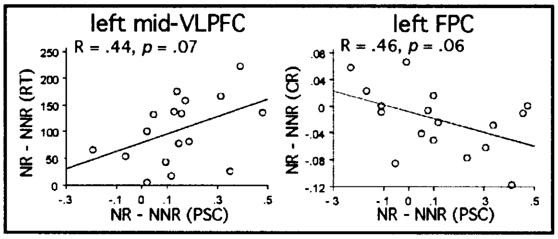

**Figure 5.** Correlations between behavioral indices of PI and PFC activation in response to PI. Individual differences in percent signal change (PSC) during Negative—Recent (NR) relative to Negative–Non-Recent (NNR) Words trials positively correlated with differences in RT slowing due to Recency. Individual differences in left FPC activation in the face of PI were negatively correlated with differences in Recency-induced declines in corrected recognition (CR).

Specifically, we assessed the relation between behavioral indices of PI—expressed as Recency-induced increases in RT and decreases in corrected recognition (computed on Hits – False Alarms) during Negative trials—and Recency-induced changes in left mid-VLPFC and FPC activation (Figure 5). Interestingly, across-subject differences in PI-related activation increases in mid-VLPFC tended to be positively related to differences in RT slowing (R = .44, *p* < .07), whereas PI-related activation increases in left FPC tended to negatively correlate with interference-related declines in response accuracy (R

= .46, $p = .06$). Though strong conclusions are not warranted based on these outcomes alone, these trends are intriguing as they suggest that mid-VLPFC and FPC were associated with different aspects of behavioral performance. This qualitative difference is consistent with distinct control processes that operate on separable components of PI resolution.

One such distinction, suggested by the extant functional literature on these regions, might lie between selection/retrieval mechanisms that overcome interference and so impact RT (correlated with left mid-VLPFC) and post-retrieval monitoring/integration processes related to arriving at a decision for action that might impact recognition outcome (correlated with FPC). It is notable that the direction of correlation also differed between left mid-VLPFC (positive with RT) and FPC (negative with corrected recognition). However, without an estimate of baseline impact of interference on behavior in each subject, it is difficult to interpret the direction of correlation. For example, the positive correlation between interference RT and left VLPFC might reflect increased processing to overcome greater interference, interference that would be even greater were that processing to have not been engaged. Indeed, damage inclusive of left mid-VLPFC produces greatly enhanced behavioral interference (Thompson-Schill et al., 2002), potentially consistent with such a complex relationship. Nevertheless, the qualitative difference in the direction of correlation further points to a process distinction between mid-VLPFC and FPC. In the final section, we consider such mechanistic perspectives in greater detail.

### *Frontal Lobe Mechanisms that Resolve Proactive Interference*

The full complement of observed results provides insight into the processes that resolve PI. Specifically, it is evident that multiple cognitive control processes contribute to performance under conditions of PI, including those mediated by left mid-VLPFC and FPC. Moreover, because Recency has the effect of up-regulating engagement of these processes, even when the Recent probe is a member of the current target set, control demands are not uniformly reduced when multiple sources of evidence converge to guide a response. However, responses to Recency cannot be simply interpreted as a general sensitivity to item familiarity. Finally, functional differences appear to distinguish the control processes mediated by FPC and left mid-VLPFC. Positive– and Negative–Recency effects were comparable in FPC whereas the consequence of Positive–Recency was more modest than that of Negative–Recency in left mid-VLPFC, and activity in FPC and mid-VLPFC tended to differentially correlate with separate components of behavior. Constrained by these findings, we will attempt to specify two classes of mechanistic perspectives on PI resolution.

Prior mechanistic accounts of PI resolution assign a response- or attribute-selection function to left mid-VLPFC that is differentially necessary when Recency induces conflict (D'Esposito et al., 1999; Jonides, Badre, Curtis, Thompson-Schill, & Smith, 2002; Jonides et al., 1998; Mecklinger et al., 2003). Such models, which we generally term *familiarity-inhibition* models, at least implicitly suggest that because all positive probes have just been encountered, and so are familiar, familiarity itself may come to be associated with a Positive response. During short-term item recognition, a process of matching a presented probe to the currently maintained target set is required on every trial. Interference during the processing of Negative–Recent probes may therefore arise from conflict between the result from this matching process ("Negative") and the learned tendency to respond "Positive" to probes that are familiar. When activation of the competing "Positive" response makes selection of the "Negative" response more difficult, a response selection processes is putatively required to select the "Negative" response. Alternatively, the attribution of familiarity itself might interfere with selecting the appropriate response, and PI resolution proceeds, in this case, through inhibition of the influence of familiarity on assigning a probe to a target context (as opposed to directly selecting the target response).

Past formulations of this familiarity-inhibition mechanism have focused on delineating this single form of control, which depends specifically on left mid-VLPFC mechanisms. However, in isolation, such a selection mechanism would need to operate without knowledge, *a priori*, of which response to select, or which attribute (familiarity or set membership) to favor. For example, there might be task contexts in which familiarity should govern a response and so selection should proceed in favor of the attribution of familiarity. By way of extending these models, a plausible variant might include an additional process—partially dependent on FPC—that monitors the results of the biasing process in the context of task-specific decision criteria, such as prioritizing set membership or temporal context. Under such a circumstance, facilitation on positive trials is not necessarily predicted and FPC activation is anticipated. However, it is not entirely clear how such a model would account for the increased activation on Positive–Recent trials, while also distinguishing between interference from familiarity due to membership in the previous set but not from familiarity due to membership in the current set. Hence, though still plausible, the modified familiarity-inhibition model appears challenged by key aspects of the present findings.

An alternative class of models, which we term *context-retrieval* models, considers PI resolution with respect to those control processes also engaged during retrieval from episodic memory. Context-retrieval proposes that PI resolution proceeds by selecting relevant episodic details in order to assign a probe to a task-relevant temporal context. Hence, the familiarity-inhibition and context-retrieval models both propose similar cognitive control processes to overcome PI, namely representational selection and monitoring, and also both recognize that the critical PI manipulation may arise from a long-term memory signal (e.g., familiarity) cued by a recently encountered item that has dropped from active memory. However, context-retrieval and familiarity-inhibition differ fundamentally with respect to the nature of the representations that give rise to PI, and thus those representations that are selected and monitored to ultimately guide action. Specifically, whereas familiarity-inhibition assigns PI to conflict between stimulus attributes or between mappings that give rise to responses, context-retrieval posits interference as competition amongst specific episodic details that can assign an item to a particular temporal context.

From one such perspective, PI resolution depends on the retrieval and evaluation of context information. Context information refers to any retrieved detail that can be used to assign a probe to the context in which it was encoded. Examples of such information may include the associated targets in a probe's memory set, its spatial location, or a temporal tag. From this perspective, multiple PFC subregions contribute to context retrieval and the integration of retrieved information with the decision criteria, including mechanisms that retrieve target contextual information (e.g., left mid-VLPFC) and mechanisms that monitor recovered information in the service of arriving at a decision for action (e.g., FPC). Because not all familiar probes are members of the current target set, when presented with a familiar probe, context retrieval processes are engaged to assign a probe to the context in which it was recently encountered. Importantly, when a probe can be associated with more than one source, interference may arise, consistent with classical accounts of PI in episodic memory. Thus, retrieval demands vary depending on the extent that one retrieved piece of contextual information needs to be favored over another interfering contextual representation.

Critically, this mechanistic hypothesis may provide a parsimonious account of the present results. On Negative–Recent trials, interference emerges because the familiar probe induces activation of associated details from the previous (*trial n-1*) context in which the probe had appeared. This competes with retrieval or selection of details from the current (*trial n*) context. Consequently, demands on selection and monitoring processes increase, as reflected in left mid-VLPFC and FPC, respectively. On Positive–Recent trials, the association of the probe with the previous trial context also results in competition during context retrieval, reflected in a Recency effect in left mid-VLPFC, and an increase in monitoring demands, reflected in a Recency effect in FPC. However, Positive trials were also members of the currently maintained target set. This may make the relevant context information prepotent and thus easier to retrieve. Hence, relative to Negative-Recent trials, selection demands on Positive-Recent trials are more modest. Accordingly, this model accounts for the Probe × Recency interaction observed in left mid-VLPFC. Though further empirical work is needed to definitively distinguish between this model and the familiarity-inhibition account, the present results lend some support to the context-retrieval hypothesis.

## Conclusions

PI places considerable constraints on cognition, and thus its resolution is critical for execution of goal-relevant behavior. This includes the demand to flexibly shift tasks as stimuli encountered during performance of a given task may cue retrieval of competing task sets and so produce PI. The evidence reported here indicates that PI during short-term item recognition is resolved by multiple cognitive control mechanisms, including those mediated by left mid-VLPFC and FPC. Additional work may further distinguish and constrain the mechanistic perspectives on PI resolution described here. This experiment has important and broad implications for the study of memory and cognitive control, as it suggests that the PFC mechanisms responsible for PI resolution during short-term item recognition may be common with those required to resist the more classic effects of PI associated with forgetting from episodic memory. Indeed, a neurally-specified mechanism of PI resolution may contribute fundamentally to our knowledge of the manner and processes through which mnemonic obstacles to successful memory performance and action are overcome. Finally, to the extent that PI plays a role in task switching, a similar mid-VLPFC mechanism to that studied here may also be critical to overcome interference from a previously performed task. Furthermore, the theoretical constraints on a PI resolution mechanism motivated by the present results may provide an important context for theorizing regarding the PFC mechanisms engaged during task switching.

## References

Badre, D., & Wagner, A. D. (2004). Selection, integration, and conflict monitoring: Assessing the nature and generality of prefrontal cognitive control mechanisms. *Neuron, 41*(3), 473-487.

Braver, T. S., Barch, D. M., Kelley, W. M., Buckner, R. L., Cohen, N. J., Miezin, F. M., et al. (2001). Direct comparison of prefrontal cortex regions engaged by working and long-term memory tasks. *Neuroimage, 14*, 48-59.

Braver, T. S., & Bongiolatti, S. R. (2002). The role of frontopolar cortex in subgoal processing during working memory. *NeuroImage, 15*, 523-536.

Braver, T. S., Reynolds, J. R., & Donaldson, D. I. (2003). Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron, 39*(4), 713-726.

Brown, J. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology, 10*, 12-21.

Buckner, R. L. (2003). Functional-anatomic correlates of control processes in memory. *Journal of Neuroscience, 23*(10), 3999-4004.

Buckner, R. L., Petersen, S. E., Ojemann, J. G., Miezin, F. M., Squire, L. R., & Raichle, M. E. (1995). Functional anatomy studies of explicit and implicit memory retrieval tasks. *Journal of Neuroscience, 15*, 12-29.

Buckner, R. L., & Wheeler, M. E. (2001). The cognitive neuroscience of remembering. *Nat Rev Neurosci, 2*(9), 624-634.

Bunge, S. A., Badre, D., Wendelken, C., & Wagner, A. D. (in press). Analogical reasoning and prefrontal cortex: Evidence for separable retrieval and integration mechanisms. *Cerebral Cortex.*

Bunge, S. A., 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*, 3419-3428.

Bunge, S. A., Ochsner, K. N., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (2001). Prefrontal regions involved in keeping information in and out of mind. *Brain, 124*(Pt 10), 2074-2086.

Cabeza, R., Dolcos, F., Graham, R., & Nyberg, L. (2002). Similarities and differences in the neural correlates of episodic memory retrieval and working memory. *Neuroimage, 16*(2), 317-330.

Cansino, S., Maquet, P., Dolan, R. J., & Rugg, M. D. (2002). Brain activity underlying encoding and retrieval of source memory. *Cereb Cortex, 12*(10), 1048-1056.

Carter, C. S., Mintun, M., & Cohen, J. D. (1995). Interference and facilitation effects during selective attention: an H215O PET study of Stroop task performance. *Neuroimage, 2*(4), 264-272.

Christoff, K., Prabhakaran, V., Dorfman, J., Zhao, Z., Kroger, J. K., Holyoak, K. J., et al. (2001). Rostrolateral prefrontal cortex involvement in relational integration during reasoning. *Neuroimage, 14*(5), 1136-1149.

Christoff, K., Ream, J. M., Geddes, L. P., & Gabrieli, J. D. E. (2003). Evaluating self-generated information: anterior prefrontal contributions to human cognition. *117, 6*(1161-1168).

Cocosco, C. A., Kollokian, V., Kwan, R. K. S., & Evans, A. C. (1997). Brainweb: Online interface to a 3D MRI simulated brain database. *NeuroImage, 5*, 425.

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

Dale, A. M. (1999). Optimal experimental design for event-related fMRI. *Hum Brain Mapp, 8*(2-3), 109-114.

Demb, J. B., Desmond, J. E., Wagner, A. D., Vaidya, C. J., Glover, G. H., & Gabrieli, J. D. E. (1995). Semantic encoding and retrieval in the left inferior prefrontal cortex: a functional MRI study of task difficulty and process specificity. *Journal of Neuroscience, 15*, 5870-5878.

Dobbins, I. G., Foley, H., Schacter, D. L., & Wagner, A. D. (2002). Executive control during episodic retrieval: multiple prefrontal processes subserve source memory. *Neuron, 35*(5), 989-996.

Dobbins, I. G., Rice, H. J., Wagner, A. D., & Schacter, D. L. (2003). Memory orientation and success: separable neurocognitive components underlying episodic recognition. *Neuropsychologia, 41*(3), 318-333.

Dobbins, I. G., & Wagner, A. D. (in press). Domain-general and domain-sensitive prefrontal mechanisms for recollecting evens and detecting novelty. *Cerbral Cortex.*

Fletcher, P. C., & Henson, R. N. (2001). Frontal lobes and human memory: insights from functional neuroimaging. *Brain, 124*(Pt 5), 849-881.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 193-225). Orlando: Academic Press.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source Monitoring. *Psychological Bulletin, 114*, 3-28.

Johnson, M. K., & Raye, C. L. (1981). Reality Monitoring. *Psychological Review, 88*, 67-85.

Johnson, M. K., Raye, C. L., Mitchell, K. J., Greene, E. J., & Anderson, A. W. (2003). FMRI evidence for an organization of prefrontal cortex by both type of process and type of information. *Cerbral Cortex, 13*(3), 265-273.

Jonides, J., Badre, D., Curtis, C., Thompson-Schill, S. L., & Smith, E. E. (2002). Mechanisms of Conflict Resolution in Prefrontal Cortex. In D. T. Stuss & R. T. Knight (Eds.), *Principles of Frontal Lobe Function* (pp. 233-245). Oxford: Oxford University Press.

Jonides, J., Marshuetz, C., Smith, E. E., Reuter-Lorenz, P. A., Koeppe, R. A., & Hartley, A. (2000). Age differences in behavior and PET activation reveal differences in interference resolution in verbal working memory. *J Cogn Neurosci, 12*(1), 188-196.

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

Kahn, I., Davachi, L., & Wagner, A. D. (2004). Functional-neuroanatomic correlates of recollection: implications for models of recognition memory. *J Neurosci, 24*(17), 4172-4180.

Keppel, G., & Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning and Verbal Behavior, 1*, 153-161.

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

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science, 302*(1181-1185).

McGeoch, J. A. (1942). *The Psychology of Human Learning*. New York: Longmans, Green.

Mecklinger, A., Weber, K., Gunter, T. C., & Engle, R. W. (2003). Dissociable brain mechanisms for inhibitory control: effects of interference content and working memory capacity. *Brain Res Cogn Brain Res, 18*(1), 26-38.

Milham, M. P., Banich, M. T., Webb, A., Barad, V., Cohen, N. J., Wszalek, T., et al. (2001). The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on the nature of conflict. *Cognitive Brain Research, 12*, 467-473.

Milham, M. P., Erickson, K. I., Banich, M. T., Kramer, A. F., Webb, A., Wszalek, T., et al. (2002). Attentional control in the aging brain: Insights from an fMRI study of the Stroop task. *Brain and Cognition, 49*, 167-202.

Monsell, S. (1978). Recency, immediate recognition memory, and reaction time. *Cognitive Psychology, 10*, 465-501.

Nelson, J. K., Reuter-Lorenz, P. A., Sylvester, C.-Y. C., Jonides, J., & Smith, E. E. (2003). Dissociable neural mechanisms underlying response-based and familiarity-based conflict in working memory. *Proc Natl Acad Sci U S A, 100*(19), 11171-11175.

Nichols, T., Brett, M., Andersson, J., Wager, T., & Poline, J.-B. (in press). Valid conjunction inference with the minimum statistic. *NeuroImage*.

Nolde, S. F., Johnson, M. K., & D'Esposito, M. (1998). Left prefrontal activation during episodic remembering: an event-related fMRI study. *Neuroreport, 9*(15), 3509-3514.

Nyberg, L., Marklund, P., Persson, J., Cabeza, R., Forkstam, C., Petersson, K. M., et al. (2003). Common prefrontal activations during working memory, episodic memory, and semantic memory. *Neuropsychologia, 41*(3), 371-377.

Nyberg, L., McIntosh, A. R., Cabeza, R., Habib, R., Houle, S., & Tulving, E. (1996). General and specific brain regions involved in encoding and retrieval of events: what, where, and when. *Proc Natl Acad Sci U S A, 93*(20), 11280-11285.

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual items. *Journal of Experimental Psychology, 58*, 193-198.

Ranganath, C., Johnson, M. K., & D'Esposito, M. (2000). Left anterior prefrontal activation increases with demands to recall specific perceptual information. *J Neurosci, 20*(22), RC108.

Rugg, M. D., Fletcher, P. C., Chua, P. M., & Dolan, R. J. (1999). The role of the prefrontal cortex in recognition memory and memory for source: an fMRI study. *Neuroimage, 10*(5), 520-529.

Rugg, M. D., Fletcher, P. C., Frith, C. D., Frackowiak, R. S. J., & Dolan, R. J. (1996). Differential activation of the prefrontal cortex in successful and unsuccessful memory retrieval. *Brain*(119), 2073-2083.

Rugg, M. D., & Wilding, E. L. (2000). Retrieval processing and episodic memory. *Trends in Cognitive Science, 4*, 108-115.

Sakai, K., & Passingham, R. E. (2003). Prefrontal interactions reflect future task operations. *Nat Neurosci, 6*(1), 75-81.

Shimamura, A. P. (1995). Memory and frontal lobe function. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences* (pp. 803-813). Cambridge, MA: MIT Press.

Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review, 99*, 195-231.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643-662.

Talairach, J., & Tournoux, P. (1988). *A Co-Planar Stereotactic Atlas of the Human Brain.* Stuttgart: Thieme.

Thompson-Schill, S. L., D'Esposito, M., & Kan, I. P. (1999). Effects of repetition and competition on activity in left prefrontal cortex during word generation. *Neuron, 23*, 513-522.

Thompson-Schill, S. L., Jonides, J., Marshuetz, C., Smith, E. E., D'Esposito, M., Kan, I. P., et al. (2002). Effects of frontal lobe damage on interference effects in working memory. *Cogn Affect Behav Neurosci, 2*(2), 109-120.

Tulving, E., Kapur, S., Craik, F. I. M., Moscovitch, M., & Houle, S. (1994). Hemispheric encoding-retrieval asymmetry in episodic memory: Positron emission tomography findings. *Proceedings of the National Academy of the Sciences, USA, 91*, 2016-2020.

Wagner, A. D. (1999). Working memory contributions to human learning and remembering. *Neuron, 22*(1), 19-22.

Wagner, A. D., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1998). Prefrontal cortex and recognition memory: fMRI evidence for context-dependent retrieval processes. *Brain, 121*, 1985-2002.

Wagner, A. D., Koutstaal, W., Maril, A., Schacter, D. L., & Buckner, R. L. (2000). Task-specific repetition priming in left inferior prefrontal cortex. *Cerebral Cortex, 10*, 1176-1184.

# Chapter 3

## Dissociable Controlled Retrieval and Generalized Selection Mechanisms in Ventrolateral Prefrontal Cortex

The fMRI experiment described in Chapter 2 associated activation in left mid-VLPFC with resolution of proactive interference analogous to that proposed to occur during task switching. Furthermore, neuroimaging studies of task switching have consistently reported increased activation in left VLPFC on switch versus repeat trials (e.g. Brass & von Cramon, 2004; Dove, Pollmann, Schubert, Wiggins, & von Cramon, 2000; see Chapter 1). However suggestive these results are, the mechanisms of control engaged generally by left VLPFC, and so its specific contribution to task switching, remain unclear. Indeed, even recognizing the potential role of left VLPFC in broadly controlling memory does not distinguish between competing perspectives. For instance, the TSR hypothesis might propose that a control process in left VLPFC acts endogenously to guide retrieval of representations in the upcoming task set. By contrast, from the TSI/TSP perspectives, left VLPFC might resolve interference coming up from memory by selecting among relevant retrieved representations. Indeed, both mnemonic control processes have been proposed for left VLPFC. Hence, to gain further leverage on the role of left VLPFC during task switching, it will be important to determine the extent to which left VLPFC control processes may be characterized as guiding retrieval and/or resolving interference from retrieved representations more generally.

How does VLPFC control mnemonic processing? Alternative models have proposed that VLPFC guides top-down (controlled) retrieval of knowledge from long-

term stores or selects goal-relevant products of retrieval from amongst competitors. To date, a paucity of evidence supports a retrieval/selection distinction, raising the possibility that these models reduce to a common mechanism. Here, four manipulations varied semantic control demands during fMRI: judgment specificity, cue–target associative strength, competitor dominance, and number of competitors. Principal components analysis revealed novel evidence for a meta-factor that accounted for common behavioral variance across these manipulations, and for functional variance in left mid-VLPFC activation. These data provide compelling evidence for a generalized control process that selects relevant knowledge under competitive conditions. By contrast, left anterior VLPFC and middle temporal cortex were sensitive to cue–target associative strength, but not competition, consistent with a control process that retrieves knowledge stored in lateral temporal cortex. Distinct PFC mechanisms mediate top-down retrieval and post-retrieval selection.

## Introduction

Over a lifetime, humans accumulate knowledge about the world, including general facts, concepts, and word meanings. Making gainful use of this knowledge to comprehend stimuli and inform action in a variable environment requires a system for retrieving and selecting stored information as goals dictate (Miller & Cohen, 2001; Shimamura, 1995). Substantial evidence indicates that left ventrolateral prefrontal cortex (VLPFC) is critical for the performance of tasks that demand access to and evaluation of semantic knowledge (Demb et al., 1995; Devlin, Matthews, & Rushworth, 2003; Kapur, Rose, Liddle, Zipursky, & et al., 1994; Noppeney & Price, 2004; Petersen, Fox, Posner, Mintun, & Raichle, 1988; Poldrack et al., 1999; Sohn, Goode, Stenger, Carter, & Anderson, 2003; Wagner, Paré-Blagoev, Clark, & Poldrack, 2001; Zhang et al., 2004), though the functional character and topographic organization of processing within left VLPFC remains highly controversial (Badre & Wagner, 2002; Dobbins & Wagner, in press; Gold, Balota, Kirchhoff, & Buckner, 2005; Moss et al., in press; Thompson-Schill, 2003). At the heart of the debate is whether left VLPFC mediates (a) the post-retrieval selection of goal-relevant knowledge over irrelevant competitors (Fletcher, Shallice, & Dolan, 2000; Moss et al., in press; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997), (b) the top-down activation (controlled retrieval) of semantic knowledge under situations in which bottom-up retrieval mechanisms fail to recover goal-relevant information (Bunge, Wendelken, Badre, & Wagner, 2005; Wagner et al., 2001), or (c) both selection and retrieval, as these putatively distinct processes may reduce to a single, shared mechanism (Badre & Wagner, 2002). Here we report novel behavioral and fMRI evidence supporting the existence of a generalized selection mechanism that accounts for behavioral variance under a variety of semantic processing contexts and that accounts for functional variance in left mid-VLPFC activation. We further report that this generalized selection mechanism is functionally and neuroanatomically distinct from a controlled retrieval process that depends on left anterior VLPFC and appears to activate stored semantic knowledge in left temporal cortex. As such, these data provide novel evidence for a mechanistic distinction between selection and retrieval, with selection operating on the products of bottom-up and top-down retrieval.

According to the selection hypothesis, left VLPFC control mechanisms are critical when a subset of knowledge that is task-relevant must be selected over a competing subset of irrelevant knowledge (Fletcher et al., 2000; Moss et al., in press; Thompson-Schill et al., 1997). Hence, selection demands can be manipulated by requiring subjects to direct attention to a subset of cue-related knowledge. For example, when the similarity between stimuli must be judged along a specific semantic dimension (e.g., color or form), other semantic features of the stimuli are task-irrelevant and must be selected against in favor of the relevant feature. Left VLPFC activation is greater during performance of such feature-based judgments relative to global similarity judgments, for which selection demands are argued to be minimal because all features are relevant (Thompson-Schill et al., 1997). Left VLPFC activation also increases during picture naming when competing knowledge is primed—and thus competition is enhanced—and this is the case even when task demands putatively require minimal controlled retrieval (Moss et al., in press). Finally, activation increases in left VLPFC accompany other circumstances in which a primed feature of a stimulus becomes irrelevant upon repetition, thus increasing selection demands during repeated stimulus processing

(Fletcher et al., 2000; Thompson-Schill, D'Esposito, & Kan, 1999). This pattern of PFC activation contrasts with across-feature priming reductions in left temporal cortices—structures thought to store long-term semantic knowledge and to mediate bottom-up retrieval (Thompson-Schill et al., 1999). These observations motivate the hypothesis that left VLPFC supports a post-retrieval selection mechanism that operates on the products of bottom-up retrieval processes, with selection demands increasing when multiple competing representations have been retrieved and when task-irrelevant representations are pre-potent.

Alternatively, left VLPFC mechanisms have been hypothesized to directly support the top-down (controlled) retrieval of knowledge when bottom-up (automatic) processes are insufficient to retrieve task-relevant knowledge (Wagner et al., 2001; Badre and Wagner, 2002). Controlled retrieval demands can be varied by manipulating the extent to which a cue is effective in eliciting retrieval of task-relevant information. For instance, greater left VLPFC activation is observed when the pre-experimental association between the retrieval cue and target knowledge is relatively weak compared to when a strong association exists (Bunge et al., 2005; Wagner et al., 2001). This is the case even within the context of a global-relatedness task in which selection demands may be minimal (Thompson-Schill et al., 1997). Importantly, according to the retrieval perspective, top-down inputs from VLPFC trigger the recovery of long-term knowledge, and thus should have a correlated activation increase in left temporal regions that store semantic knowledge. It is important to note, however, that a manipulation of associative strength may also result in increased selection demands, as weak activation of relevant information may make this knowledge more susceptible to interference, a case similar to an underdetermined response (Thompson-Schill et al., 2005).

Given these competing models of VLPFC function, a critical challenge for theorists of cognitive control is to specify the relation between selection and controlled retrieval. One possibility is that a common process biases retrieval under any circumstance in which relevant knowledge does not come to mind automatically, either due to poor cue support (e.g., weak cue–target associative strength) or to competition from automatically retrieved, irrelevant competitors (Badre & Wagner, 2002). Alternatively, controlled retrieval and selection may be mechanistically and anatomically distinct processes mediated by left VLPFC (Dobbins & Wagner, in press; Martin & Chao, 2001), with the former guiding retrieval of knowledge stored in temporal cortex and the latter operating on the products of retrieval to select relevant representations from amongst competitors. This latter possibility receives indirect support, as studies putatively varying selection demands have typically identified activation in left mid-VLPFC (~Brodmann areas |BA| 45/44), whereas those putatively varying controlled retrieval have localized activation in a more anterior and ventral region of left VLPFC (~BA 47). However, at present, direct evidence for a functional anatomic dissociation between selection and retrieval is lacking. Moreover, a clear mechanistic distinction between these two processes has not been articulated nor empirically supported. These limits partly stem from the fact that, to date, no study has directly manipulated both selection and controlled retrieval demands, and because of the exclusive reliance upon task analyses to support past inferences about the processes correlated with VLPFC activation.

The present functional MRI (fMRI) study directly examined the functional and neuroanatomic relation between selection and controlled retrieval, combining four manipulations of control demands across two experiments (Fig. 1). In both experiments, Judgment Specificity (Fig. 1A) varied whether subjects selected a target based on its global relatedness to a cue (*Related*; low selection) or its similarity to a cue along a specific dimension (*Feature*; high selection). Within the Related task, Associative Strength (Fig. 1B) varied whether the correct target was a *Strong* (low controlled retrieval; potentially low selection) or *Weak* (high controlled retrieval; potentially higher selection) associate of the cue. During Exp. 2, Number of Targets varied whether the correct target was selected from amongst *Two* or *Four* alternatives, providing an

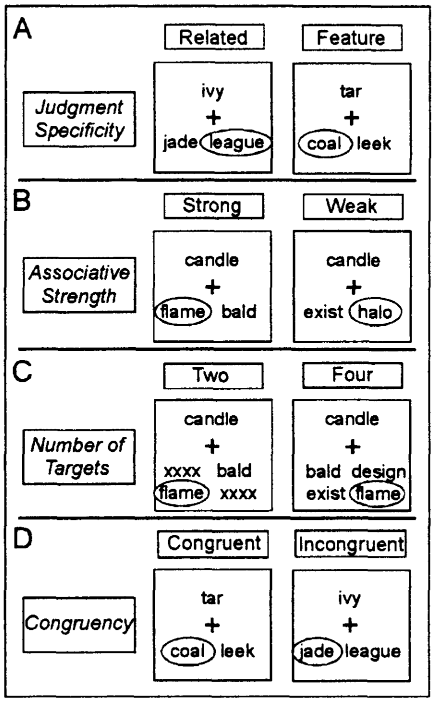

**Figure 1.** Task schematic depicting four manipulations of control at retrieval. On all trials, subjects selected a target (below fixation) based on its relation to the cue (above fixation). (A) Judgment Specificity was manipulated by either requiring selection of the target most globally Related to the cue (left) or most similar to the cue along a specific semantic Feature (right), such as color in this example. (B) Within Related blocks, Associative Strength manipulated whether the correct target was a Strong (left) or Weak (right) associate of the cue. (C) In Exp. 2, the Number of Targets during Related blocks varied between Two (left) or Four (right). (D) Within the Feature task, a trial was Congruent (left) if the correct target was also a pre-experimental associate of the cue and Incongruent (right) if the correct target was not the pre-experimental associate.

additional manipulation of retrieval (be it automatic or controlled) and selection demands (Fig. 1C). Finally, during the Feature task, one of the targets was a normative associate of the cue (e.g., TAR → COAL), and Congruency varied whether this associate was the target most similar to the cue along the relevant dimension (*Congruent*) or was the competing distractor (*Incongruent*) (Fig. 1D). Selection demands were greater during Incongruent trials as information retrieved automatically due to the associative linkage between the cue and distractor was irrelevant, yielding greater competition.

To assess the possible contribution of a common control process across these semantic processing contexts, a principal components analysis (PCA) of behavior was performed to extract a meta-variable that accounted for common behavioral variance across the manipulations. This meta-variable then served as a covariate during fMRI analysis to examine whether it accounted for functional variance within VLPFC.

## Methods

### *Participants*

Twenty-two right-handed, native English speakers (13 female; ages 18-25 yrs) were enrolled in Exp. 1 and an independent sample of 11 right-handed, native English speakers (4 female; ages 18-30 yrs) were enrolled in Exp. 2. Data from two additional subjects recruited for Exp. 1 were excluded due to significant artifacts. All participants received $50 remuneration for participation. Informed consent was obtained in a manner approved by the Human Subjects Committee of the Massachusetts General Hospital and the Committee on the Use of Humans as Experimental Subjects at MIT.

*Design and Logic*

Event-related trials were separated in time by jittered (0-8 s) null fixation periods and were grouped into task blocks. Blocks began with a baseline fixation period (12 s and 16 s for Exps. 1 and 2, respectively), followed by a 4-s instruction cue indicating the task (Feature or Related) to be performed for that block. On each trial, a cue word and a set of target words were presented for 3 s (Fig. 1). Subjects chose one of the targets based on its semantic relationship with the cue and indicated their response on a keypad positioned under their left hand. Subjects were given 4 s to respond (inclusive of the 3-s cue-target set presentation). When the instruction cue was "RELATED", subjects were to select the target that was most globally related to the cue. Alternatively, if the instruction specified a semantic feature (e.g., "COLOR", "SHAPE", "SIZE", or "TEXTURE"), subjects were to select the target most similar to the cue along this dimension. This design permitted manipulation of Judgment Specificity (Feature vs. Related), cue-target Associative Strength (Strong vs. Weak), Number of Targets (Two vs. Four), and Congruency (Congruent vs. Incongruent) during semantic processing (Fig. 1). The order of experimental and fixation events within a block was determined by optimizing the efficiency of the design matrix so as to permit event-related analyses (Dale, 1999); efficiency was equated across Related and Feature blocks.

Exp. 1 was designed to factorially combine control demands, crossing the Associative Strength, Judgment Specificity, and Congruency manipulations within-subject. Across four fMRI scan runs, subjects encountered 240 trials divided equally among the four Associative Strength × Judgment Specificity condition crossings. Furthermore, of the 120 Feature trials, half were Congruent and half were Incongruent

(Fig. 1). Each scan contained four experimental blocks, two Related and two Feature, counterbalanced in an ABBA/BAAB fashion.

Exp. 2 was designed to maximize sensitivity and power of the control manipulations while still permitting within-subject analysis. This goal was achieved by isolating control manipulations into separate processing epochs within a single scan session. During an initial epoch, subjects performed the Related task alone, with Associative Strength and the Number of Targets being manipulated (Fig. 1A and B). In a second epoch, subjects alternated between Feature and Relatedness judgments, as in Exp. 1. However, unlike in Exp. 1, only Judgment Specificity and Congruency manipulations were included. Otherwise, trial events in this epoch unfolded as with Exp. 1. Each of the epochs consisted of two fMRI scan runs. During epoch one, subjects encountered 288 Related trials (Fig. 1A). In the second epoch, subjects performed 80 Related and 80 Feature trials grouped into 8 Related and 8 Feature blocks counterbalanced in an ABBA/BAAB fashion. These blocks were divided equally and counterbalanced across the two scan runs. Furthermore, subjects encountered 40 Congruent and 40 Incongruent trials mixed across the Feature blocks.

### *Stimuli*

Stimuli for all experiments were chosen from single-response free association norms (Moss & Older, 1996; Postman & Keppel, 1970) and were equated for word length and for normative frequency of use (Kucera & Francis, 1967) across experimental conditions. For each of 240 cues in Exp. 1, one strongly associated, one weakly associated, and one unassociated target were chosen. The mean normative probability that the item was generated as the associate of the cue differed across Strong (.11) and Weak (.02) targets, yielding a pre-experimental associative strength ratio of approximately 5:1 for Strong:Weak trials. This ratio was markedly lower than the 22:1 ratio adopted in epoch one of Exp. 2 (see below) due to the additional counterbalancing constraints of the Exp. 1 factorial design. Unassociated targets were determined based on their absence from a cue's normative list of associates.

Stimuli for epoch one of Exp. 2 were taken directly from a prior study, and thus details of stimulus selection and counterbalancing have been described previously (Wagner et al., 2001). The mean normative probability that an item was generated as the associate of the cue differed substantially between Strong (.22) and Weak (.01) targets. For each of the 160 cues in epoch two of Exp. 2, one associated and one unassociated target were selected. The mean normative probability of item generation for the associated target (.19) was comparable to Strong trials of epoch one. Again, unassociated targets were determined based on their absence from a cue's normative list of associates.

### *fMRI Procedures*

Whole-brain imaging for both experiments was performed on a 3T Siemens MRI system (Exp 1: 3T Allegra MRI system; Exp. 2: 3T Trio MRI system). Functional data were acquired using a gradient-echo echo-planar pulse sequence (Exp. 1: TR=2 sec, TE=40 msec, 21 axial slices, 3.125 x 3.125 x 5 mm, 1 mm inter-slice gap, 208 volume acquisitions per run; Exp. 2: TR=2 sec, TE=30 msec, 20 axial slices, 3.125 x 3.125 x 5 mm, 1 mm inter-slice gap, 408/284 volume acquisitions per epoch1/epoch2 run). High-resolution T1-weighted (MP-RAGE) anatomical images were collected for anatomical visualization. Head motion was restricted using firm padding that surrounded the head. Visual stimuli were projected onto a screen, and were viewed through a mirror attached to the head coil.

### *Principal Components Analysis*

PCA analysis of the behavior was performed using standard procedures (Harris, 1967). Differences in RT and errors for Associative Strength, Judgment Specificity, and Congruency were computed and standardized within experimental group for inclusion in PCA analysis. The six eigenvalues describing the variance-covariance matrix of these six scores were then extracted. Factors with eigenvalues greater than 1 were selected for additional analysis. Selected factors underwent oblique rotation using the Varimax algorithm. Regression estimate factor scores for inclusion in fMRI analysis were derived for each subject based on the oblique factor solution (Harris, 1967). Principal components analysis was performed in StatView 5.0.1 (SAS Institute).

### *fMRI Data Analysis*

Data were preprocessed using SPM99 (Wellcome Dept. of Cognitive Neurology, London). Functional images were corrected for differences in slice acquisition timing by resampling all slices in time to match the first slice, followed by motion correction across all runs (using sinc interpolation). Structural and functional data were spatially normalized to an EPI template based on the MNI stereotactic space (Cocosco, Kollokian, Kwan, & Evans, 1997) using a 12-parameter affine transformation along with a nonlinear transformation using cosine basis functions. Images were resampled into 3-mm cubic voxels and then spatially smoothed with an 8-mm FWHM isotropic Gaussian kernel.

Statistical models were constructed using SPM99 under the assumptions of the general linear model. Exp. 1 and the second epoch of Exp. 2 used a mixed fMRI design, such that Judgment Specificity was manipulated across blocks and Associative Strength and Congruency were manipulated in an event-related manner within blocks (Donaldson, Petersen, Ollinger, & Buckner, 2001). Because event and block regressors were correlated in these instances, all conditions were solely modeled as events by constructing regressors for each cell of the design (i.e., any effect of task 'state' was not separately assessed). Similarly, the first epoch of Exp. 2 was modeled in a standard event-related manner. Correct and incorrect trials were modeled separately; all statistical contrasts were restricted to correct trials.

Effects were estimated using a subject-specific fixed-effects model, with session-specific effects and low-frequency signal components treated as confounds. Linear contrasts were used to obtain subject-specific estimates for each effect. These estimates were entered into a second-level analysis treating subjects as a random effect, using a one-sample t-test against a contrast value of zero at each voxel. Correlations of individual effects of control with factor scores were estimated using a multiple regression that included the factor scores as independent measures and the subject-specific estimate for each control contrast as the dependent measure at each voxel.

Voxel-based group effects were considered reliable to the extent that they consisted of at least 5 contiguous voxels that exceeded an uncorrected threshold of $p < .001$. Moreover, maxima reported in left VLPFC survived correction for multiple comparisons (corrected $p < .05$) over the search volume using Gaussian random field theory (Friston et al., 1995). Where effects in left VLPFC did not pass the corrected threshold, uncorrected results are reported to the extent that they constitute replications of findings from an independent data set. The volume used for correction included gray matter within left VLPFC and was generated in an unbiased manner based on the intersection of the automated anatomical labeling (AAL) regions (Tzourio-Mazoyer et al., 2002) that comprise the entire inferior frontal gyrus (AAL regions: 11 - inferior frontal gyrus pars opercularis, 13 - inferior frontal gyrus pars triangularis, and 15 - inferior frontal gyrus pars orbitalis) and the SPM *a priori* gray image (50% prior probability of gray matter). Group statistical maps were rendered on a canonical brain using SPM99. For the purpose of additional anatomical precision, group contrasts were also rendered on an MNI canonical brain that underwent cortical "inflation" using FreeSurfer (CorTechs Labs, Inc.) (Dale, Fischl, & Sereno, 1999; Fischl, Sereno, & Dale, 1999).

To reveal common effects at the voxel-level across independent conditions and data sets, conjunction analyses were performed. Unless otherwise noted, conjunction analyses were assessed as significant at a conjoint alpha level of $p < .001$. That is, a significant conjunction does not indicate that both contrasts were individually significant at standard thresholds (Nichols, Brett, Andersson, Wager, & Poline, in press), but rather means that both were significant at more lenient thresholds (with a joint probability of a Type I being less than .001).

The group-level voxel-based contrasts were supplemented with region-of-interest (ROI) analyses. All significant voxels within a 6-mm radius of a chosen maximum defined an ROI, and unless otherwise noted, were defined from the conjunction of All

Control conditions > Fixation for the first and second epochs of Exp. 2 (Fig. 4A). Selective averaging with respect to peristimulus time allowed assessment of the signal change associated with each condition. Integrated percent signal change (iPSC) was then computed based on the peak plus and minus one TR. The peak was defined neutrally for each ROI based on the average timecourse across all conditions. The resultant data were subjected to repeated-measures analyses of variance (ANOVA).

## Results

### *Simple Behavioral Effects*

Initial analyses of reaction time (RT) and errors confirmed the efficacy of the four control manipulations (Fig. 2). Judgment Specificity reliably impacted RT, such that Feature judgments took longer than Relatedness judgments (Exp. 1: $F(1,21) = 70.6$, $p < .0001$; Exp. 2: $F(1,10) = 114.1$, $p < .0001$; Fig. 2A), indicating that RT slowed as putative selection demands increased. Though errors were slightly higher during Related (12% errors) than Feature judgments (10% errors) in Exp. 1 ($F(1,21) = 8.1$, $p < .01$), there was no such difference in Exp. 2 ($F(1,10) = 2.7$, $p = .13$).

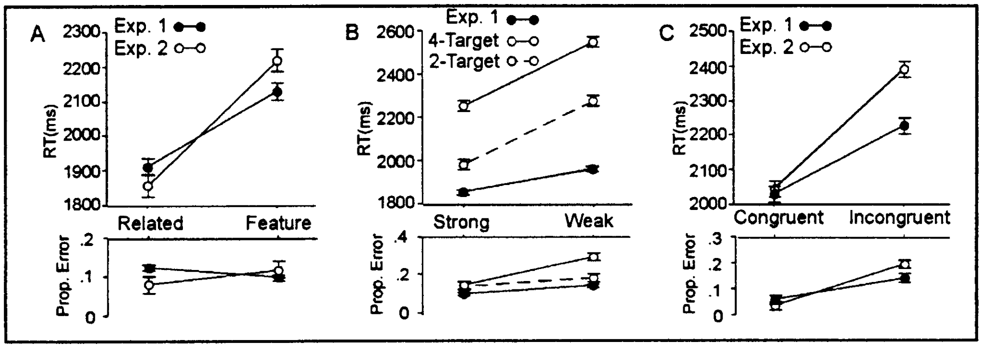

**Figure 2.** Impact on RT and Errors of (A) Judgment Specificity, (B) Associative Strength and Number of Targets (Four Targets = solid line; Two Targets = dashed line), and (C) Congruency in Exp. 1 (solid circles) and Exp. 2 (open circles).

Cue–target Associative Strength impacted both RT and errors. RT was longer (Exp. 1: $F(1,21) = 17.5$, $p < .0005$; Exp. 2: $F(1,10) = 171.4$, $p < .0001$) and errors were higher (Exp. 1: $F(1,21) = 7.5$, $p < .05$; Exp. 2: $F(1,10) = 60.9$, $p < .0001$) when one of the targets was a Weak associate of the cue than when one of the targets was a Strong associate (Fig. 2B). The crossing of Associative Strength with Judgment Specificity in Exp. 1 revealed a reliable interaction (RT: $F(1,21) = 11.4$, $p < .005$; errors: $F(1,21) = 5.4$, $p < .05$), with the effects of Associative Strength being reliable during Relatedness judgments (RT: $F(1,21) = 36.4$, $p < .0001$; errors: $F(1,21) = 12.1$, $p < .002$) but not during Feature judgments (RT: $F = 1.6$; Errors: $F < 1$). This pattern is consistent with Associative Strength impacting controlled retrieval demands during the Relatedness task, but not during the Feature task (which requires selection of specific stimulus features).

Number of Targets impacted performance, such that selecting from amongst Four targets in Exp. 2 slowed RT ($F(1,10) = 9.1$, $p < .05$) and increased errors ($F(1,10) = 4.7$, $p = .055$; Fig. 2B) relative to when there were Two targets. Number of Targets and Associative Strength did not interact ($F < 1$). Importantly, central to a subsequent analysis conducted to rule out time-on-task accounts of the fMRI data, there was no behavioral difference between Weak–Two vs. Strong–Four trials (RT: $F < 1$; Errors: $F(1,10) = 1.6$, $p = .23$). Finally, the Congruency manipulation of selection demands affected both RT and errors, such that RT slowed (Exp. 1: $F(1,21) = 66.2$, $p < .0001$; Exp. 2: $F(1,10) = 142.9$, $p < .0001$) and Errors increased (Exp. 1: $F(1,21) = 21.1$, $p < .0005$; Exp. 2: $F(1,10) = 24.9$, $p < .0005$) on Incongruent relative to Congruent trials (Fig. 2C).

*Principal Components of Behavior*

Factors extracted from PCA can serve as meta-variables that account for more total variance across behavioral measures than any of the measures contribute in isolation. In the current context, a component that accounts for variance across the experimental manipulations might reflect the influence of a common control process. Accordingly, principal components were extracted from the standardized differences in errors and RT due to Associative Strength (Weak - Strong), Judgment Specificity (Feature - Related), and Congruency (Incongruent - Congruent) for subjects in Exp. 1 and Exp. 2 (total N = 33). Number of Targets was not included in this analysis as it only varied in Exp. 2.

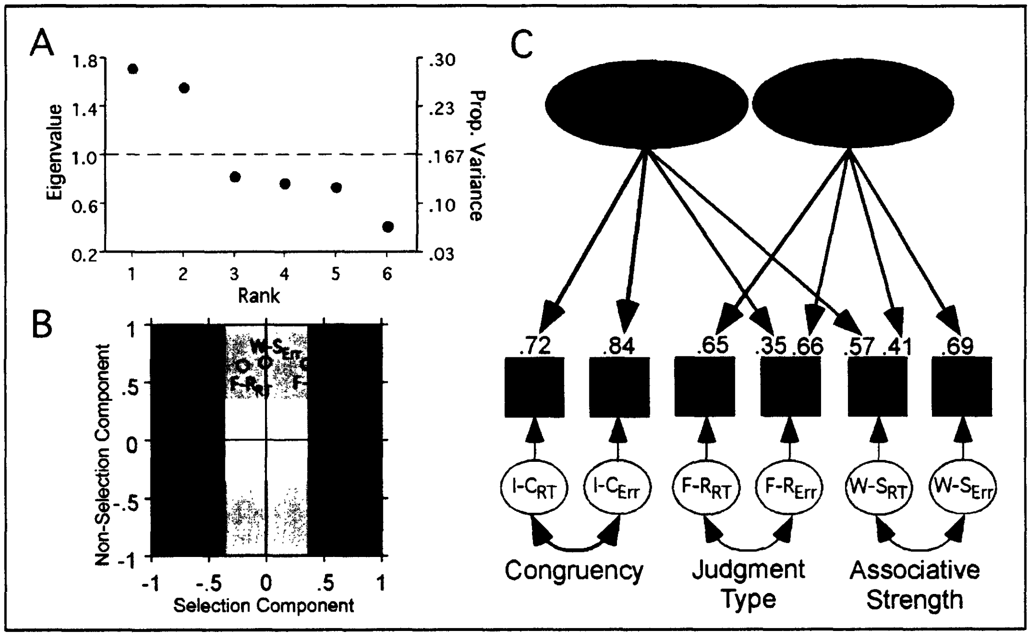

**Figure 3.** Results from the PCA. (A) The scree plot depicts the six initial unrotated factors ("Selection Component" = red; "Non-Selection Component" = blue) including rank (x-axis), eigenvalue (left y-axis), and proportion of overall variance accounted for by each factor (right y-axis). (B) Spatial representation of the factor loadings of the six behavioral measures (points) plotted in a space defined on the Selection Component (x axis) and Non-Selection Component (y-axis). The further along a given axis a point is from the origin, the stronger its relationship with that component. Points in the red zone may be considered strongly related to the Selection Component and points in the blue zone are strongly related to the Non-Selection Component. (C) Diagram depicting the mapping of the two components (ovals) onto the variances (boxes) associated with each measure. Numbers represent factor loadings, curved arrows connect correlated factors, and colored shading represents the proportion of explained variance accounted for by the Selection Component (red) and Non-Selection Component (blue). Together, (B) and (C) illustrate that the Selection Component is shared across all three manipulations of control. [Note: F-R=Feature-Related; I-C=Incongruent-Congruent; W-S=Weak-Strong; RT=reaction time; Err=Errors]

PCA revealed two factors that accounted for over half (54%) the variance in the six behavioral measures (Fig. 3A). Loadings of the six individual measures on the first factor revealed a common component accounting for variance due to Congruency, Judgment Specificity, and Associative Strength (Fig. 3B and C; Table 1). Based on its pattern of loadings—particularly noting the strong association with Congruency—we suggestively refer to this meta-variable as the "*Selection Component*". Congruency loaded heavily and almost exclusively on the Selection Component (Fig. 3B and C), with this factor accounting for 51% and 71% of the variance in the Congruency RT and Error effects, respectively. Critically, manipulations of Associative Strength and Judgment Specificity also produced behavioral effects that loaded on the Selection Component (Fig. 3B and C; Table 1).

By contrast, the second factor (*"Non-Selection Component"*) accounted for variance in RT and Error effects of Associative Strength and Judgment Specificity, but accounted for practically no variance due to Congruency (Fig. 3B and C; Table 1). Hence, only the Selection Component indexed behavioral variance common to all three control manipulations, whereas the Non-Selection Component, though accounting for variance common to Associative Strength and Judgment Specificity, was not associated with Congruency, and thus does not likely reflect a source of variance due to selection demands.

**Table 1.** Percentage of variance in each behavioral measure accounted for by the two extracted components

| Measure | Selection Component | Non-Selection Component | Total |
|---|---|---|---|
| **Congruency RT** | 51.3 | 0.3 | 51.5 |
| **Congruency Err.** | 70.7 | 4.6 | 75.3 |
| **Judg. Type RT** | 4.3 | 42.5 | 46.8 |
| **Judg. Type Err.** | 12.3 | 43.8 | 56.1 |
| **Assoc. Str. RT** | 32.6 | 16.7 | 49.3 |
| **Assoc. Str. Err.** | 0.0 | 47.2 | 47.2 |

*Correlates of Semantic Processing*

Relative to fixation, semantic processing (collapsed across condition and restricted to accurate responses) elicited activation throughout left VLPFC, as well as in posterior cortices (Fig. 4A). Within left VLPFC, activation extended from a posterior region (~BA 44/6) to a more anterior mid-VLPFC region (~BA 45), corresponding to inferior frontal gyrus pars opercularis and pars triangularis (Fig. 4B), and also to an anterior and ventral VLPFC region (~BA 47), corresponding to inferior frontal gyrus pars orbitalis (Fig. 4B).

It should be noted that the anterior and mid-VLPFC subregions defined here, correspond to a division of what has been previously termed anterior left inferior prefrontal cortex (aLIPC). As noted above, these subregions correspond to existing anatomical and approximate cytoarchitechtonic subdivisions of the inferior frontal gyrus, and may be largely distinguished in anterior slices based on their relationship to the horizontal ramus of the lateral fissure (Fig. 4B). In part, this finer fractionation has been adopted in light of recent observations (Badre and Wagner, in press; Dobbins and Wagner, in press) suggesting functional distinctions among these subregions. Posterior VLPFC corresponds to what has been previously referred to as posterior LIPC (pLIPC).

Beyond PFC, activation was evident in left middle temporal cortex (~BA 21/22), a region previously associated with semantic retrieval (Bokde, Tagamets, Friedman, & Horwitz, 2001; Dobbins & Wagner, in press; Martin, Wiggs, Lalonde, & Mack, 1994; Petersen et al., 1988; Wagner et al., 2001) and that functionally couples with left anterior

VLPFC (Bokde et al., 2001; Dobbins & Wagner, in press). Given the present focus on selection and controlled semantic retrieval, subsequent analyses focused on responses in left VLPFC and middle temporal cortex.

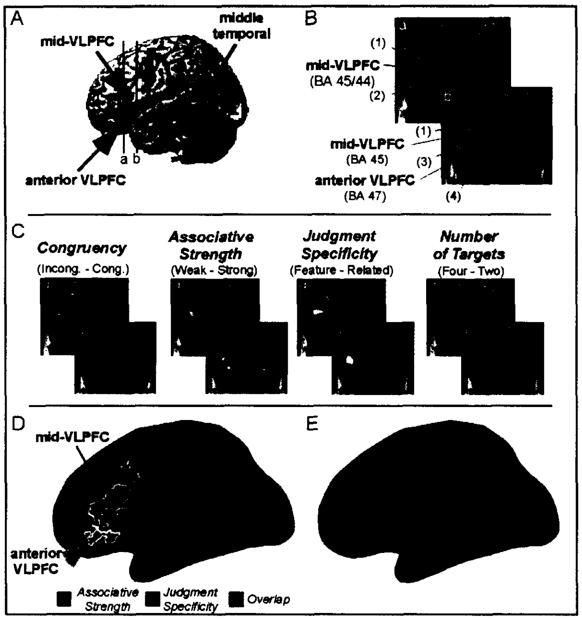

**Figure 4.** VLPFC responses to control manipulations. (A) Surface rendered conjunction of All Control conditions > Fixation for the first and second epochs of Exp. 2 with critical ROIs labeled. (B) Coronal slices ($y = 16$ and $y = 30$) from canonical brain demarcating the anatomical boundaries by which activation foci were assigned to mid-VLPFC (inferior frontal gyrus pars triangularis and pars opercularis) or anterior VLPFC (inferior frontal gyrus pars orbitalis). Labeled anatomical landmarks are (1) inferior frontal sulcus, (2) insular sulcus, (3) horizontal ramus of the lateral fissure, (4) orbital gyrus. (C) Contrasts of Congruency (conj. of Exp. 1 and Exp. 2, $p < .001$), Associative Strength (Exp. 2, $p < .001$), Judgment Specificity (conj. of Exp. 1 and Exp. 2, $p < .001$), and Number of Targets (Exp. 2, $p < .005$). The cross-experiment conjunction of Associative Strength was not calculated, as differences in Associative Strength between experiments were not comparable (see Methods). (D) Contrasts of Associative Strength (blue) and Judgment Specificity (red) and their overlap (purple) are rendered on an inflated MNI canonical surface. Substantial anatomical and functional separability is observed between anterior VLPFC, which was selectively sensitive to Associative Strength, and mid-VLPFC, which was sensitive to Associative Strength and Judgment Specificity. Moreover, the effect of Congruency (not plotted for ease of viewing) overlapped with those of Associative Strength and Judgment Specificity in mid-VLPFC (see C). (E) Rendering of the Weak-Two > Strong-Four convergence map between Exp. 2 from the present study and the corresponding contrast from Wagner et al. (2001) highlighting the replication of the ventral anterior VLPFC focus.

### *Neural Effects of Congruency*

The Congruency manipulation loaded most specifically on the Selection Component meta-variable. Hence, this control manipulation putatively provides a relatively pure starting point for indexing the neural substrates of a generalized selection process. In Exp. 1, an Incongruent > Congruent effect was observed in left fronto-operculum ($p < .05$, corrected), with greater activation on Incongruent trials also present in left mid-VLPFC (-54 15 18) at an uncorrected threshold ($p < .001$). Replicating this mid-VLPFC effect, comparison of Incongruent > Congruent Feature trials in Exp. 2 ($p <$ .05, corrected) revealed reliable activation in left mid-VLPFC (-45 18 24) that extended posteriorly (-45 9 27). A formal test of the convergence of the Congruency effects across Exps. 1 and 2 revealed activation in left mid-VLPFC (-48 18 18), posterior (-42 9 21) and dorsal, anterior (-45 39 3) VLPFC subregions (Fig. 4C), and bilateral fronto-operculum.

*Neural Effects of Judgment Specificity, Associative Strength, and Number of Targets*

The Selection Component also accounted for a portion of the behavioral variance due to Judgment Specificity and Associative Strength, suggesting a common source of variance between these control manipulations and the Congruency manipulation. This PCA outcome predicts a convergence in the patterns of left VLPFC activation engaged by these control manipulations. Consistent with this perspective, contrasts of Associative Strength (Weak > Strong) and Judgment Specificity (Feature > Related) revealed activation in left mid-VLPFC ($p < .05$, corrected), inclusive of the voxels showing a Congruency effect (Fig. 4C). Direct overlap of the Judgment Specificity and Associative Strength contrast maps revealed extensive convergence in their engagement of left mid-VLPFC (~BA 45) extending into posterior VLPFC (Fig. 4D; purple region). Given this high overlap, it is notable that the most anterior and ventral extent of left VLPFC (-51 27 –3 and -48 30 –12), corresponding to the inferior frontal gyrus pars orbitalis (~BA 47), appeared selectively sensitive to Associative Strength (Fig. 4D; blue region), a finding to which we return below.

The Number of Targets manipulation (Four > Two) in Exp. 2 revealed no reliable activation in left VLPFC at the corrected threshold. At a moderately reduced threshold ($p < .005$, uncorrected), activation was observed in left posterior (-45 12 27) and mid-

VLPFC (-51 27 15) (Fig. 4C), replicating Number of Targets effects identified in a prior experiment (Wagner et al., 2001) (-39 6 24 and -45 27 9). Importantly, localization of these effects converged with the mid-VLPFC activation common to the Congruency, Judgment Specificity, and Associative Strength contrasts (Fig. 4C).

### *Mid-VLPFC and the Selection Component*

The three manipulations included in the PCA behavioral analysis resulted in functional effects within left VLPFC, with overlapping activation in mid- to posterior VLPFC (Fig. 4). This finding raises the possibility that this region supports a common control process that may be indexed by the Selection Component meta-variable. To test this hypothesis, the fMRI indices of each control manipulation were correlated with the factor scores derived from the two extracted principal components. Specifically, a conjunction analysis was conducted to test for the convergence of regions showing a correlation between the Selection Component and the associated neural effects of Congruency,

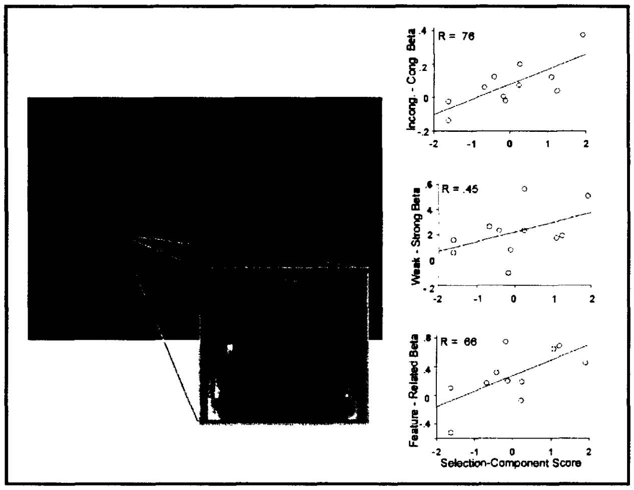

**Figure 5.** The Selection Component accounted for variance in left mid-VLPFC functional activation. The PCA meta-variable served as a covariate during fMRI analyses of Congruency, Associative Strength, and Judgment Specificity. Conjunction of these covariate effects (conjoint $p < .000125$) revealed that the Selection Component reliably accounted for function variance in left mid-VLPFC activation (–54 21 12), here rendered on an inflated canonical surface. Also, plotted are Beta values extracted from left mid VLPFC (y-axis) against the Selection Component factor score (x-axis) for Congruency, Associative Strength, and Judgment Specificity manipulations.

Associative Strength, and Judgment Specificity. Significance was assessed at a conservative threshold (conjoint alpha = .000125), providing confidence in rejection of the conjunction null (Nichols et al., in press). Strikingly, this analysis implicated left mid-VLPFC (-54 21 12) as the only convergent site at which all control contrasts were correlated with the Selection Component (Fig. 5). This novel analysis strongly implicates neural processes in left mid-VLPFC as coupled with the common variance in behavior indexed by the Selection Component.

The Non-Selection Component did not account for variance in the Congruency manipulation, but loaded strongly on the effects of Associative Strength and Judgment Specificity (Fig. 3). A conjunction analysis (conjoint alpha = .0025) between the correlation of the Non-Selection Component and the Associative Strength and the Judgment Specificity neural effects revealed convergent activation in left fronto-polar cortex (FPC; -42 45 –3), well rostral to the anterior VLPFC region, described in detail below, that was selectively sensitive to Associative Strength (Fig. 4D).

### *Anterior VLPFC and Controlled Semantic Retrieval*

In contrast to left mid-VLPFC, which was engaged across all control manipulations and was associated with the Selection Component, an anterior and ventral focus in left VLPFC was specifically sensitive to Associative Strength (Fig. 4C and D). Moreover, the contrast of Weak–Two > Strong–Four trials, which is behaviorally matched for time-on-task, revealed differential activation restricted precisely to this anterior and ventral locus of left VLPFC (-45 27 –15), with this pattern converging with that seen in a prior study of controlled semantic retrieval (Fig. 4E; Wagner et al., 2001).

The selective nature of the response in the ventral anterior extent of left VLPFC was confirmed by ROI analyses (Fig. 6). Specifically, the ventral anterior extent of left VLPFC (-45 27 –15) revealed a robust effect of Associative Strength (F(1,10) = 20.1, *p* < .005), but did not show reliable effects of Judgment Specificity (F(1,10) = 2.5, *p* = .14), Congruency (F(1,10) = 2.1, *p* = .18), nor Number of Targets (F(1,10) = .11, *p* = .75).

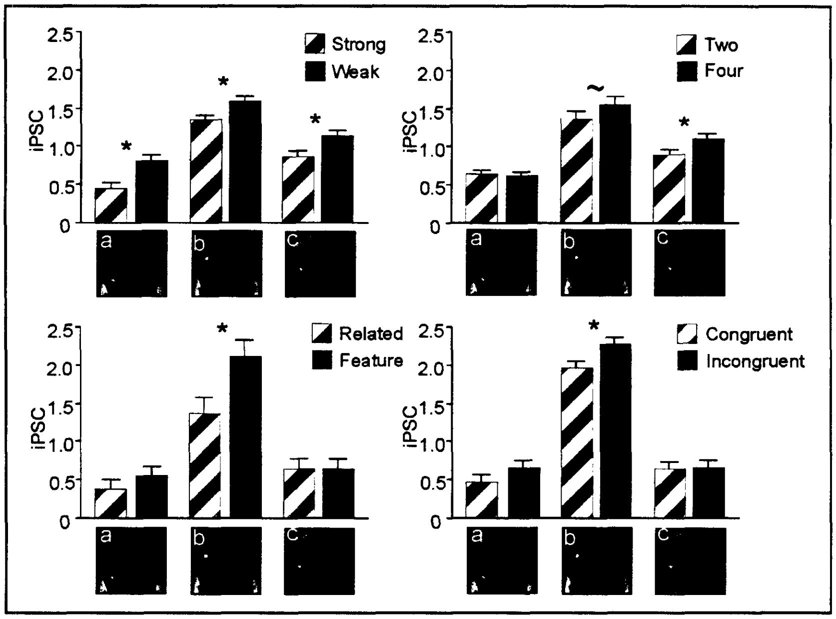

**Figure 6.** Integrated percent signal change data from ROIs in (a) anterior VLPFC (-54 27 –9), (b) posterior/mid-VLPFC (-51 15 33), and (c) middle temporal cortex (-48 –48 3) reveal the sensitivity of each region to Associative Strength (top-left), Number of Targets (top-right), Judgment Specificity (bottom-left), and Congruency (bottom-right) manipulations. Anterior VLPFC showed selective sensitivity to Associative Strength, middle temporal cortex showed sensitivity to Associative Strength and Number of Targets, whereas mid-VLPFC that was sensitive to all control manipulations.

This pattern qualitatively differed from that in left mid- VLPFC (-51 15 33), which showed effects of all four manipulations, as evident in a Manipulation [Associative Strength, Number of Targets, Judgment Specificity, Congruency] x Region [anterior VLPFC, mid-VLPFC) interaction (F(3,30) = 5.0, *p* < .01). This outcome strongly supports an anatomical and functional delineation between controlled retrieval and selection.

### *Middle Temporal Cortex and Semantic Retrieval*

Semantic processing elicited activation in a left middle temporal region (Fig. 4A) previously implicated in studies of semantic retrieval. Given the distinct predictions made by the selection and controlled retrieval hypotheses regarding activation in regions thought to store long-term semantic knowledge, ROI analyses assessed the sensitivity of left middle temporal cortex (-48 –48 3) to the four control manipulations (Fig. 6). There were two important findings. First, left middle temporal cortex was sensitive to the two manipulations that putatively varied the amount of semantic knowledge retrieved (Associative Strength: $\underline{F}(1,10) = 13.3, p < .005$; Number of Targets: $\underline{F}(1,10) = 11.9, p < .01$). Semantic retrieval demands putatively increase across Weak versus Strong cue–target Associative Strength trials because during Weak trials additional knowledge, above and beyond that emerging through automatic retrieval processes, must be recovered in a top-down manner to guide the decision. Semantic retrieval is also greater when there are Four versus Two targets, as semantic knowledge is recovered about more stimuli in the former case. In this instance, the differential semantic retrieval may emerge from bottom-up (automatic) processes, as argued by others (Thompson-Schill et al., 1997), thus resulting in an effect of Number of Targets in middle temporal cortex but not in left anterior VLPFC. Consistent with this perspective, left anterior VLPFC and middle temporal ROIs functionally dissociated across the two retrieval manipulations, as evidenced by a Manipulation [Associative Strength, Number of Targets] × Region interaction ($\underline{F}(1,10) = 10.7, p < .001$).

Second, whereas left middle temporal cortex was sensitive to semantic retrieval, it was insensitive to Judgment Specificity and Congruency ($\underline{F}$s < 1), providing novel evidence that the selection processes subserved by left mid-VLPFC operate post-retrieval. Importantly, the pattern of left middle temporal activation, which may mark semantic retrieval, dissociated from that in left mid-VLPFC, which putatively marks selection, as evidenced by a Manipulation [Associative Strength, Number of Targets, Judgment Specificity, Congruency] × Region interaction ($\underline{F}(3,30) = 9.4, p < .0005$) and by a Selection Demands [Judgment Specificity, Congruency] × Region interaction ($\underline{F}(1,10) = 7.8, p < .05$).

## Discussion

The present data indicate that controlled retrieval and selection processes make distinct contributions to the regulation of memory and are mediated by anatomically separable subregions of left VLPFC. As such, these data offer resolution to the debate over left VLPFC function, and advance mechanistic understanding of the relation between top-down retrieval and selection. Two central findings warrant attention.

First, our data provide novel evidence for a general selection process that operates across multiple semantic control conditions and is mediated by left mid-VLPFC. Process commonality was initially established through detection of functional overlap in left mid-VLPFC (~BA 45) across the four control manipulations, consistent with task analyses suggesting that each varied selection demands. Process commonality was further established through identification of a meta-variable that accounted for behavioral variance common to three of the control manipulations. Strikingly, the variance in this meta-variable correlated with modulations in left mid-VLPFC activation.

Second, left anterior VLPFC (~BA 47) was exclusively engaged in response to increased demands on the top-down retrieval of semantic knowledge, rather than post-retrieval selection. That is, left anterior VLPFC was selectively sensitive to cue-target Associative Strength, with this functional pattern dissociating from that in left mid-VLPFC, thus suggesting a role in activating long-term knowledge rather than resolving competition. This interpretation garners further support when considering the pattern of activation in left middle temporal cortex, a region that stores semantic knowledge and thus was expected to be sensitive to amount of semantic retrieval, be it knowledge accessed via controlled retrieval (indexed by Associative Strength) and via more automatic retrieval routes (indexed by Number of Targets). Importantly, left middle temporal activation varied with Associative Strength and Number of Targets, but showed little sensitivity to selection demands (Congruency and Task Specificity).

Collectively, these findings motivate a two-process model of fronto-temporal control of semantic memory. Retrieval of semantic knowledge stored in lateral temporal cortex may emerge through bottom-up (automatic) and/or top-down (controlled) mechanisms, with the latter mediated by left anterior VLPFC. Once retrieved, selection of task-relevant representations from amongst retrieved competitors is required, with selection being mediated by left mid-VLPFC.

### *Post-Retrieval Selection*

The demand to select task-relevant representations from retrieved alternatives may be common to many contexts. Because stimuli are capable of automatically cueing more than one associate, any retrieval act holds the possibility of some competition from irrelevant, retrieved information (M. C. Anderson & Spellman, 1995; Badre & Wagner, 2002). Indeed, even the manipulation of Associative Strength, which we previously argued to impact controlled retrieval demands without consequences for selection

(Wagner et al., 2001), can also result in variable competition due to the presence of irrelevant competitors or an "underdetermined response" on Weak trials (Kan and Thompson-Schill, 2004; Thompson-Schill et al., 2005). Consistent with this perspective, increased semantic retrieval due to top-down (Associative Strength) or more automatic (Number of Targets) processes served to up-regulate demands on left mid-VLPFC. In contrast to left middle temporal cortex, however, up-regulation of left mid-VLPFC activation did not simply track the amount of information retrieved. Rather, increased activation also accompanied manipulations that directly varied the degree of competition between retrieved alternatives, while putatively holding semantic retrieval constant (Judgment Specificity and Congruency). This pattern, together with the striking observation that across-manipulation behavioral variance in the "Selection Component" accounted for functional variance in left mid-VLPFC, provides particularly compelling evidence in favor of a selection interpretation of left mid-VLPFC function.

Critically, the dissociation between left mid-VLPFC and lateral temporal cortex suggests that the representations on which this general selection process operates are not necessarily long-term semantic representations, of the sort thought to be stored in lateral temporal regions. In particular, it appears reasonable to designate this selection process as occurring post-retrieval, operating on active representations that perhaps are being maintained in working memory. As noted above, this proposal does not require that retrieval itself is all-or-none. Indeed, active representations entering working memory may be partial, transient, and even weak. However, it does require a distinction between active representations (putatively maintained in working memory) and long-term memory representations. There is evidence for such a distinction in the non-human primate

(Miller et al., 1996), where disruption of active neural representations in temporal cortices through interference does not disrupt representations actively maintained in PFC, which are capable of subsequently guiding action. Furthermore, a distinction between stored long-term representations and working memory representations that guide action does have theoretical precedence (e.g., O'Reilly et al., 2002).

Although the present manipulations of selection demands were within the context of semantic processing, it is not necessarily the case that the operation of this mechanism must be restricted to task contexts of semantic control or even memory in general. Indeed, the common factor influencing whether tasks elicit activation in this region appears to be whether they involve selection or interference resolution en route to generating a response. For example, left mid-VLPFC has been associated with increased interference within working memory (Badre & Wagner, in press; Jonides, Smith, Marshuetz, Koeppe, & Reuter-Lorenz, 1998; Thompson-Schill et al., 2002), during task switching (Brass & von Cramon, 2004), and in response selection (Jiang & Kanwisher, 2003; Milham et al., 2001). Furthermore, during episodic remembering, left mid-VLPFC has been associated with selection of perceptual and conceptual episodic details (Dobbins & Wagner, in press), suggesting that this region resolves conflict across memory and content domains. At a mechanistic level, this domain-general selection process may bias active representations maintained in working memory to overcome conflict, thereby permitting selection of relevant representations from "noise" due to other active competitors.

### *Controlled Semantic Retrieval*

Though a generalized selection process may play a role in resolving interference, the present data provide evidence of a dissociation across left VLPFC subregions for which a single–process model does not provide an account. The anterior, ventral extent of left VLPFC dissociated from mid-VLPFC, as left anterior VLPFC was exclusively sensitive to Associative Strength (Fig. 4D). Importantly, this was the case even when pitting controlled retrieval demands (Associative Strength) against overall retrieval (Number of Targets), as left anterior VLPFC was the only region to show a Weak–Two > Strong–Four effect. This pattern suggests that left anterior VLPFC is uniquely sensitive to the need to control retrieval when available cues are insufficient to activate relevant knowledge through bottom-up processes.

In operation, a controlled retrieval mechanism may accumulate and maintain cues or retrieval goals to mediate retrieval of additional relevant information stored in left temporal cortices (e.g., Badre & Wagner, 2002). Consistent with this interpretation, left anterior VLPFC activation due to Associative Strength was accompanied by similar activation in left middle temporal cortex. Moreover, prior studies have demonstrated a functional coupling between left anterior VLPFC and left middle temporal cortex during semantic processing (Bokde et al., 2001) and episodic recollection of conceptual event details (Dobbins & Wagner, in press). Collectively, these data suggest that left anterior VLPFC may operate on representations in middle temporal cortex, though a metric of causality or directionality of information flow awaits future research (Friston, Harrison, & Penny, 2003; Goebel, Roebroeck, Kim, & Formisano, 2003; Sun, Miller, & D'Esposito, 2004).

In our account for the observed dissociation between anterior VLPFC and mid-VLPFC, the critical distinction between controlled retrieval and selection putatively derives from the nature of the representations on which each process operates. The controlled retrieval process subserved by left anterior VLPFC may directly influence long-term semantic representations stored in lateral temporal regions. By contrast, the generalized selection process supported by left mid-VLPFC may be critical in resolving interference among active representations maintained in working memory. One implication of the close association of controlled retrieval with the activation of stored representations is that this process should be tied more directly to tasks that demand access to long-term memory, whereas the selection process may be required to resolve interference among representations in working memory that came to be activated through means other than semantic retrieval.

The ability to flexibly and strategically access knowledge is a central feature of an adaptive control system (J. R. Anderson et al., 2004; Sohn et al., 2003). The present results argue that distinct control mechanisms in left VLPFC contribute to this process by guiding access to semantic knowledge not retrieved automatically and then selecting from amongst retrieved representations. The network proposed here may be central to a number of task contexts in which representations must be retrieved or selected en route to generating a response. Further research promises to further specify the nature of these control mechanisms so as to better understand when they are necessary for successful adaptive behavior. In line with this goal, Chapter 4 reports an experiment designed to investigate how a selection process that is sensitive to conflict among competing conceptual representations can account for behavioral and fMRI data, particularly from left mid-VLPFC, during task switching.

## References

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychol Rev, 111*(4), 1036-1060.

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: memory retrieval as a model case. *Psychological Review, 102*(1), 68-100.

Badre, D., & Wagner, A. D. (2002). Semantic retrieval, mnemonic control, and prefrontal cortex. *Behavioral and Cognitive Neuroscience Reviews, 1*(3), 206-218.

Badre, D., & Wagner, A. D. (in press). Frontal lobe mechanisms that resolve proactive interference. *Cereb Cortex.*

Bokde, A. L. W., Tagamets, M.-A., Friedman, R. B., & Horwitz, B. (2001). Functional interactions of the inferior frontal cortex during the processing of words and word-like stimuli. *Neuron, 30*, 609-617.

Brass, M., & von Cramon, D. Y. (2004). Decomposing components of task preparation with functional magnetic resonance imaging. *J Cogn Neurosci, 16*(4), 609-620.

Bunge, S. A., Wendelken, C., Badre, D., & Wagner, A. D. (2005). Analogical reasoning and prefrontal cortex: Evidence for separable retrieval and integration mechanisms. *Cereb Cortex, 15*, 239-249.

Cocosco, C. A., Kollokian, V., Kwan, R. K. S., & Evans, A. C. (1997). Brainweb: Online interface to a 3D MRI simulated brain database. *NeuroImage, 5*, 425.

Dale, A. M. (1999). Optimal experimental design for event-related fMRI. *Hum Brain Mapp, 8*(2-3), 109-114.

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage, 9*(2), 179-194.

Demb, J. B., Desmond, J. E., Wagner, A. D., Vaidya, C. J., Glover, G. H., & Gabrieli, J. D. (1995). Semantic encoding and retrieval in the left inferior prefrontal cortex: a functional MRI study of task difficulty and process specificity. *J Neurosci, 15*(9), 5870-5878.

Devlin, J. T., Matthews, P. M., & Rushworth, M. F. (2003). Semantic processing in the left inferior prefrontal cortex: a combined functional magnetic resonance imaging and transcranial magnetic stimulation study. *Journal of Cognitive Neuroscience, 15*(1), 71-84.

Dobbins, I. G., & Wagner, A. D. (in press). Domain-general and domain-sensitive prefrontal mechanisms for recollecting events and detecting novelty. *Cereb Cortex.*

Donaldson, D. I., Petersen, S. E., Ollinger, J. M., & Buckner, R. L. (2001). Dissociating state and item components of recognition memory using fMRI. *Neuroimage, 13*(1), 129-142.

Dove, A., Pollmann, S., Schubert, T., Wiggins, C. J., & von Cramon, D. Y. (2000). Prefrontal cortex activation in task switching: an event-related fMRI study. *Brain Res Cogn Brain Res, 9*(1), 103-109.

Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage, 9*(2), 195-207.

Fletcher, P. C., Shallice, T., & Dolan, R. J. (2000). "Sculpting the response space"--an account of left prefrontal activation at encoding. *Neuroimage, 12*(4), 404-417.

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *Neuroimage, 19*(4), 1273-1302.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-B., Frith, C. D., & Frackowiak, R. S. J. (1995). Statistical parametric maps in functional imaging: A general linear approach. *Hum Brain Mapp, 2*, 189-210.

Goebel, R., Roebroeck, A., Kim, D. S., & Formisano, E. (2003). Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. *Magn Reson Imaging, 21*(10), 1251-1261.

Gold, B. T., Balota, D. A., Kirchhoff, B. A., & Buckner, R. L. (2005). Common and Dissociable Activation Patterns Associated with Controlled Semantic and Phonological Processing: Evidence from fMRI Adaptation. *Cereb Cortex.*

Harris, C. W. (1967). On Factors and Factor Scores. *Psychometrika, 32*(4), 363-379.

Jiang, Y., & Kanwisher, N. (2003). Common neural substrates for response selection across modalities and mapping paradigms. *J Cogn Neurosci, 15*(8), 1080-1094.

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

Kan, I. P., & Thompson-Schill, S. L. (2004). Effect of name agreement on prefrontal activity during overt and covert picture naming. *Cogn Affect Behav Neurosci, 4*(1), 43-57.

Kapur, S., Rose, R., Liddle, P. F., Zipursky, R. B., & et al. (1994). The role of the left prefrontal cortex in verbal processing: Semantic processing or willed action? *Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience, 5*(16), 2193-2196.

Kucera, H., & Francis, W. N. (1967). *Computational Analysis of Present-day English.* Providence, RI: Brown University Press.

Martin, A., & Chao, L. L. (2001). Semantic memory and the brain: structure and processes. *Curr Opin Neurobiol, 11*(2), 194-201.

Martin, A., Wiggs, C. L., Lalonde, F., & Mack, C. (1994). Word retrieval to letter and semantic cues: a double dissociation in normal subjects using interference tasks. *Neuropsychologia, 32*(12), 1487-1494.

Milham, M. P., Banich, M. T., Webb, A., Barad, V., Cohen, N. J., Wszalek, T., et al. (2001). The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on the nature of conflict. *Cognitive Brain Research, 12*, 467-473.

Miller, E. K., Erickson, C. A., & Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *Journal of Neuroscience, 16*, 5154-5167.

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

Moss, H., & Older, L. (1996). *Birkbeck Word Association Norms.* Hove, UK: Psychology Press.

Moss, H. E., Fletcher, P. C., Abdallah, S., Pilgrim, L. K., Acres, K., Bright, P., et al. (in press). Selecting among competing alternatives: Selection and controlled retrieval in the left prefrontal cortex. *Cereb Cortex.*

Nichols, T., Brett, M., Andersson, J., Wager, T., & Poline, J.-B. (in press). Valid conjunction inference with the minimum statistic. *NeuroImage.*

Noppeney, U., & Price, C. J. (2004). Retrieval of abstract semantics. *Neuroimage, 22*(1), 164-170.

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

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single- word processing. *Nature, 331*(6157), 585-589.

Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage, 10*(1), 15-35.

Postman, L., & Keppel, G. (1970). *Norms of Word Association.* New York: Academic Press.

Shimamura, A. P. (1995). Memory and frontal lobe function. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences* (pp. 803-813). Cambridge, MA: MIT Press.

Sohn, M. H., Goode, A., Stenger, V. A., Carter, C. S., & Anderson, J. R. (2003). Competition and representation during memory retrieval: roles of the prefrontal cortex and the posterior parietal cortex. *Proc Natl Acad Sci U S A, 100*(12), 7412-7417.

Sun, F. T., Miller, L. M., & D'Esposito, M. (2004). Measuring interregional functional connectivity using coherence and partial coherence analyses of fMRI data. *Neuroimage, 21*(2), 647-658.

Thompson-Schill, S. L. (2003). Neuroimaging studies of semantic memory: inferring "how" from "where". *Neuropsychologia, 41*(3), 280-292.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proc Natl Acad Sci U S A, 94*(26), 14792-14797.

Thompson-Schill, S. L., D'Esposito, M., & Kan, I. P. (1999). Effects of repetition and competition on activity in left prefrontal cortex during word generation. *Neuron, 23*(3), 513-522.

Thompson-Schill, S. L., Jonides, J., Marshuetz, C., Smith, E. E., D'Esposito, M., Kan, I. P., et al. (2002). Effects of frontal lobe damage on interference effects in working memory. *Cogn Affect Behav Neurosci, 2*(2), 109-120.

Thompson-Schill, S. L., Bedny, M., & Goldberg, R. F. (2005). The frontal lobes and the regulation of mental activity. *Current Opinion in Neurobiology, 15,* 219-224.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage, 15*(1), 273-289.

Wagner, A. D., Paré-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning: Left prefrontal cortex guides controlled semantic retrieval. *Neuron, 31*, 329-338.

Zhang, J. X., Zhuang, J., Ma, L., Yu, W., Peng, D., Ding, G., et al. (2004). Semantic processing of Chinese in left inferior prefrontal cortex studied with reversible words. *Neuroimage, 23*(3), 975-982.

# Chapter 4

## Resolution of Interference by Left Ventrolateral Prefrontal Cortex During Task Switching

The fMRI experiments described in Chapters 2 and 3 have highlighted a portion of left mid-VLPFC (Figure 1), corresponding to the left inferior frontal gyrus *pars triangularis* (Brodmann's Area [BA] 45), under conditions of interference similar to those proposed to be the source of RT task switching costs (e.g., Allport & Wylie, 2000; Koch, Prinz, & Allport, 2005; Waszak, Hommel, & Allport, 2003; Wylie & Allport, 2000). In particular, the experiment reported in Chapter 2 demonstrated that, in the context of proactive interference during a short-term item memory task, left mid-VLPFC was part of a network of control processes resolving proactive interference (Figure 1A), and argues from the full complement of results that this region may be up-regulated to the extent that an encounter with a recent target can cue retrieval of irrelevant associated information and so produce conflict among retrieved representations. Likewise, in Chapter 3, left mid-VLPFC was active across four manipulations of control during semantic memory tasks, and was associated with an empirically derived meta-variable that indexed common variance associated with selection from competition across all the

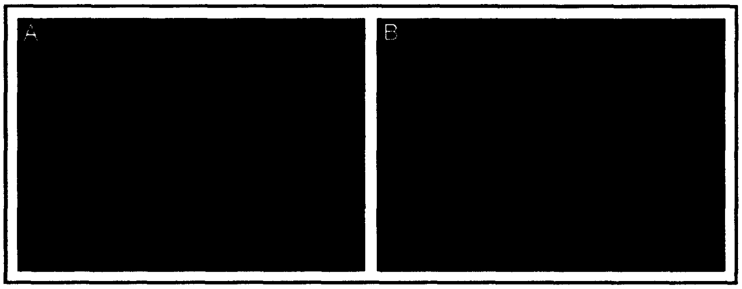

**Figure 1.** Inflated surface renderings showing the convergence between the sites in left mid-VLPFC sensitive to (A) proactive interference (Negative-Recent versus Non-Recent) from Chapter 2 and (B) the Selection-Component identified in Chapter 3.

manipulations (Figure 1B). Critically, these data dissociated this post-retrieval selection process in left mid-VLPFC from a control process that guides the act of retrieval itself. Hence, if cued retrieval of competing, irrelevant representations is a source of interference in task switching, these experiments strongly motivate the hypothesis that left mid-VLPFC will be critical in resolving this interference. The present study sought to test this hypothesis more directly.

The ability to intentionally shift between different tasks and the representations that govern action is a fundamental requirement for flexible behavior. Despite established behavioral consequences, the mechanisms underlying task switching remain controversial, being differentially conceptualized as reconfiguring the system for an upcoming task or overcoming long-term or transient interference from a previous task. This chapter describes an event-related fMRI experiment (currently in preparation for submission) that indexed neural responses while subjects switched tasks under differing levels of preparedness and interference. Preparation-related decay of switching effects in VLPFC was accounted for by conflict among competing conceptual representations, as defined by a novel computational model of task switching that derives switch costs from changes in its long-term associative structure. These data and the associated theoretical framework provide a critical step forward in the controversy surrounding the obstacles to flexible performance and the mechanisms by which these obstacles are overcome.

## Introduction

The course of modern life is often interrupted by demands that do not await our disposition but that must be addressed immediately. Though these bids on our cognitive resources are irritating and, indeed, are often obstacles to achieving goals, the capability of our cognitive system to reconfigure itself to meet shifting task demands is evident and remarkable. A fundamental problem in the study of cognitive control is specification of the psychological and neural processes by which we achieve this flexibility. An important instance of flexible behavior, that has garnered considerable interest, is task switching.

Task switching can be studied by comparing episodes in which subjects *switch* between two simple tasks (e.g., odd/even or vowel/consonant judgments about stimuli) to those in which they *repeat* the same task (Allport & Wylie, 2000; Jersild, 1927; Logan, 2003; Monsell, 2003). In such comparisons, task switching incurs a slowing or cost in response time (RT). Though repeatedly observed, the nature of this RT switch cost remains highly controversial, as it does not directly reflect the time required to reconfigure the system for a new task (Allport, Styles, & Hsieh, 1994; Rogers & Monsell, 1995). Introduction of a preparation interval prior to presentation of a target does not extinguish switch costs even when subjects are aware of a forthcoming switch. Rather, following an initial decline, a residual switch cost persists even at very long preparation intervals (Dejong, 2000; Meiran, 2000; Rogers & Monsell, 1995).

Three alternative theoretical positions have framed the debate over the interpretation of task switch costs, and hence the processes by which task switching is achieved. The *task-set reconfiguration* (TSR) hypothesis interprets the residual switch cost as reflective of a set of exogenous reconfiguration processes that require input of the target stimulus to reach convergence (Rogers & Monsell, 1995). These processes are distinct from endogenous reconfiguration processes that can initiate reconfiguration independent of the presentation of a target, and so contribute to the initial reduction in switch cost with preparation.

By contrast, the *task-set inertia* (TSI) hypothesis proposes that switch costs are substantially or wholly attributable to interference arising from transient carry-over of residual activation from a recently performed task during a task switch (Allport et al., 1994; Gilbert & Shallice, 2002; Yeung & Monsell, 2003). From this perspective, preparation-related decay in switch costs is attributable to the diminishing influence of the competing task over time.

Finally, the *task-set priming* (TSP) hypothesis has emphasized the contribution of long-term proactive interference in task switch costs (Allport & Wylie, 2000; Koch et al., 2005; Waszak et al., 2003; Wylie & Allport, 2000). Though similar to TSI in ascribing switch costs primarily to interference from an activated, competing task set, TSP proposes that this interference arises due to encounter with cues that activate or retrieve competing task representations from long-term associative memory (Allport & Wylie, 2000; Waszak et al., 2003). From this perspective, performance of a given task primes task associations among available cues and weakens associations with competing tasks; a case potentially analogous to that proposed for forgetting in declarative memory (e.g., M. C. Anderson, Bjork, & Bjork, 1994). Consequently, subsequent encounter with these cues in the context of a new task facilitates retrieval of irrelevant task information and impairs retrieval of task-relevant information. Though TSR, TSI, and TSP find behavioral evidence in their support, to date, behavioral data resolving this controversy remain equivocal (Allport & Wylie, 2000; Logan, 2003; Monsell, 2003).

Central to the TSR, TSI, and TSP hypotheses is activation of a task set from memory (Allport & Wylie, 2000; Mayr & Kliegl, 2000; Rubinstein, Meyer, & Evans, 2001; Waszak et al., 2003), alternatively emphasizing intentional, controlled access to task-relevant representations or interference arising from active long-term memory associations. Thus, leverage on the TSR/TSI/TSP debate may emerge through consideration of VLPFC contributions to task switching. Outside of the context of task switching, VLPFC has been associated with retrieval and selection of task-relevant long-term memory representations (Badre & Wagner, 2002; Petersen, Fox, Posner, Mintun, & Raichle, 1988; Poldrack et al., 1999; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997; Thompson-Schill, D'Esposito, & Kan, 1999; Wagner, Paré-Blagoev, Clark, & Poldrack, 2001). Of particular note, mid-VLPFC (BA 45; left inferior frontal gyrus *pars triangularis*) has been associated with resolution of proactive interference (Badre & Wagner, in press; Jonides, Smith, Marshuetz, Koeppe, & Reuter-Lorenz, 1998), overcoming interference from competing semantic representations (Thompson-Schill et al., 1997, also see Chapter 3), and overcoming interference due to irrelevant primed associations (Fletcher, Shallice, & Dolan, 2000; Thompson-Schill et al., 1999). Hence, it is reasonable to hypothesize a similar role for VLPFC in regulating memory during a task switch, particularly under conditions of mnemonic interference such as those proposed by TSI and TSP.

Consistent with this hypothesis, neuroimaging studies have revealed a replicable fronto-parietal network engaged during task switching that includes VLPFC, in addition to supplementary motor area (SMA), and inferior/superior parietal cortex (Brass & von Cramon, 2002, 2004a, 2004b; Braver, Reynolds, & Donaldson, 2003; Dove, Pollmann, Schubert, Wiggins, & von Cramon, 2000; Dreher & Grafman, 2003; Dreher, Koechlin, Ali, & Grafman, 2002; Meyer et al., 1998; Meyer et al., 1997; Reynolds, Donaldson, Wagner, & Braver, 2004; Rushworth, Hadland, Paus, & Sipila, 2002). Imaging studies attempting to identify endogenous processes, by manipulating a subject's foreknowledge of an upcoming switch (but not preparation time, *per se*), have observed switch-related foreknowledge effects in VLPFC (Brass & von Cramon, 2002, 2004a; Luks, Simpson, Feiwell, & Miller, 2002; Sohn, Ursu, Anderson, Stenger, & Carter, 2000). However, studies including interference manipulations have also demonstrated sensitivity to task switch interference in VLPFC (Brass & von Cramon, 2004b; Dreher & Berman, 2002) ; Fig. 2). Hence, activation in VLPFC is modulated by foreknowledge, putatively attributable to preparation-related endogenous reconfiguration processes, and is also modulated by task-level interference, potentially serving to overcome the impact of TSI/TSP. These results raise the possibility that specification of VLPFC contributions to task switching will advance understanding of the mechanisms underlying task switching more generally. More specifically, when considered in the context of the broader literature on left VLPFC function, these results motivate the hypothesis that VLPFC is engaged to overcome interference between competing representations retrieved during task switching.

The present study contributes important new evidence that may serve to resolve the TSR/TSI/TSP controversy, focusing on characterization of the impact of preparation time and proactive interference on VLPFC activation during task switching. To provide an explicit theoretical context, a novel computational model, termed control of associative memory during task switching (CAM-TS), was specified and evaluated based on behavioral results from Experiment 1. An index of mnemonic control demands, computed as the level of conflict derived from the model, then served as an explicit theoretical context for understanding the results from the fMRI experiment described in Experiment 2, and provided quantitative predictions for the patterns of fMRI response. Experiment 2 was designed to assess the impact of two central manipulations on regions engaged during task switching: (1) A manipulation of preparation time permitted estimation of time-dependent changes in switch effects; and (2) a manipulation of task-level proactive interference permitted direct characterization of the modulation of observed switching effects due to interference from a prior task. These two factors served to dissociate the pattern of response in VLPFC from that observed in other switch-related regions, and to relate this response directly to the quantitative pattern of conflict simulated by CAM-TS.

## Experiment 1 – Behavioral characterization of preparation independent of decay

Experiment 1 provided data on which to specify and evaluate the computational model of task switching. The primary goal of Experiment 1 was to characterize the change in RT switch cost due to preparation, independent of transient decay. On each trial, subjects (N = 24) were explicitly cued (200ms) as to which task (vowel/consonant or odd/even judgment) was to be performed on the subsequently presented stimulus, with half of the trials requiring a task *Switch* and half entailing a task *Repeat*. As subjects had foreknowledge on every trial, preparedness was manipulated by varying CSI from 50ms to 950ms (plus 200ms for cue presentation) thereby impacting the opportunity for preparatory processing (greater with increasing CSI). Critically, time from the response on the previous trial until presentation of the cue (RCI) was also varied from 50ms to 950ms, thus decoupling decay time from preparation time. To provide an empirical basis for assessing simulated effects of proactive interference on switch costs, interference due to response repetition (RR) was directly manipulated across all CSI and RSI combinations. RR across a task switch has been consistently shown to increase RT switch costs relative to emitting a different motor response (RD) (Meiran, Chorev, & Sapir, 2000; Rogers & Monsell, 1995).

## Methods

### *Subjects*

Twenty-four right-handed, native English speakers (16 female; ages 18-25 yrs) were remunerated $10/hour for participation. Informed consent was obtained as approved by the Administrative Panel on Human Subjects in Non-medical Research at Stanford University.

*Design*

Stimuli consisted of number–letter pairs (e.g. "2b") presented centrally in 32-point Monaco font. Pairs were constructed from a set of 10 letters, consisting of five consonants ('p','f','n','k','s') and five vowels ('a','e','i','o','u'), and a set of 10 digits, consisting of five odd numbers ('1','3','5','7','9') and five even numbers ('0','2','4','6','8'). The spatial positions of the number and letter were counterbalanced across pairs (e.g., “2b” or “b2”).

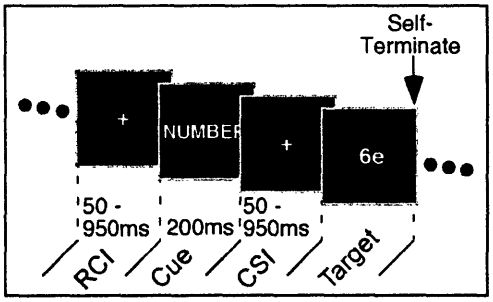

**Figure 2.** Schematic depicting events during a trial of Experiment 1. The pre-target portion of the trial began with a variable RCI, over which only passive decay could occur, followed by a task cue (LETTER or NUMBER), then a CSI during which active preparation could also occur. Then a number–letter target was presented until the subject made their response, after which the trial was terminated.

Subjects performed one of two simple categorization tasks (Figure 2) with each stimulus pair. In the Number task, subjects categorized the number as odd or even by pressing the Left or Right button on the computer keyboard under their right hand. In the Letter task, subjects categorized the letter as vowel or consonant, again pressing the Left or Right button on the computer keyboard.

During each block of the main experiment, the Number and Letter tasks were intermixed. An instruction cue (LETTER or NUMBER) preceded the onset of each target stimulus (Figure 2) and signaled the task to be performed for that target. The task cued for the upcoming target could entail a task Repeat or a task Switch. Once presented, the target stimulus remained on the screen until a response was made, upon which the trial terminated. The response cued by the target could be the same (Response Repetition, RR) or different (Response Different; RD) than the response emitted on the previous trial. The interval between the previous trial response and the task cue (RCI) was varied on a trial-to-trial basis among four values (50 ms, 226 ms, 506 ms, 950 ms) that expanded logarithmically. Likewise, the CSI varied among the same four values (50 ms, 226 ms, 506 ms, 950 ms). To maximize our ability to locate switch related declines (Rogers & Monsell, 1995), CSI was blocked, and the order of blocks was fully counterbalanced between the 24 subjects.

Each CSI-defined block consisted of 256 trials divided evenly among all remaining experimental conditions, plus 4 warm-up trials at the beginning of each block that were excluded from analysis. In addition to experimental conditions, trials were counterbalanced for (a) the match or mismatch of the response cued by the irrelevant flanking stimulus to the correct response, and (b) whether the position of the correct target switched from the previous trial.

### *Procedure*

All behavioral testing was conducted on a Macintosh G4 computer in a darkened testing room at Stanford University.

Prior to data collection, subjects received extensive training on the two tasks. After receiving instructions on the task, subjects continuously practiced one of the tasks (e.g., Letter or Number) by itself and then practiced the other task (e.g., Number or Letter) by itself. Subsequently, subjects practiced switching between the tasks in four short blocks of practice at each CSI used in the main experiment. To allow subjects to familiarize themselves with the CSIs used in the actual experiment, the practice blocks were performed in the same order used in the main experiment.

## Results

Task switching resulted in a median RT cost ($F(1,23) = 30.6$, $p < .0001$). Furthermore, switch costs were greater on RR trials than RD trials (Task Switch [TS] × RR: $F(1,23) = 29.9$, $p < .0001$). Critically, RT costs declined with CSI when collapsed across RCI intervals (see Figure 3B; $F(3,69) = 5.4$, $p < .005$). There was no interaction of this CSI-based decline in switch costs with RCI ($F(9,207) = 1.5$). By contrast, beyond a small quantitative decline evident at the shortest CSI, a general decline in switch cost with RCI when collapsed across CSI was not reliable ($F(3,69) = 1.4$). Hence, changes in RT costs deriving from the manipulation of CSI, in Experiment 1, are primarily reflective of preparation-related declines rather than other contributors such as passive decay, and so these patterns provide a robust source of data on which to base the model.

## The Control of Associative Memory during Task Switching

The architecture and dynamics of CAM-TS (Figure 3A, detailed in Chapter 5), share a number of basic features with other models of cognitive control (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Cohen, Dunbar, & McClelland, 1990; Gilbert & Shallice, 2002; Yeung, Cohen, & Botvinick, 2004). However, CAM-TS differs significantly from previous approaches to task switching in its exclusive reliance on small performance-dependent changes in association strengths between nodes (i.e., connection weights) to produce task switching effects, rather than other features such as short-term transient carry-over (Gilbert & Shallice, 2002). In this sense, the model most closely follows the assumptions of TSP.

CAM-TS consists of three layers of nodes corresponding to the Task (Letter and Number), Concept (Odd, Even, Vowel, and Consonant), and Response (Left and Right) alternatives in the explicit cueing task. Feedforward connections proceed from the Task layer to relevant nodes in the Concept layer, and from the Concept layer to the appropriate nodes in the Response layer. Furthermore, reciprocal feedback connections loop from nodes in subordinate layers (i.e., Response and Concept) back to superordinate layers (i.e. Concept and Task). Nodes within a layer compete via mutual inhibitory connections.

Task switching costs emerge from two features of the model. (1) The feedback connections between layers enable activated Response and Concept nodes to elicit activation in task-irrelevant nodes and so produce competition. (2) Following each trial, the baseline connection weights in the Task layer change based on a simple learning rule (Gilbert & Shallice, 2002), such that the weights among convergently activated nodes increase and among divergently activated nodes decrease prior to the next trial. On a Switch trial, irrelevant associations are stronger and relevant associations are weaker relative to a Repeat trial. The result is interference in the form of enhanced conflict in the Conceptual layer (i.e., multiple activated competing nodes) and so a switch cost.

CAM-TS also differs from previous models (Gilbert & Shallice, 2002) in the manner by which preparatory control during a task switch is enacted to overcome interference in the Conceptual layer. Preparatory task control is implemented by the top-down influence of the Task layer on the Concept layer, activating the relevant conceptual nodes prior to the presentation of the target stimulus. Nodes in the Task layer may be activated by external input, an instance of external or stimulus control, or even by feedback connections from the Concept and Response layers. Hence, the preparatory application of endogenous control in the Task layer is implemented through the modulation of a gain term (e.g., Cohen & Servan-Schreiber, 1992) that up-regulates the top-down influence of this layer on the Concept layer during the preparatory interval of a switch trial. Increasing this top-down influence for a short, fixed duration (3 cycles) during a preparation interval of a switch trial serves to overcome switch-related conflict in the Conceptual layer. Over increasing CSI, this results in a decrease in the switch cost.

The strong fit of the simulated task switching RT effects from CAM-TS with the results from Experiment 1 are depicted in Figure 3B and C below. The simulated decline in RT switch cost closely matched the decline in the RT data (R = .99). Furthermore, when divided based on RR and RD trials, CAM-TS closely simulated the impact of proactive interference across different preparation intervals (R = .97).

Of central importance, the model may perform task switches successfully and obtains a switch cost without up-regulating endogenous control differentially during a task switch. However, without up-regulation of control prior to a Switch there is no CSI-dependent decline in simulated switch costs (Figure 3B) and the fit of the model is relatively poor (R = .486). To produce CSI-dependent declines, CAM-TS requires a control process that increases the bias on the Conceptual layer during the preparation interval of a task switch.

As stated above, feedback connections and performance-dependent changes in association weights increase conflict in the model during a Switch which produces task switch costs. This conflict is due to activation being spread over more competing nodes in a layer and likewise the diminished activation of relevant nodes. A quantifiable index of this conflict, emerging from the Conceptual and Response layers, may be computed as Hopfield energy (Botvinick et al., 2001; Hopfield, 1982) (Figure 3A). The computational properties of Hopfield energy correspond to desirable conceptual features of conflict,

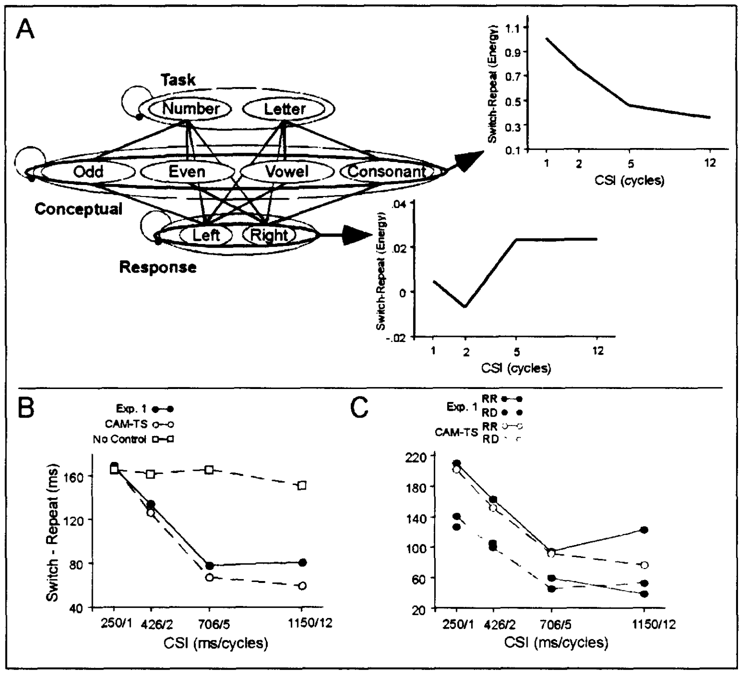

**Figure 3.** Results from Experiment 1 and CAM-TS simulations. (A) CAM-TS a connectionist architecture consisting of Task, Conceptual, and Response layers from which task switch costs emerge due to performance-based changes in its connection weights. The Switch versus Repeat difference in conflict computed from the model's Conceptual layer (Red) declines over CSI, whereas the difference in conflict computed from the Response layer (Blue) roughly increases over increasing CSI. (B) Declines in switch RT cost from Experiment 1 were modeled well by CAM-TS when control was enacted during the preparation interval. (C) The model also simulated the decline in RT cost from Experiment 1 for both RR and RD trials.

both in the model and more generally (Berlyne, 1957), in that energy increases exponentially with numbers of nodes active, and so competing, within a layer and with the strength of their mutual activation levels (Botvinick et al., 2001).

As plotted in Figure 3A, enhanced conflict in the Conceptual layer during a task switch, as indexed by energy, declined across CSI intervals. By contrast, the Response layer showed a roughly increasing conflict function over increasing CSI, potentially reflective of the accumulation of evidence in these nodes and the transfer of selection demands from Conceptual to Response levels. Interestingly, RR impacted both Concept and Response layers by modestly increasing the conflict overall, suggesting that changes in long-term associations may give rise to RR effects but that these effects may be expressed as conflict at multiple representational levels. Hence, unlike the decline in Conceptual conflict over CSI, which clearly distinguished Conceptual from Response conflict, this RR-induced modulation of conflict in the Conceptual layer was evident at multiple levels, likely indicating that RR effects emerge from both the Conceptual and Response layers of the model.

In summary, from CAM-TS, a neural control processor that receives information about conflict in the Conceptual layer during a task switch as its input should be marked by greater conflict during a switch trial that declines at longer CSIs. In addition, an enhancement in this declining conflict signal due to proactive interference from RR may be evident. The conflict signals defined by CAM-TS provided quantitative predictions to be used in the analysis of the fMRI results from Experiment 2.

## Experiment 2 - VLPFC and Control of Memory during Task Switching

An independent sample of 10 subjects were scanned using fMRI while performing a task analogous to that used in Experiment 1. Again, a CSI manipulation (250ms to 1150ms) varied preparation time, and an RR manipulation varied task-level proactive interference. Together, these factors permitted dissociation of the role of left VLPFC in task switching from other regions implicated during task switching, and further characterized left VLPFC response as sensitive to conflict among retrieved conceptual representations.

### Methods

#### *Subjects*

Thirteen right-handed, native English speakers (8 female; ages 18-25 yrs) were remunerated $50 for participation. Data from three of these subjects (1 female) were excluded prior to fMRI analysis because of high non-response rates due to a difficulty with responding prior to the response deadline. An additional subject was recruited but was not scanned due to difficulty with learning the task. Informed consent was obtained as approved by the Human Subjects Committee of Massachusetts General Hospital and the Committee on the Use of Humans as Experimental Subjects at MIT.

#### *Design*

Stimuli were identical to those used in Experiment 1. Stimulus position and the congruency of responses cued by the stimuli were counterbalanced across experimental conditions. As in Experiment 1, subjects performed one of two simple categorization tasks (Figure 4) with each stimulus pair. In the Number task, subjects categorized the number as odd or even by pressing the Left or Right button on a response pad under their left hand. In the Letter task, subjects categorized the letter as vowel or consonant, again pressing the Left or Right button on the response pad. Prior to fMRI data collection, subjects received extensive training on the two tasks: (a) outside the scanner, subjects continuously performed the Letter task and then the Number task, and (b) subsequently, subjects practiced switching between the tasks during collection of structural MR images.

During fMRI scanning, the Number and Letter tasks were intermixed. An instruction cue (LETTER or NUMBER) preceded the onset of each stimulus (Figure 4), which signaled the task to be performed for that stimulus. During the cue-to-stimulus interval (CSI) within a trial and the inter-trial-interval (ITI) separating trials, a white fixation cross was presented centrally (as a preparatory warning, the fixation cross turned from white to red immediately prior to cue/stimulus presentation). To allow for estimation and deconvolution of the hemodynamic response as a function of relatively small changes in CSI duration, trials were grouped into pairs of two task events (T-1 and T-2), though to the subject the experiment appeared as a continuous stream of task

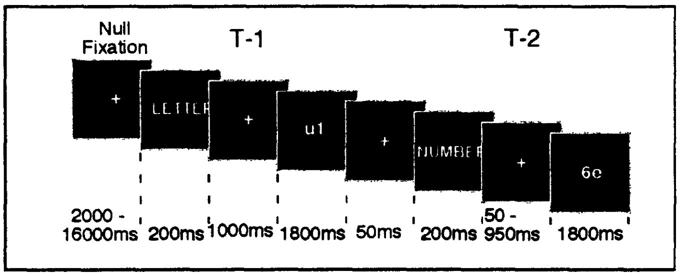

**Figure 4.** Schematic depicting events during a task event pair in Experiment 2. Task events (T-1 and T-2) consisted of a task cue (LETTER or NUMBER), fixation during the CSI, and a number–letter stimulus that required a response. Task events were grouped into sets of an initial event (T-1), during which experimental variables were held constant, and a second event (T-2), during which the experimental factors were manipulated. For fMRI analysis, each set was coded as an epoch starting at the onset of the T-1 cue; these epochs could be readily compared because, across the T-2 experimental conditions, the epoch history was identical up to presentation of the T-2 cue.

events. The onset of the first task event (T-1) in each pair followed a variable duration null fixation period (2s-16s) that intervened between the prior pair and the present pair. To permit event-related fMRI analyses, T-1 always required performance of the same task that had just been performed (Task Repeat) during the T-2 phase in the prior pair. Furthermore, the consonance of the T-1 manual response to the manual response emitted during the T-2 phase of the prior pair (RD/RR) and the target and flanker relationship within T-1 were counterbalanced across experimental conditions at T-2 (see below). The CSI duration for T-1 was fixed at 1000ms and the ITI between T-1 and the cue for the second task event in the pair (T-2) was always 50ms.

The critical experimental variables were manipulated during T-2. The task for T-2 was either the same as (Task *Repeat*) or different from (Task *Switch*) that was performed during T-1. The T-2 manual response (Left or Right) was either the same as (*RR*) or different from (*RD*) that required during T-1. Finally, the duration of the CSI (inclusive of 200ms task cue) for the T-2 event varied (250ms, 426ms, 706ms, or 1150ms). For half of the pairs, the target stimulus in T-2 was in the same position as the target stimulus in T-1; for the other half, the positions of the targets differed. Collectively, this design allowed for analysis of fMRI signal differences at T-2 according to the critical factor manipulations (Repeat vs. Switch, RR vs. RD, and CSI duration). For both T-1 and T-2 events, a response deadline of 1800ms was imposed. For the purposes of imaging analysis, a pair was considered incorrect if a subject responded incorrectly or failed to respond prior to the response deadline on the T-1 and/or T-2 task events within the pair. Analysis of RT was restricted to the trials included in the imaging analysis. However, error analysis was based only on T-2 events and was not dependent on T-1 accuracy.

During fMRI scanning, subjects encountered 480 pairs of trials across 4 scan epochs. To bolster efficacy of the CSI manipulation (Rogers & Monsell, 1995), events were grouped into blocks of 30 pairs on the basis of T-2 CSI duration. Subjects encountered a block of each CSI duration during each scan epoch; importantly, the variable duration null fixation events interposed between pairs allowed for event-related analyses. Response mappings and condition order were counterbalanced across subjects.

### *MRI Procedures*

Whole-brain imaging was performed on a 3T Siemens Trio MRI system. Functional data were acquired using a gradient-echo echo-planar pulse sequence (TR=2 s, TE=30 ms, 21 axial slices, 3.125 x 3.125 x 5 mm, 1 mm inter-slice gap, 4 runs x 756 volume acquisitions). High-resolution T1-weighted (MP-RAGE) structural images were collected for anatomical visualization, during which subjects practiced the tasks. Head motion was restricted using firm padding that surrounded the head. Projected visual stimuli were viewed through a mirror attached to the standard head coil.

Data were preprocessed using SPM99 (Wellcome Dept. of Cognitive Neurology, London). Functional images were corrected for differences in slice acquisition timing, followed by motion correction (using sinc interpolation). Structural and functional data were spatially normalized to a template based on the MNI305 stereotactic space (Cocosco, Kollokian, Kwan, & Evans, 1997) using a 12-parameter affine transformation along with a nonlinear transformation using cosine basis functions. Images were resampled to 3-mm cubic voxels and spatially smoothed with an 8-mm FWHM isotropic Gaussian kernel.

Statistical models were constructed using SPM99 under the assumptions of the general linear model. The units of analysis were the task pairs described above (Figure 4). Because the T-1 phase of each pair was counterbalanced across the T-2 conditions, the unique contribution to the overall variance due to manipulations of T-2 could be estimated. Epochs of 6 s, beginning at the onset of the T-1 cue, were used to model each pair; the 6-s epochs were convolved with a canonical HRF. Correct and incorrect trials were modeled separately, and subsequent contrasts were restricted to correct trials. Effects were estimated using a subject-specific fixed-effects model, with session-specific effects and low-frequency signal components treated as confounds. Linear contrasts were used to obtain subject-specific estimates for each effect. These estimates were entered into a second-level analysis treating subjects as a random effect, using a one-sample t-test against a contrast value of zero at each voxel. Effects in the whole brain analysis were considered reliable to the extent that they consisted of at least 5 contiguous voxels that exceeded an uncorrected threshold of $p < .001$.

The voxel-based contrasts were supplemented with region-of-interest (ROI) analyses to further characterize the effects of CSI and interference (RR vs. RD) in *a priori* expected regions, including VLPFC. ROI analyses also provided quantitative characterization of the effects observed in the voxel-based analyses. The ROI analyses were performed using a toolbox for use with SPM99 (written by R. Poldrack; http://sourceforge.net/projects/spm-toolbox/). ROIs were defined on the basis of the Switch versus Repeat contrast ($p < .001$, uncorrected). ROIs included all significant (uncorrected $p < .001$) voxels within a 6-mm radius of the chosen maximum. Selective averaging with respect to peristimulus time allowed assessment of the signal change associated with each condition, thus permitting ROI analyses based on the data rather than on the parameter estimates. ROI analyses were performed on measures of integrated percent signal change (peak ±2 TRs), which were subjected to repeated-measures analyses of variance (ANOVA). Finally, assessment of monotonic decay components over CSI in ROIs was performed by estimating a logarithmic decay for each subject individually and then entering these estimates into a second-level analysis that treated subject as a random variable, using a one-sample $t$-test against a null effect value of 0. Decay in RT cost was assessed similarly except that a linear model was used rather than a logarithmic, after residual analysis revealed that the distribution of errors deviated from normal (also see Chapter 5 – Global Preparation). Likewise, convergence between the conflict output of the computational model and fMRI signal was evaluated within-subject based on a linear predictive relationship.

## Results

As plotted in Figure 5, switching incurred a cost in median RT (F(1,9) = 82.8, $p <$ .0001). Critically, this cost declined linearly with increasing CSI (t(9) = 2.4, $p <$ .05), and a marked residual cost (105ms) was evident even after a 1150ms CSI (F(1,9) = 31.4, $p <$ .0001).

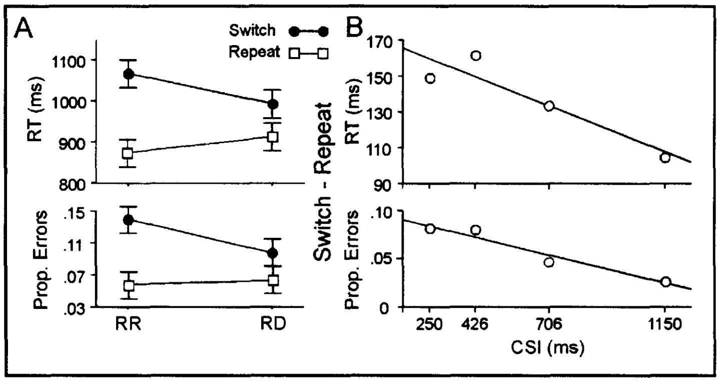

**Figure 5.** Plots of reaction time (RT) and error rate during task switching. (A) Depiction of differences between Switch (filled circle) and Repeat (open circle) as a function of Response Repetition (RR) interference versus Response Different (RD). (B) Though the linear decline in RT cost was reliable ($p <$ .05), the quantitative decline in error cost did not reach significance (t(9) = 1.9, $p =$ .08).

Whole-brain voxel-wise contrast of Switch versus Repeat trials revealed activation in anticipated regions of PFC and posterior neocortices (Figure 6). Specifically, greater activation on Switch relative to Repeat trials was observed in posterior (-48 9 27) and mid-VLPFC (-45 18 24; -54 33 18), and along the medial surface in SMA (0 18 48). Outside of PFC, Switch versus Repeat main effects were also evident in inferior (-36 –54 51; -51 –33 48) and bilateral superior parietal cortex (-27 –66 57; 21 –60 51).

Region of interest (ROI) analyses assessed the impact of preparation time on the switching effects in fMRI measures (Figure 6A). The only region to show a CSI-dependent decline in Switching effects was mid-VLPFC (-54 33 18), in that the Switch versus Repeat difference was reliably greater at the longest CSI (1150ms) relative to the shortest CSI (250ms; F(1,9) = 5.0, $p < .05$), and the decline was marginally well fit by a monotonically declining function (Fig. 6A; t(9) = 2.1, $p = .06$). By contrast, Switch versus Repeat differences were stable across CSI in SMA (t(9) = .06, $p = .95$). In inferior parietal cortex, not only was no decline evident, the quantitative pattern was the opposite, with the Switch versus Repeat difference tending to increase with longer preparation intervals, though not reliably (t(9) = 1.4, $p = .18$).

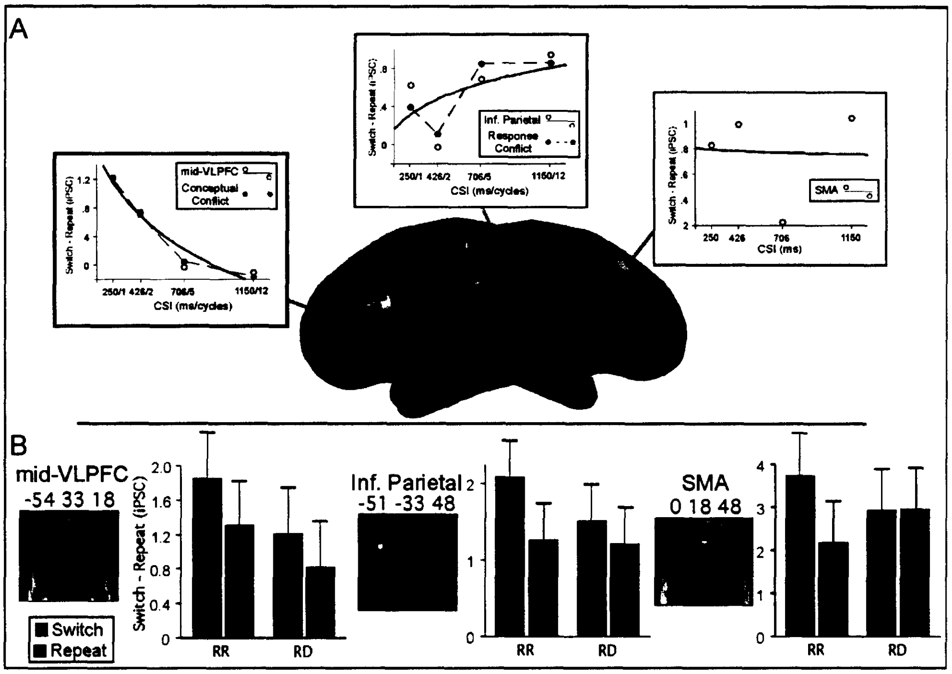

**Figure 6.** (A) Surface rendering of Switch versus Repeat on an inflated MNI canonical reveals activations in left mid-VLPFC, inferior and superior parietal cortex, and SMA on the medial surface. ROI analyses revealed a declining Switch versus Repeat difference in left mid-VLPFC (-54 33 18), an increasing difference in inferior parietal cortex (-51 -33 48), and a relatively stable response in SMA (0 18 48). The linearly scaled conflict signal from Conceptual (Red dashed line) and Response (Blue dashed line) layers is also depicted for comparison. The differential pattern of decline in left mid-VLPFC dissociated this region from inferior parietal cortex. (B) Bar graphs depicting enhancement of switch costs from RR versus RD. Whereas RR fully accounted for the Switch (Red bar) versus Repeat (Green bar) difference in SMA, RR only reliably increased activation for Switch-RR versus Switch-RD trials in left mid-VLPFC. The pattern of interference dissociated mid VLPFC from SMA.

Of central importance, the Switch versus Repeat decline in left mid-VLPFC corresponded reliably with the conflict signal defined from the Conceptual layer of CAM-TS, as assessed within subject (t(9) = 2.3, $p < .05$; Figure 6A). By contrast, the ramping response observed in inferior parietal cortex appeared to correspond to the conflict signal arising from the Response layer of CAM-TS[1] ($t(9) = 2.6$, $p < .05$; Figure 6A). Critically, the difference between inferior parietal cortex and left mid-VLPFC in CSI-dependent switch functions was verified within-subject and reliably dissociated these regions ($t(9) = 2.5$, $p < .05$).

The impact of task-level proactive interference through RR was evident in an enhanced RT switch cost (Figure 5A; TS × RR: $F(1,9) = 24.5$, $p < .001$). In fMRI measures, a TS × RR interaction was evident in SMA (Figure 6B; $F(1,9) = 51.2$, $p < .05$). Indeed, the Switch versus Repeat main effect in SMA was entirely accounted for by RR interference, as Switch was only greater than Repeat on RR trials ($F(1,9) = 10.1$, $p < .05$), with no statistical effect present on RD trials ($F(1,9) = .003$, $p = .99$).

In contrast to SMA, the quantitative TS × RR interaction across all CSIs (Figure 6B) was not reliable in mid-VLPFC ($F(1,9) = .16$, $p = .69$). However, indicative of RR interference and consistent with the Conceptual conflict signal in CAM-TS, planned contrasts revealed that activation in mid-VLPFC was reliably greater for Switch-RR than Switch-RD trials ($F(1,9) = 5.6$, $p < .05$). Furthermore, there was a marginal CSI-dependent decline only in the RR switch cost in mid-VLPFC ($t(9) = 2.1$, $p = .06$). However, this should not be interpreted as indicating that the RR condition fully accounted for the switch decline in mid-VLPFC, as there was also a quantitative, though not reliable ($t(9) = 1.4$), decline in Switching effects on RD trials. Critically, this

---

[1] It should be noted that the fit of the reported simulation to the specific pattern of fMRI response in parietal cortex (inclusive of the initial dip at the CSI of 426 ms) proved to be variable in additional simulation runs. Hence, though the Response conflict signal emerging from the model always increased over CSI across all simulations, and so matches the ramping response of parietal cortex qualitatively, variability in this signal was such that it was not always as quantitatively characteristic of the exact parietal response.

declining sensitivity to RR dissociated mid-VLPFC from SMA, evident in a reliable Region [mid-VLPFC/SMA] × RR × TS interaction (F(1,9) = 7.2, $p < .05$) and a marginal Region × CSI × TS interaction (F(3,9) = 2.8, $p = .06$).

## Conclusions

Task switching is fundamentally an act of memory. Consequently, the behavioral and psychological consequences of task switching may be understood in terms of the structures, processes, and failures of memory. This proposition entails that control processes contributing to task switching may be indistinguishable from the control processes engaged to overcome interference arising during any other act of memory.

The results and computational framework introduced here strongly support these conclusions. In particular, we demonstrate that a simple connectionist model that derives its task switching cost entirely from performance-dependent changes in its associative structure is capable of accounting for preparation-related declines in switch costs by increasing top-down control during the preparation interval. This model further defined a signature of declining conflict among activated concepts during longer preparation intervals prior to a task switch. Critically, this signature was characteristic of activity in left mid-VLPFC and dissociated this region from other regions active during task switching. Indeed, parietal cortex was associated with a ramping response over CSI, a pattern consistent with the increased conflict in the Response layer of CAM-TS at longer CSIs. This pattern might be broadly consistent with perspectives on the role of inferior parietal cortex during response selection (Bunge, Hazeltine, Scanlon, Rosen, & Gabrieli, 2002; Schumacher & D'Esposito, 2002). Indeed, a shift over time in selection demands from competition among retrieved conceptual representations mediated by left VLPFC to competition among responses or response mappings mediated by parietal cortex may also be consistent with results showing the global timecourse for task switching to be characterized by two separable temporal components: an early frontal component (~300-500ms) followed by a subsequent parietal component (~500-1000ms; Karayanidis, Coltheart, Michie, & Murphy, 2003; Lorist et al., 2000; Rushworth, Passingham, & Nobre, 2002); though a strong connection between these effects awaits greater anatomical precision potentially afforded by source constrained data.

The present results provide a reconciliation to the long-standing debate over the origins of task switch costs and the control processes that engage in prospectively switching tasks. In particular, the model presented here suggests that a substantial portion of switch costs can, indeed, be accounted for by long-term carry-over due to retrieval/activation of a primed, competing task set, consistent with TSP. However, analogous to more traditional proactive interference effects, control processes that regulate memory may overcome this interference and resolve competition from irrelevant retrieved representations, and may do so prospectively during the preparation interval by biasing relevant retrieved representations. Outside of the context of task switching, these interference resolution processes have been associated with neural computations in mid-VLPFC (Badre & Wagner, in press; Jonides et al., 1998; Thompson-Schill et al., 1999). The present results suggest that task switching is not an exception to this pattern and directly link preparation in task switching to a mnemonic control process in left mid-VLPFC that is sensitive to conflict among conceptual representations.

As modeled by CAM-TS, an active control process in mid-VLPFC may work to resolve interference during a task switch by relying on a maintained representation of a task goal to bias task-relevant, retrieved conceptual representations. As such, this process may be particularly sensitive to increases in conflict (i.e. selection demands) among active conceptual representations. In its current implementation, this process acts somewhat like a "homunculus", knowing when to up-regulate control during a task switch. However, without this control, conceptual conflict in CAM-TS is greatly enhanced, and differentially so for Switch trials. Hence, it is conceivable, that left mid-VLPFC monitors the conceptual layer for conflict and up-regulates control accordingly (Botvinick et al., 2001), or alternatively, through experience the system develops procedural propositions to up-regulate this control prior to a task switch (J. R. Anderson et al., 2004). Thus, in line with TSR, control processes in mid-VLPFC may come on-line prior to presentation of a target, and in effect, prospectively reconfigure the system by biasing the relevant task representations over any competing retrieved representations.

Critically, CAM-TS shows that task switching, at least in this explicit cueing variant, may be able to proceed without this control process (perhaps similar to arguments of Logan & Bundesen, 2003, 2004), but will be more vulnerable to interference. Consistent with this latter point, damage inclusive of left mid-VLPFC does not prohibit performance of task switching (Rogers et al., 1998) nor even simple short-term item recognition under general conditions of proactive interference (Thompson-Schill et al., 2002). Rather, task switch costs and proactive interference effects are enhanced relative to controls.

Finally, the present work does not preclude the participation of other control processes or interference effects during task switching. For example, a process of goal setting is likely required in many task switching contexts (e.g., Rubinstein et al., 2001), and indeed, may contribute to the system's ability to determine when to up-regulate mid-VLPFC control mechanisms. Such demands may have been rendered constant and minimal by the explicit cueing procedure used here, but such a process is likely integral to task switching, as it is in other contexts. Indeed, goal setting may be associated with distinct regions of prefrontal cortex, such as frontal polar cortex (Braver et al., 2003), during task switching. Furthermore, additional factors contributing to switch costs, such as transient carry-over proposed by TSP and intentionally excluded from the present theoretical framework, may also elicit additional compensatory processes. Notably, the distinct pattern in SMA was not directly accounted for under the current theoretical framework, and likely reflects important additional processes in task switching (Rushworth, Hadland et al., 2002). Moreover, the emergent properties of CAM-TS and the imaging results suggest that though behavioral RR effects may partially emerge from conceptual conflict and find their source in performance-dependent changes in long-term associations, they may also arise from conflict contributed from other layers, such as the Response layer. The downstream effects of these multiple sources of conflict might converge in processors like SMA prior to generation of a response. The contribution of these additional components to task switching and their interaction with the mnemonic control process characterized here await direct consideration.

## References

Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and Performance XV: Conscious and nonconscious information processing* (pp. 421-452). Cambridge, MA: MIT Press.

Allport, A., & Wylie, G. (2000). 'Task-switching', stimulus-response bindings and negative priming. In S. Monsell & J. Driver (Eds.), *Control of Cognitive Processes: Attention and Performance XVIII* (pp. 35-70). Cambridge, MA: MIT Press.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychol Rev, 111*(4), 1036-1060.

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(1063-1087).

Badre, D., & Wagner, A. D. (2002). Semantic retrieval, mnemonic control, and prefrontal cortex. *Behavioral and Cognitive Neuroscience Reviews, 1*(3), 206-218.

Badre, D., & Wagner, A. D. (in press). Frontal lobe mechanisms that resolve proactive interference. *Cereb Cortex.*

Berlyne, D. E. (1957). Uncertainty and conflict: A point of contact between information-theory and behavior-theory concepts. *Psychological Review, 64*, 329-339.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol Rev, 108*(3), 624-652.

Brass, M., & von Cramon, D. Y. (2002). The role of the frontal cortex in task preparation. *Cereb Cortex, 12*(9), 908-914.

Brass, M., & von Cramon, D. Y. (2004a). Decomposing components of task preparation with functional magnetic resonance imaging. *J Cogn Neurosci, 16*(4), 609-620.

Brass, M., & von Cramon, D. Y. (2004b). Selection for cognitive control: a functional magnetic resonance imaging study on the selection of task-relevant information. *J Neurosci, 24*(40), 8847-8852.

Braver, T. S., Reynolds, J. R., & Donaldson, D. I. (2003). Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron, 39*(4), 713-726.

Bunge, S. A., Hazeltine, E., Scanlon, M. D., Rosen, A. C., & Gabrieli, J. D. (2002). Dissociable contributions of prefrontal and parietal cortices to response selection. *Neuroimage, 17*(3), 1562-1571.

Cocosco, C. A., Kollokian, V., Kwan, R. K. S., & Evans, A. C. (1997). Brainweb: Online interface to a 3D MRI simulated brain database. *NeuroImage, 5*, 425.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97*, 332-361.

Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: A connectionist approach to behavior and biology in schizophrenia. *Psychological Review, 99*, 45-77.

Dejong, R. (2000). An intention-activation account of residual switch costs. In S. Monsell & J. Driver (Eds.), *Control of Cognitive Processes: Attention and Performance XVIII* (pp. 357-376). Cambridge, MA: MIT Press.

Dove, A., Pollmann, S., Schubert, T., Wiggins, C. J., & von Cramon, D. Y. (2000). Prefrontal cortex activation in task switching: an event-related fMRI study. *Brain Res Cogn Brain Res, 9*(1), 103-109.

Dreher, J.-C., & Grafman, J. (2003). Dissociating the roles of the rostral anterior cingulate and the lateral prefrontal cortices in performing two tasks simultaneously or successively. *Cerebral Cortex, 13*, 329-339.

Dreher, J. C., & Berman, K. F. (2002). Fractionating the neural substrate of cognitive control processes. *Proc Natl Acad Sci U S A, 99*(22), 14595-14600.

Dreher, J. C., Koechlin, E., Ali, S. O., & Grafman, J. (2002). The roles of timing and task order during task switching. *Neuroimage, 17*(1), 95-109.

Fletcher, P. C., Shallice, T., & Dolan, R. J. (2000). "Sculpting the response space"--an account of left prefrontal activation at encoding. *Neuroimage, 12*(4), 404-417.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-B., Frith, C. D., & Frackowiak, R. S. J. (1995). Statistical parametric maps in functional imaging: A general linear approach. *Hum Brain Mapp, 2*, 189-210.

Gilbert, S. J., & Shallice, T. (2002). Task switching: a PDP model. *Cognit Psychol, 44*(3), 297-337.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc Natl Acad Sci U S A, 79*, 2554-2558.

Jersild, A. T. (1927). Mental set and shift. *Archives of Psychology*(Whole No. 89).

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

Karayanidis, F., Coltheart, M., Michie, P. T., & Murphy, K. (2003). Electrophysiological correlates of anticipatory and poststimulus components of task switching. *Psychophysiology, 40*(3), 329-348.

Koch, I., Prinz, W., & Allport, A. (2005). Involuntary retrieval in alphabet-arithmetic tasks: Task-mixing and task-switching costs. *Psychol Res, 69*(4), 252-261.

Logan, G. D. (2003). Executive control of thought and action: In search of the wild homunculus. *Current Directions in Psychological Science, 12*(2), 45-48.

Logan, G. D., & Bundesen, C. (2003). Clever homunculus: is there an endogenous act of control in the explicit task-cuing procedure? *J Exp Psychol Hum Percept Perform, 29*(3), 575-599.

Logan, G. D., & Bundesen, C. (2004). Very clever homunculus: compound stimulus strategies for the explicit task-cuing procedure. *Psychon Bull Rev, 11*(5), 832-840.

Lorist, M. M., Klein, M., Nieuwenhuis, S., De Jong, R., Mulder, G., & Meijman, T. F. (2000). Mental fatigue and task control: planning and preparation. *Psychophysiology, 37*(5), 614-625.

Luks, T. L., Simpson, G. V., Feiwell, R. J., & Miller, W. L. (2002). Evidence for anterior cingulate cortex involvement in monitoring preparatory attentional set. *Neuroimage, 17*(2), 792-802.

Mayr, U., & Kliegl, R. (2000). Task-set switching and long-term memory retrieval. *J Exp Psychol Learn Mem Cogn, 26*(5), 1124-1140.

Meiran, N. (2000). Modeling cognitive control in task-switching. *Psychological Research, 63*, 234-249.

Meiran, N., Chorev, Z., & Sapir, A. (2000). Component processes in task switching. *Cognit Psychol, 41*(3), 211-253.

Meyer, D. E., Evans, J. E., Lauber, E. J., Gmeindl, L., Rubinstein, J., Junck, L., et al. (1998, April). *The role of dorsolateral prefrontal cortex for executive cognitive processes in task switching*. Paper presented at the Cognitive Neuroscience Society, San Francisco, CA.

Meyer, D. E., Lauber, E. J., Rubinstein, J., Gmeindl, L., Junck, L., & Koeppe, R. A. (1997, March, 23-25). *Activation of brain mechanisms for executive mental processes in cognitive task switching*. Paper presented at the Cognitive Neuroscience Society, Boston, MA.

Monsell, S. (2003). Task switching. *Trends Cogn Sci, 7*(3), 134-140.

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single- word processing. *Nature, 331*(6157), 585-589.

Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage, 10*(1), 15-35.

Reynolds, J. R., Donaldson, D. I., Wagner, A. D., & Braver, T. S. (2004). Item- and task-level processes in the left inferior prefrontal cortex: positive and negative correlates of encoding. *Neuroimage, 21*(4), 1472-1483.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124*(2), 207-231.

Rogers, R. D., Sahakian, B. J., Hodges, J. R., Polkey, C. E., Kennard, C., & Robbins, T. W. (1998). Dissociating executive mechanisms of task control following frontal lobe damage and Parkinson's disease. *Brain, 121 ( Pt 5)*, 815-842.

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *J Exp Psychol Hum Percept Perform, 27*(4), 763-797.

Rushworth, M. F., Hadland, K. A., Paus, T., & Sipila, P. K. (2002). Role of the human medial frontal cortex in task switching: a combined fMRI and TMS study. *J Neurophysiol, 87*(5), 2577-2592.

Rushworth, M. F., Passingham, R. E., & Nobre, A. C. (2002). Components of switching intentional set. *J Cogn Neurosci, 14*(8), 1139-1150.

Schumacher, E. H., & D'Esposito, M. (2002). Neural implementation of response selection in humans as revealed by localizing effects of stimulus-response compatibility on brain activation. *Human Brain Mapping, 17*, 193-201.

Sohn, M. H., Ursu, S., Anderson, J. R., Stenger, V. A., & Carter, C. S. (2000). Inaugural article: the role of prefrontal cortex and posterior parietal cortex in task switching. *Proc Natl Acad Sci U S A, 97*(24), 13448-13453.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proc Natl Acad Sci U S A, 94*(26), 14792-14797.

Thompson-Schill, S. L., D'Esposito, M., & Kan, I. P. (1999). Effects of repetition and competition on activity in left prefrontal cortex during word generation. *Neuron, 23*(3), 513-522.

Thompson-Schill, S. L., Jonides, J., Marshuetz, C., Smith, E. E., D'Esposito, M., Kan, I. P., et al. (2002). Effects of frontal lobe damage on interference effects in working memory. *Cogn Affect Behav Neurosci, 2*(2), 109-120.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage, 15*(1), 273-289.

Wagner, A. D., Paré-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning: Left prefrontal cortex guides controlled semantic retrieval. *Neuron, 31*, 329-338.

Waszak, F., Hommel, B., & Allport, A. (2003). Task-switching and long-term priming: role of episodic stimulus-task bindings in task-shift costs. *Cognit Psychol, 46*(4), 361-413.

Wylie, G., & Allport, A. (2000). Task switching and the measurement of "switch costs". *Psychol Res, 63*(3-4), 212-233.

Yeung, N., Cohen, J. D., & Botvinick, M. M. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychol Rev, 111*(4), 931-959.

Yeung, N., & Monsell, S. (2003). The effects of recent practice on task switching. *J Exp Psychol Hum Percept Perform, 29*(5), 919-936.

# Chapter 5

## The Control of Associative Memory during Task Switching Model

Chapter 4 introduced a novel computational model, termed the Control of Associative Memory during Task Switching (CAM-TS) model, that derives a task switching cost from performance-dependent changes in its associative structure, and that obtains preparation-related declines through a control mechanism that resolves conflict in the model's Conceptual layer by up-regulating top-down input to this layer during the preparatory interval of a switch trial. In addition to demonstrating the capability of the model to simulate preparation-related declines in task switching costs, a conflict signal computed from the model's conceptual layer was characteristic of the fMRI response measured from left mid-VLPFC, linking this region to a control process sensitive to conflict among retrieved, competing conceptual representations.

In this chapter, CAM-TS is considered in greater depth, making explicit all of the assumptions, parameters, and dynamics that underlie its implementation. Following a full specification of the architecture of the model and its dynamics, a series of simulations are described that demonstrate the capability of CAM-TS to account for a number of phenomena in the task switching literature. Also emerging from these simulations are phenomena not accounted for by CAM-TS, such as a process of goal setting. Finally, the relationship between CAM-TS and three previous formal models of task switching is considered. In sum, this in-depth treatment of CAM-TS seeks to situate the model firmly in the theoretical landscape of task switching and serve as the basis for future applications of the model.

## Details of CAM-TS

CAM-TS was designed to model task switching during the explicit cueing task based on associative memory theories of task switching, such as TSP (Allport & Wylie, 2000). As such, other factors, like short-term transient carry-over, or a time-consuming reconfiguration process are not implemented in the model. The critical assumptions of TSP and similar associative memory theories captured by this model are (1) that encountering stimuli associated with a competing task during performance of a given task may cue retrieval of competitive information and (2) recent performance of a given task may facilitate or prime these associative links, giving rise to enhanced competition during a task switch. It is a further goal of this model to be explicit about the presence and impact of control on resolution of an associative memory-based source of task switching.

CAM-TS implements task switching using established mechanisms of connectionist modeling (McClelland & Rumelhart, 1981), and further many of the computational features of this model are adapted or taken directly from other established bias competition models of cognitive control (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Cohen, Dunbar, & McClelland, 1990; Cohen & Servan-Schreiber, 1992; Gilbert & Shallice, 2002; Yeung, Cohen, & Botvinick, 2004). Hence, most fundamental computational assumptions in CAM-TS are those implicit and explicit in any parallel distributed processing system, and have been fully evaluated elsewhere.

*Architecture of CAM-TS*

CAM-TS is implemented as a simple parallel distributed processing network (Figure 1) (McClelland & Rumelhart, 1981) that consists of three layers of nodes corresponding to the Task (Letter and Number), Concept (Odd, Even, Vowel, and Consonant), and Response (Left and Right) representations in the explicit cueing task (e.g., Methods, Chapter 4). Feedforward connections proceed from the Task layer to relevant nodes in the Concept layer, and from the Concept layer to the appropriate nodes in the Response layer. Connection strengths are determined by a set of weights (*w*) such that a higher weight indicates a stronger connection between two nodes (see Table 1). In this way, activation of a task concept in the Task layer will result in activation of associated Conceptual and Response nodes, instantiating retrieval of a task set in the model. Likewise, input to the Concept layer due to presentation of a target stimulus will activate the Response associated with that concept. Furthermore, reciprocal feedback connections loop from nodes in subordinate layers back to superordinate layers. Thus, feedback from active Concept and Response nodes can activate other associated Concept, Response, and Task nodes, including irrelevant and so competitive ones. Finally, within a layer, nodes are connected via mutual inhibitory connections (negative *w*). This feature makes these layers recurrent, capable of maintaining information in the absence of input

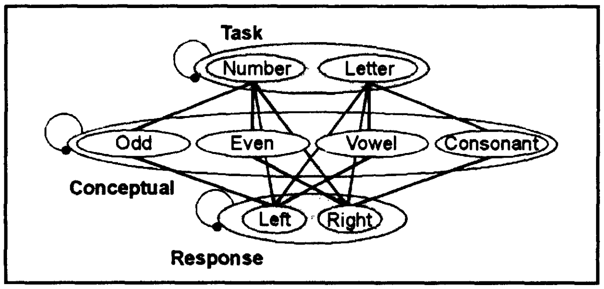

**Figure 1.** Schematic representation of the CAM-TS architecture.

(such as during the preparation interval), and further permits a computation of conflict (see below) based on the energy in the layer (Botvinick et al., 2001; Hopfield, 1982).

*Processing in CAM-TS*

At the start of a simulation trial, the activation values of all nodes in the Task, Concept, and Response layers were set to 0, regardless of what occurred on the previous trial, thereby nullifying the possibility of any transient carry-over. Presentation of a task cue initiated the first cycle of a trial and was simulated by delivering an input to one of the nodes in the Task layer and computing activation values across the network.

On each cycle, the net input (*n*) to node *i*, including external input, was computed according to Equation 1 (Botvinick et al., 2001; Yeung et al., 2004)

$$n_i = \sum_j a_j w_{ij} s_{ij} + \varepsilon \qquad \text{(E1)}$$

such that $a_j$ is the positive activation value of the *j*th input node on the last cycle, $w_{ij}$ is the weight between the *j*th input node and node *i*, and $s_{ij}$ is a constant scaling factor that only differed between external (0.4) and internal inputs (0.08). For external inputs, activation values were always 0 or 1, and *w* and *s* were determined by the external input weight (.15) and scaling (see Table 1). The noise term (ε) was distributed normally with a mean of 0 and standard deviation 0.01.

Based on $n_i$, the change in activation ($\Delta a_i$) was then computed according to Equation 2:

$$\Delta a_i = \begin{cases} (\text{max} - a_i) a_i n_i g_i - |(a_i - rest) * decay|, n_i \geq 0 \\ (a_i - \text{min}) a_i n_i g_i - |(a_i - rest) * decay|, n_i < 0 \end{cases} \qquad \text{(E2)}$$

This equation produces changes in activation equivalent to a logistic function such that as $a_i$ approaches *max* or *min*, the influence of $n_i$ diminishes, and $a_i$ tends to decay toward resting activation set by *rest* at the rate determined by *decay*. The net input ($n_i$) is scaled by the gain term ($g_i$) which, when increased, has the effect of making the activation function more sensitive to inputs, and can allow a recurrent network to maintain information in the absence of external input (Cohen & Servan-Schreiber, 1992). The gain was always uniform across all nodes of a layer, and unless otherwise noted, gain was set to 1.0. Computed values of $a_i$ were further bounded at *max* and *min*.

External input to the Task layer was maintained for 6 cycles after which all external input to the model was ceased and the model was allowed to cycle for a duration of cycles determined by CSI. During this preparation interval, activation in the relevant Task node, along with relevant activation in the associated Concept and Response nodes, comes to dominate. As with the experimental CSI manipulation, variation of CSI used in CAM-TS expanded with a logarithmic schedule setting preparation at 1, 2, 5 and 12 cycles.

The activation of the task set during the preparation interval is particularly enhanced by increasing the gain on the Task layer during the preparatory interval. To simulate a strategic control process coming on-line during a task switch, the gain term in the Task layer was increased to 2.0 (from 1.0) during the preparation interval of switch trials. The dynamics of this control process was such that gain was not increased until after the first cycle of the preparation interval, reflective of the slow onset of control, and was further only increased for a maximum of 3 cycles regardless of the duration of the CSI, after which it was reduced to its default level (1.0).

Following the preparation interval, external input was delivered to the Concept layer, reflecting the presentation of the target stimulus. As with the experimental paradigm used in Experiments 1 and 2 of Chapter 4, in which both a letter and number are presented as a target, input to the Concept layer of CAM-TS was applied equally to one of the number Concept nodes (Odd/Even) and one of the letter Concept nodes (Vowel/Concept). Identification of the target stimulus (such as recognizing the digit "1" as the number one) and its subsequent categorization were not modeled, as these factors were not manipulated in the present theoretical context and so would add little to what is already captured by the scaled external input to the Concept layer. Furthermore, such features have been shown not to interact with long-term carry-over effects (Sohn & Anderson, 2003). However, one could easily model these additional levels. Indeed, doing so might be helpful in detailing predictions regarding the modulation of switch costs by categorization difficulty (though see the associative strength simulation below).

External input to the Concept layer was maintained for 6 cycles after which all external input to the model was ceased and the model was allowed to cycle until it generated a response, up to a maximum of 100 cycles. A response was recorded once sufficient evidence accumulated in one of the Response nodes that its activation value exceeded a set threshold (.25). Following emission of a response, the gain in all layers was dropped to .5 to allow information to decay. This is reasonable as it is not likely that human subjects continue to actively maintain the task, target, or response following a button press.

On the cycle that a response was emitted, changes in the baseline weight ($\Delta w$) for a given input to the Task layer was computed according to Equation 3 (Gilbert & Shallice, 2002).

$$\Delta w_{ij} = a_i a_j \lambda \qquad \text{(E3)}$$

Such that $\Delta w_{ij}$ is determined by the product of the activation values ($a$) of the $i$th and $j$th nodes scaled by the learning rate ($\lambda$). In CAM-TS, $\lambda$ was always set to 1 emulating fast one trial learning. So that weights did not increase indefinitely, changes in weights were always made to baseline weight values rather than to the modified values from the previous trial. This is a simplifying feature of the model and follows others (Gilbert & Shallice, 2002), but is not meant as a theoretical position regarding repetition priming. Modifying the baseline weights after each trial using this equation has the effect of increasing the connection strength among nodes that are similarly active at the response (i.e., both are active) and diminishing the connection among those concepts and responses not similarly active at the response (i.e., one is active and the other is not). Hence, when a task switch occurs, associations with competitive representations will have been enhanced by the previous trial and associations with relevant representations will have been diminished. As this is a modulation of connection weights rather than activation values, this feature of the model is intended to reflect the sorts of small changes in long-term representations and pathways that underlie long-term repetition priming, as suggested by TSP.

*Simulation of RT and Conflict in CAM-TS*

The conversion of cycles to simulated RT used identical parameters to those already estimated by others in the context of other control tasks (Botvinick et al., 2001; Yeung et al., 2004) and was based on Equation 4.

$$RT = \alpha + cycles * \rho \qquad \text{(E4)}$$

The constant value $\alpha$ is meant to capture early perceptual processes not modeled by CAM-TS and was set at 200 ms for all simulations. The cycle conversion rate ($\rho$) was set at 16 ms for all simulations.

Equation 5 determined conflict in each layer at any cycle *c* based on the integrated computation of energy (*E*) (Hopfield, 1982) within the active portion of each layer.

$$E(c) = \int_{1}^{c} -\sum_{i}\sum_{j} a_{ic} a_{jc} w_{ijc} \qquad \text{(E5)}$$

Energy (*E*) was thus computed based on the integral across cycles of a trial (c = 100) of the sum of the products of the activation values (*a*) in the *i*th and *j*th nodes of a given layer weighted by their connection strength $w_{ij}$. Only nodes with activation values greater than or equal to 0 were included, as this reflects conflict specifically among retrieved representations in working memory. As noted by Botvinick (2001), computational features of Hopfield energy correspond to desirable conceptual features of conflict (Berlyne, 1957) in that energy increases exponentially with numbers of nodes actively competing within a layer and with the strength of their mutual activation levels.

Direct simulation of the BOLD response, for instance using a gamma function (e.g., Boynton, Engel, Glover, & Heeger, 1996), was unnecessary in the present model. The present estimates of conflict using CAM-TS would not have taken more than 1.5 seconds, less than a single TR of the imaging experiment. As such, the integrated conflict metric for a trial would act as a single instantaneous input to a gamma function that would elicit a corresponding linear increase in amplitude. Hence, for simplicity, the conflict scores were considered directly with respect to integrated measures from the deconvolved BOLD responses from individual ROI in Chapter 4, rather than transforming them linearly using a gamma function or other estimate of BOLD response.

CAM-TS was programmed in MATLAB (Mathworks, Inc.) and run on a Macintosh G4 computer. Where parameters were not taken directly from previously established models, parameters were either set by hand or were optimized using a cost minimization algorithm (Bogacz & Cohen, 2004) that searched the parameter space. As a source of independent output criteria for use with the algorithm, we used the basic switch versus repeat RT and error costs from Experiment 1 of Rogers and Monsell (1995). Hence, parameters were set before simulating the data for Chapter 4, Experiment 1. Once parameters were set, they were maintained unchanged for all subsequent simulations. Simulations of Experiments 1 and 2 were based on 50,000 trials per CSI condition. Outside of noted exceptions, all additional simulations consisted of 50,000 trials at a fixed CSI of 6 cycles and conditions were divided equally among trials.

**Table 1.** Summary of Parameters in CAM-TS

| Parameter | Value | Interpretation |
|---|---|---|
| Task to Concept *w* | 2.2 | Feedforward projection from Task nodes to relevant Concept nodes |
| Concept to Response *w* | 1.5 | Feedforward projection from Concept nodes to linked Response nodes |
| Concept to Task *w* | 1.7 | Feedback projection from Concept nodes to Task nodes |
| Response to Concept *w* | 0.2 | Feedback projection from Response nodes to Category nodes |
| Response to Task *w* | 0.5 | Feedback projection from Response nodes to Task nodes |
| Task to Task *w* | -3.7 | Mutual inhibitory connection in Task layer |
| Concept to Concept *w* | -1.0 | Mutual inhibitory connection in Concept layer |
| Response to Response *w* | -1.5 | Mutual inhibitory connection in Response layer |
| External input to Task *w* | 0.15 | Weight on external input to the Task layer |
| External input to Concept *w* | 0.15 | Weight on external input to the Concept layer |
| Internal scaling *s* | 0.08 | Scales product of input activation and weights |
| External scaling *s* | 0.4 | Scales external input |
| Response Threshold | 0.25 | Threshold for Response nodes at which response is recorded |
| Decay Rate | 0.1 | Controls rate at which activation values go to resting |
| Max (Activation) | 1.0 | Maximum activation value |
| Min (Activation) | -0.2 | Minimum activation value |
| Rest | -0.1 | Resting activation value |
| Noise $\mu$ | 0.0 | Mean of noise distribution |
| Noise SD | 0.01 | Standard deviation of noise distribution |
| Default *g* | 1.0 | Default gain term |
| Preparation *g* | 2.0 | Increased preparatory gain term to up-regulate control |
| Learning rate $\lambda$ | 1.0 | Post-response learning rate for weight modification |

## CAM-TS Simulations of Common Phenomena in Task Switching

Though CAM-TS was built to provide an explicit theoretical context for understanding the results from the fMRI experiment of task switching presented in Chapter 4, it also provides a means of assessing the extent to which a number of established behavioral phenomena associated with task switching may be accounted for by the control of associative memory perspective. In the following section, a number of CAM-TS simulations determine the model's ability to account for a number of extant behavioral results.

### *Explicit Cueing Task – Global Preparation and Errors*

The CAM-TS simulation of the explicit cueing task is already described in Chapter 4, including its strong fit to the preparation effect in RT switch costs from Experiment 1 and its characterization of a conceptual conflict signal during task switching. Two additional points about this simulation may be worth noting: global preparation and errors.

A variant of CAM-TS might still include control but would prepare by increasing control on both Switch and Repeat trials, rather than only on switch trials. In CAM-TS, global preparation of this kind results in a decrease in switch costs over CSI, still demonstrating the advantage of prospective control, but the decrease is considerably more linear (Figure 2A) than the log decrease evident from Experiment 1.

Interestingly, the linear decrease due to global preparation also appears to fit the decrease in RT cost from the fMRI experiment. Though the signature of conceptual conflict does not change qualitatively with global preparation, and so this does not undermine the central arguments regarding a preparatory control process, this raises the possibility that some subjects may have adopted a more general preparatory strategy in the fMRI experiment. This strategy difference may be due to the procedure used in the fMRI for varying CSI. Whereas in the behavioral experiment, CSI could be fully blocked, in fMRI, deconvolution required a different CSI between the first event in a task pair and the second event. Under such variable preparation intervals, switch cost declines can be difficult to obtain (Rogers & Monsell, 1995), and so uncertainty regarding preparation time might result in strategy differences in the application of control by subjects.

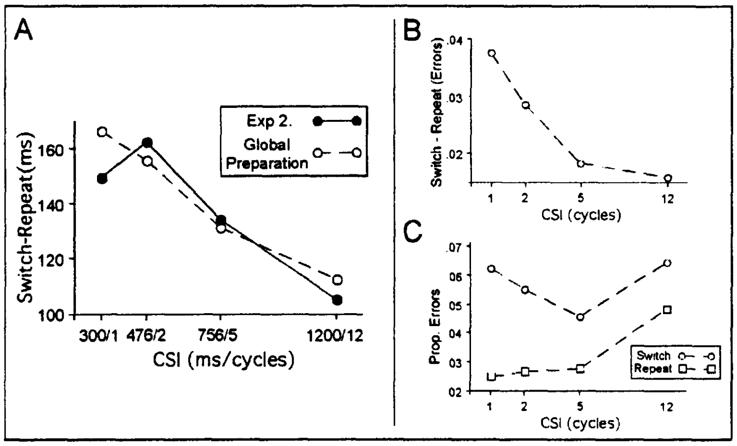

**Figure 2.** Additional explicit cueing task simulation results. (A) Global preparation resulted in a linear decline in switch costs that resembled the median decay pattern from Experiment 2 (Chapter 4). (B) Error cost declined with increasing CSI. (C) Across condition, errors increased slightly at the longest CSI though the Switch versus Repeat difference in errors still diminished.

Errors have not, to date, been a major focus in the task switching literature, perhaps because they do not lend themselves easily to interpretation in terms of time-consuming processes along serial information processing stages. Nevertheless, task switch costs in errors are evident and also tend to decline with increasing CSI. Errors in CAM-TS, as with other bias competition models (Yeung et al., 2004), arise entirely from the noise term which, in addition to making the model non-deterministic, also can produce loss of information and even slips of action if chance pushes activation values too high too quickly and in the wrong response node, before evidence has had time to accumulate in the appropriate response node.

Qualitatively similar to error costs reported in the task switching literature, a higher proportion of Switch than Repeat trials in CAM-TS produced an error. Furthermore, this error cost declined with increasing CSI (Figure 2B). At the longest CSI, overall errors increased for Switch and Repeat trials (Figure 2C). It is important to note, however, that the Switch versus Repeat error difference still decreased at the longest CSI, indicating that the critical decline in switch RT cost was not due to the model trading speed for accuracy differentially at longer CSI. However, the rise in errors at the longest CSI is likely due to the fact that at longer retention intervals there is a greater probability that the noise term will artificially enhance activation values in irrelevant nodes. This may suggest that up-regulation of gain may need to dynamically adjust to longer retention intervals.

### *Alternating Runs and First Trial Effects*

An important outcome to emerge from the classic alternating runs procedure of Rogers and Monsell (1995; Monsell, Sumner, & Waters, 2003) was the *first trial effect*, in that only the first trial in a run of trials of a given task, which constituted a task switch, demonstrated an elevated RT cost. Critically, all subsequent repeat trials were equally facilitated relative to the first trial. This effect was originally considered difficult to account for from an interference perspective, particularly theories relying on short-term transient interference such as TSI, as it suggested that passive decay did not carry-over beyond the first trial. Gilbert and Shallice (2002) used a connectionist model of task switching, similar to the present framework but relying principally on transient carry-over, to show that no additional decay was required by transient carry-over. Hence, such effects are important to demonstrate in CAM-TS.

CAM-TS performed 50,000 trials consisting of alternating runs of 4 trials of each task. Hence, the first trial of a run constituted a task switch and the subsequent trials were repeats. As plotted in Figure 3A (see below), following an initial elevated switch RT, subsequent trials in the run did not show any further substantial declines in RT.

### *Cross-Talk Interference*

In addition to response repetition (RR) effects, highlighted in the present empirical work, another interference effect commonly observed during task switching is due to *cross-talk* (Meiran, Chorev, & Sapir, 2000; Rogers & Monsell, 1995). Typically, target stimuli during task switching paradigms are bivalent, in that they hold information associated with tasks other than the currently relevant task (as in a paired letter and number in the present experiment). Trials on which the irrelevant cue in the target stimulus is relevant to a competing task (e.g., "a1") may be compared to those in which it is neutral (e.g., "a#"). The former condition, referred to as the cross-talk condition, produces elevated switch costs relative to the latter, neutral condition (Meiran et al., 2000; Rogers & Monsell, 1995).

In CAM-TS, the flanking stimulus may act as a cue for primed representations of a competing task. Cross-talk priming of this type would likely increase conflict in the Conceptual layer and so increase activation in left VLPFC. Potentially consistent with this interpretation from the theoretical perspective regarding VLPFC function forwarded here, an imaging study that manipulated cross-talk during task switching located a modulation of switch effects in left mid-VLPFC in association with the presence of cross-talk (Brass & von Cramon, 2004). It follows, then, that CAM-TS should show an enhancement in switch cost during cross-talk relative to neutral conditions, and further should show an enhanced conflict signal in its Conceptual layer.

Cross-talk and neutral conditions were simulated in CAM-TS by comparing neutral trials on which only the task-relevant conceptual node received external input at the target presentation phase to cross-talk trials on which both a number conceptual node (Odd/Even) and a letter conceptual node (Vowel/Consonant) received equal external input, as under the standard simulation conditions. Consistent with behavioral cross-talk effects, there was a larger RT cost for cross-talk relative to neutral trials (Figure 3A). Moreover, conflict in the Conceptual layer was higher in cross-talk relative to neutral conditions.

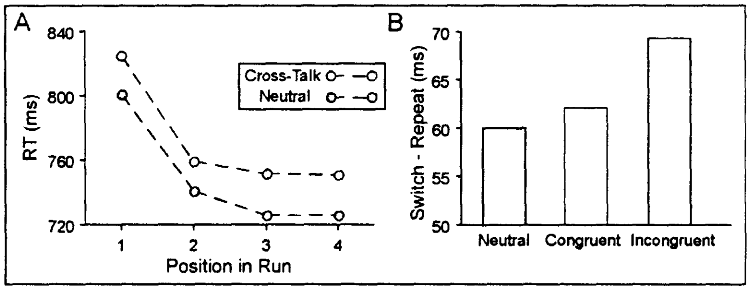

**Figure 3.** Alternating runs simulation results. (A) Most of the switch cost is resolved after the first trial of a run of a new task for both Cross-Talk and Neutral trials. (B) Simulations with both Congruent and Incongruent Cross-Talk trials produce a larger switch cost than Neutral trials.

One may further specify cross-talk by the congruency of the response cued by the flanking stimulus with the response cued by the task relevant stimulus. If they cue the same response, they are Congruent and otherwise they are Incongruent. Though reliable congruency modulations of task switching costs in the behavioral literature have been difficult to obtain (Meiran et al., 2000; Rogers & Monsell, 1995), it appears that the typical quantitative trend is for Incongruent trials to elicit a greater switch cost than Congruent trials, though both Congruent and Incongruent trials have greater switch costs that neutral trials. This latter effect suggests that the mere presence of a cue relevant to another task may produce interference by cueing non-response representations. The former effect does suggest that the conflicting response may also produce some additional conflict.

Interestingly, the qualitative pattern from CAM-TS is consistent with the behavioral literature (Figure 3B). Though both Congruent and Incongruent trials show greater switch costs than neutral trials, switch costs are greater on Incongruent than Congruent trials (Figure 3B). Hence, CAM-TS may account for cross-talk effects. It may be of further interest to note that the model does not predict any interaction between the increments in cost due to Incongruent versus Congruent and that due to RR versus RD. This is consistent with the results from Experiment 1 (TS × RR × Cong: F(1,23) = .007, $p = .93$) and from Experiment 2 (F(1,9) = 3.4, $p = .1$).

### *Asymmetrical Costs*

Switching from an a less familiar to a more familiar, and so easier, task results in a larger switch cost than the reverse case (Allport, Styles, & Hsieh, 1994). This counterintuitive *asymmetrical cost* was among the first findings in favor of interference effects in task switching, as it was difficult to propose a reconfiguration process that would need to overcome increased reconfiguration demands when accessing a better-learned task.

Asymmetrical costs were established behaviorally in modified versions of the Stroop task where there is an asymmetry in response conflict such that color naming in the presence of an incongruent word produces interference, whereas word reading in the presence of a congruent color produces no interference (MacLeod, 1991; Stroop, 1935). Interestingly, switching from the easier word reading to the more difficult color naming task produces less of a switch cost than switching from the more difficult color naming task to the less difficult word reading task (Allport et al., 1994).

CAM-TS might account for asymmetrical costs by noting that the increased control during color naming differentially enhances associative pathways during color naming, which negatively impacts a subsequent attempt at word reading. A special simulation procedure sought to test the capability of CAM-TS to produce an asymmetrical cost.

First, a prepotent advantage was arranged for one of the two tasks (Number/Letter) modeled by CAM-TS. In an initial training phase, the model performed the tasks in a random sequence for 100 trials. At the end of each trial, connection weights were calculated according to E3 for the Concept layer. However, in this training stage, the weights were allowed to accumulate, rather than only being added to the baseline weights after each trial. As the model did not perform the two tasks with equal frequency, a slight bias for one of the two tasks emerged in the accumulating connection weights. Following this training period, the model was run with these new connection weights, using only incongruent target input, and otherwise standard parameters for 1000 trials at a CSI of 5 cycles.

Figure 4 depicts the results from the asymmetrical cost simulation. First, it is clear that the initial training period successfully produced an advantage for the Number task, as both Switch and Repeat were faster than the Letter task. However, the Switch cost for the more difficult Letter task (64 ms) was less than the Switch cost for the easier Number task (72 ms). Hence, CAM-TS can produce asymmetrical costs, even in the absence of transient carry-over required by TSI.

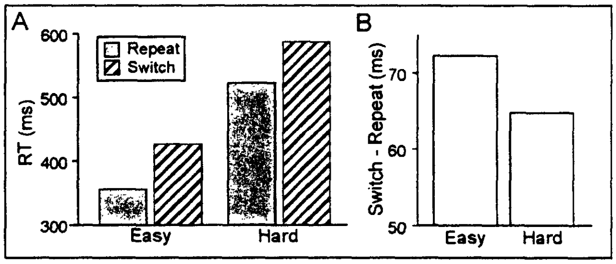

**Figure 4.** Results from asymmetrical cost simulation. (A) The training successfully produced an advantage for one of the tasks for both Switch and Repeat trials. (B) However, the switch cost was larger for the easier task.

*Associative Strength and Task Cueing Effects*

An early indication that the processes that guide access to memory may play an important role in task switching was evident from the earliest task switching experiments (Jersild, 1927; Spector & Biederman, 1976) that manipulated the difficulty of retrieval during a task switch. Such manipulations typically relied on arithmetic operations of varying difficulty and demonstrated increased switch costs under these more difficult retrieval conditions. The increased switch costs have been interpreted as arising from increased demands on retrieval processes that must access memory to reconfigure the system for an upcoming task (Mayr & Kliegl, 2000; Rubinstein, Meyer, & Evans, 2001). Interestingly, from CAM-TS, task switching costs would also be greater under such conditions, as the weaker activation of relevant Concept nodes due to the weak association of a target to a Concept node would reduce the efficacy of these Conceptual nodes in selecting a response and, critically for task switching, make them more susceptible to conflict. By contrast, during a task repeat, priming of the relevant Task to Concept pathway helps overcome the weak association between the target and the Concept node.

To simulate associative strength effects in CAM-TS, the weight (*w* in E1) on the external input to the concept layer was reduced from .15 to .12. This was meant to emulate a weaker connection between presented targets and their associated Concept nodes. With this parameter change, CAM-TS performed 50,000 trials at otherwise standard parameters.

The results from the associative strength simulation are plotted in Figure 5 relative to the results from the standard simulation using the default external weight value ($w = .15$). Consistent with the empirical data, a lower associative strength resulted in a larger switch cost (88.8 ms versus 72.1 ms).

A second factor that influences switch costs is the specification of a task cue. Indeed, a large independent portion of the switch cost may be attributed to cue switches (Logan & Bundesen, 2003; 2004 2001; Mayr & Kliegl, 2003). Rubinstein and Meyer (2001) showed that an increase in switch cost due to a manipulation of associative strength did not interact with an increase in switch cost due to a manipulation of task specification (making it easier or harder to determine what the next task was in the series). Critically, these results have been interpreted as evidence for a dissociable process of rule activation, perhaps analogous to the top down mnemonic control process in CAM-TS, and a process of goal setting process, not modeled in CAM-TS, that determined what the next task to perform might be.

A goal setting process is not inconsistent with CAM-TS, and indeed, may be an integral part of the task switching network. However, similar to the associative strength simulation, we may simulate the impact of a deficient goal setting process in the model by reducing the weight on the external input to the Task layer (to *w* = .12 from *w* = .15). Simulating task switching in CAM-TS under this modified parameter produced an enhanced switch cost (111.4 ms; Figure 5) greater than under standard conditions (72.1 ms).

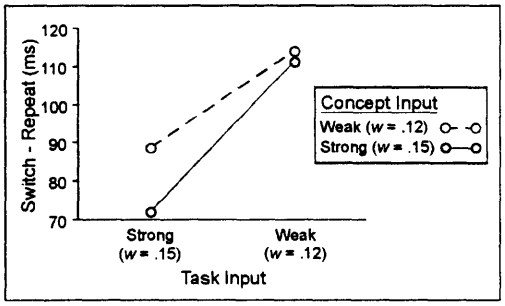

**Figure 5.** Reducing the input weights to the Task and Concept layers produces increased switch costs. But when both are reduced the increments combine underadditively, supporting their independence.

The finding of an increased cost for a weak input to the Task layer motivates testing whether the enhanced switch costs in CAM-TS due to simultaneously weak input to the Task and Concept layers will share the independence cited as evidence for a dissociation between rule activation and goal setting processes. In particular, if these effects, in concert, result in an overadditive interaction in switch cost, it would argue against their independence in CAM-TS, and this would be difficult to reconcile with the results from Rubinstein and Meyer (2001). Hence, it is an important test of CAM-TS to determine whether these effects are independent.

An additional simulation was run using CAM-TS in which the external input connection strength was reduced to $w = .12$ for both the Task and Concept layers. Critically, the results revealed an underadditive effect (Figure 5), such that the switch cost under weakened Task and Concept input was 114.2 ms. This switch cost was still enhanced relative to the standard switch cost (72.1 ms) but was practically equal to the enhancement due to a weak Task input (111.4 ms), the larger of the two input effects. Not only does this underadditivity suggest independence between these effects, but it further suggests that these effects take place in parallel and so are delimited by the longer of the two enhancements (the Task effect). An additional, independent process, such as goal setting, that increases RT when the task is ambiguous would add on independently to the Task effect and so would result in the additive effect reported by Rubinstein and Meyer (2001). Hence, CAM-TS is consistent with a goal setting/rule activation distinction.

### *Mixing Costs*

Mixing costs are derived from list completion times, in which averaged RT from pure, single task lists are faster than list completion times from an alternating task list (Meiran, 2000; Monsell, 2003). However, a portion of the mixing cost may be independent of alternation costs, which arise as a Switch versus Repeat trial difference in paradigms like the explicit cueing task and alternating runs. This independence is evident in that RT on Repeat trials during alternating runs is slower than the mean trial RT from a pure task list (Meiran, 2000). One possible account of this divergence notes a difference in the demand to maintain the current task context in working memory in order to determine what task to do next. Hence, such a demand may stem from processes, such as goal setting, not modeled in the current context. Alternatively, long-term carry-over effects may also impact Repeat trials during an alternating block (Koch, Prinz, & Allport, 2005). Thus, it is of central interest to assess whether any portion of the independent portion of mixing costs can still be accounted for by CAM-TS.

To simulate mixing costs, two additional simulations were run using CAM-TS. First, a pure block (25,000 trials) of the Letter task was run, during which no task switches occurred. Then, for the next 25,000 trials the model alternated on every trial between the Letter and Number task. The simulated trial-averaged RT from these blocks was compared to the Switch and Repeat RT from the alternating runs simulation (Figure 6). Though the differences between mixing and alternating blocks were small, the results are qualitatively similar to the mixing costs obtained in the literature. There was a larger mixing cost (79.5 ms versus 72.1 ms alternation cost), due to a slightly faster pure block RT (749.5 ms versus 753.6 ms Repeat RT) and a slightly slower alternating block RT (829.1 ms versus 825.7 ms Switch RT). Hence, CAM-TS appears capable of capturing at least a portion of mixing costs, as well.

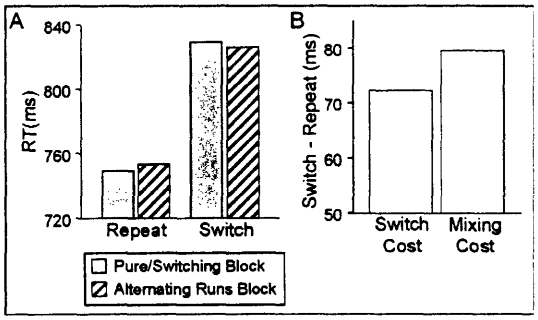

**Figure 6.** (A) Average RT from a Pure single task block was slightly faster than RT from a Repeat trial of an Alternating block. Likewisethe average RT from a Switching block (switch on every trial) was slower than the Switch RT from an alternating block. (B) This resulted in a larger mixing cost (Switching - Pure block) than Switch cost (Switch - Repeat trials).

Though qualitatively consistent with the pattern observed in the literature, there may be a number of reasons for the relatively small mixing costs (as well as other small effects reported here) simulated by CAM-TS. First, as an effort was made not to change any parameters for any of the above simulations, it is likely that searching the parameter space more exhaustively for each simulation might optimize the model to produce specific effects reported in the literature. Hence, it is significant that the qualitative effect was present without changing any parameter beyond those required to simulate the experiment. With respect to mixing costs in particular, a portion of behavioral mixing costs may also come from demands on goal setting or task specification processes that are not modeled in CAM-TS. This is reasonable in that during an alternating block, there is a stronger demand to determine what task to do next than during a pure block. A second reason that CAM-TS may not account for a full mixing cost is that the long-term repetition priming effects are only modeled on a trial-to-trial basis in CAM-TS. Hence, a portion of mixing costs may be due to long-term carry-over accumulating over multiple preceding trials. Trial order effects beyond one trial back are currently not modeled by CAM-TS, but may be the source of additional predictions and explanatory power in more sophisticated future versions of the model.

## Relationship to Previous Models

In preceding section, the performance of CAM-TS across a variety of different task switching manipulations demonstrated the capability of the model to account for a number of established behavioral phenomena, and also suggest the boundaries of the model, particularly with respect to goal setting/task specification processes. CAM-TS is not the first formal approach to task switching, though it is the first explicit model to account for both such a broad range of task switching phenomena and also to provide a theoretical context for understanding the impact of interference on BOLD signal differences among critical regions in the brain. It is, however, important to establish how CAM-TS is consistent with and differs from previous models devoted to task switching.

### *Gilbert and Shallice, 2002*

CAM-TS is highly similar in architecture to an independently developed parallel distributed processing model by Gilbert and Shallice (2002). The Gilbert and Shallice (2002) model was designed primarily to simulate short-term transient interference effects, of the type proposed by TSI, and to demonstrate how these effects were capable of accounting for asymmetrical costs as well as first trial effects, the latter of which had been, to that point, considered incommensurate with TSI. However, this model also included a weight adjustment procedure identical to the one used in CAM-TS, so TSP was also included in some simulations of this model. However, though the source of task switch interference partially overlaps between the models, they differ fundamentally with respect to the implementation of control, which is a central component of CAM-TS. We consider this and other differences among the models below.

The Gilbert and Shallice (2002) architecture was developed by modifying the parallel distributed processing model of the Stroop task developed by Cohen et al. (1990). The model consists of three color input units (red, green, blue) that feedforward to three color output units and likewise three word input units ("red", "green", "blue") that feedforward to three word output units. The output units are in recurrent layers and share inhibitory connections with the output layers from the other task (e.g., word or color).

The connection strengths between word input and word output units are stronger than between the color input and output units, simulating the dominance of this task. This prepotency ensures that the model will always respond with the word rather than the color, in the absence of control. Hence, to allow the color output to be read off when appropriate, a task demand unit is connected to the output units of each pathway, and each task demand unit shares an excitatory connection with one set of output units and an inhibitory connection with the other. The task demand units also receive feedback connections from the input and output units. Control is exerted through an additional external input to the task demand layer that is differentially strong when the color task is required, thereby allowing the task demand layer to bias the appropriate color pathway in the face of the prepotent word pathway. Finally, the model uses a "random walk" response threshold (e.g., Logan & Gordon, 2001) based on the maximum difference among response alternatives as opposed to the absolute threshold used in CAM-TS[2] and other models (e.g., Cohen et al., 1990). Task switching effects were elicited (1) by allowing activation values in the model to carry-over at 20% of their value thus implementing transient carry-over (e.g., TSI), and (2) implementing a weight adjustment computed the same as in CAM-TS, but restricted to the connections from the stimulus input units to the task demand units rather than all inputs to the task layer, as in CAM-TS.

The model of Gilbert and Shallice (2002) was capable of producing asymmetrical switch costs, first trial effects, and even a decline in switch cost with increasing preparation time that declined to zero at the longest interval (150 cycles). However, there

---

[2] Inclusion of a random walk response threshold in CAM-TS reduces the overall switch costs, but still results in qualitatively similar preparation effects. It may be an interesting question for future research what differs between random walk and absolute thresholds.

was apparently no strategic implementation of control during a task switch and the source of task switching effects was somewhat different, particularly in including transient carry-over in addition to long-term carry-over. Hence, these two apparent differences between the models merit additional consideration.

Control is not absent from the Gilbert and Shallice (2002) model. Indeed, it comes in the form of a weighted input that is delivered throughout the preparation interval and persists during target presentations. To some extent, an input delivered in this manner may be more consistent with experiments during which the task cue is presented to the subject throughout the preparation interval and remains on the screen after presentation of the target. This procedure was not used in the experiments described in Chapter 4 nor in the simulations of CAM-TS, in which the task cue, in the form of external input, was modeled separately from control, and was removed following an initial encoding period (250 ms/6 cycles). To some extent, as modeled by Gilbert and Shallice (2002), the model is stimulus controlled, in that the external stimulus provides a bias on the task layer and acts in a manner similar to a gain modulation on the task layer. However, it is not the case that the stimulus entirely controls the task layer. The stimulus input is weighted, and that weight depends on the task to be performed. More specifically, when preparing to perform the more difficult color task, the weighted input (effectively a gain modulation) is higher than when preparing to perform the color task. It is clear that no differentiation is made in the application of control between Switch and Repeat trials, but this is essentially the same as with global preparation in CAM-TS. Control is still required during the preparation interval to achieve a task switch.

Control in the Gilbert and Shallice (2002) model also differs from CAM-TS in that the biased control input is sustained throughout the preparation interval rather than having the slow-onset/fixed-duration characteristic of CAM-TS. This is, in part, possible because of a structural feature of the Gilbert and Shallice (2002) model not included in CAM-TS, in that activation in the output and input layers (e.g., the Response and Concept layers in CAM-TS) is clamped at 0 during the preparation interval. Were this not to be the case, the strong control signal from the task layer during the retention interval would push these subordinate layers to settle onto a local maximum prior to presentation of a target. However, clamping these other layers during the preparation interval potentially negates some important benefits of preparation, namely activating relevant representations and pathways in anticipation of a target. Indeed, preparation prior to a task switch results in an anticipatory parietal discrimination potential determined by the upcoming task (Wylie, Javitt, & Foxe, 2003), evidence of the impact of preparation on subsequent processing. Hence, clamping other layers should not be necessary in a network during preparation. Two mechanisms can be implemented to overcome this limitation and not require clamping of other layers in the model during preparation: (1) Control can act dynamically, as in CAM-TS, onsetting and offsetting over a fixed portion of any preparation interval or (2) the gain terms of other layers can also be under control, in that they can be reduced so they don't maintain information as readily and so do not settle as quickly toward a local maximum. However, both of these implementations require a control process. Hence, though different in its dynamics, control is at least implicitly required by the model of Gilbert and Shallice (2002).

A second central difference between CAM-TS and the Gilbert and Shallice (2002) model concerns the source of task switching effects. Whereas CAM-TS models switching effects as exclusively arising from long-term carry-over interference and locates conflict in the Conceptual layer, the Gilbert and Shallice (2002) model includes both short- and long-term carry-over, with focus primarily on conflict within the Task layer. Hence, the conflict largely determining the switch cost in the Gilbert and Shallice (2002) model comes from short-term carry-over of activation values in the Task level, with activation of the irrelevant task representation making activation of the relevant task representation more difficult. As activation values in the Concept and Response layers are clamped at 0, the benefit in the preparation interval does not come as prospective activation of relevant representations in advance of interference, as with CAM-TS, but rather provides the opportunity for the relevant Task representation to compete with the irrelevant Task representation and come to dominate the control layer. As noted by the authors, should sufficient controlling input to the Task layer not be provided, the model would perseverate on the irrelevant task. When the activation values in the Task layer of the Gilbert and Shallice (2002) model are plotted over cycles (Gilbert and Shallice, 2002 [Figure 5]), one can see that the number of cycles taken for the relevant Task unit to overcome the transient carry-over of irrelevant Task unit is the determining factor for the switch cost and its reduction with preparation.

This is an important difference from CAM-TS. The position of CAM-TS is that conflict is in the Conceptual layer and the up-regulation of the Task layer enacts control to overcome this interference. There is no carry-over in the Task layer, so the appropriate Task representation always comes to dominate this layer upon presentation of the Task cue. However, the efficacy of this layer in biasing the representations in the Conceptual layer in advance of conflict upon presentation of a task stimulus is diminished. This means that in the absence of up-regulation of control, the model will not perseverate but will simply take longer to converge on a response because of interference in the Conceptual layer from weak but active competitors. This mechanism appears to correspond better to the neuropsychological data in that patients with left lateral PFC damage show enhanced switch costs rather than perseveration (e.g., Rogers et al., 1998). Furthermore, the evidence described in Chapters 2, 3, and 4 links left mid-VLPFC to a process that controls memory to overcome interference, and so the perspective of CAM-TS may fit better with the broader neuroimaging data, as well.

As a final note, it is not necessarily the position of CAM-TS that short-term carry-over effects do not exist. Indeed, one can implement these effects in CAM-TS, as well, either multiplying a squashing term by the activation values of the Task layer, as with Gilbert and Shallice (2002), or preferably, by reducing the gain in the Task layer but not resetting the values to 0 at the start of the next trial, thus allowing the activation value to decay in a time-dependent manner prior to the next trial. Indeed, this latter implementation would permit implementation of short-term decay effects. However, CAM-TS does demonstrate that a number of phenomena, both behaviorally and in fMRI, may be accounted for solely on the basis of long-term carry-over effects.

### *The Intention-Activation Model*

The intention-activation model (Dejong, 2000) has been an influential analytical approach to understanding switch cost preparation effects. Intention-activation relies on the statistics of finite mixture distributions, and so fundamentally depends upon a discrete binary state assumption, similar to the view that prevailed in traditional theories of task switching. However, unlike the traditional theory, intention-activation is capable of accounting for the problematic residual switch cost—the phenomenon whereby a switch cost persists even after a very long preparation interval—and it can do so without recourse to a second exogenous process as with TSR. Hence, intention-activation has been a very influential approach and represents an important alternative to CAM-TS and other interference-based models. Thus, this alternative is considered in some detail in this section.

Intention-activation assumes that at the presentation of a given target, a subject will be in one of two states, prepared or unprepared, and the state in which she/he resides will determine the RT for that trial. Furthermore, depending on experimental conditions, a probability exists, called the mixing probability, which describes the likelihood that the subject is in one state or the other. Consequently, the RT for a given trial will come from a finite mixture distribution (Evritt, 1985; Thomas, 1969), $m(x)$, that can be expressed as

$$m(x) = \alpha f_{prepared}(x) + (1-\alpha) f_{unprepared}(x) \qquad \text{(E6)}$$

where $f_{prepared}(x)$ is the basis distribution of RT when the subject is prepared, $f_{unprepared}(x)$ is the basis distribution of RT when the subject is unprepared, and $\alpha$ is the mixing probability. Hence, unless $\alpha$ is 1, there is always a chance that an RT will come from the unprepared distribution, and so the result will be a residual switch cost at a long CSI because this condition averages together a mixture of two distributions, a high proportion of prepared trials on which the switch cost was eliminated and a smaller set of unprepared trials when it was not. From intention-activation, what determines this mixing probability is the intent to engage in preparation. However, more generally, for any binary state model, α indicates the probability of successfully reconfiguring the system for the upcoming task, even including a successful memory retrieval process (Mayr & Kliegl, 2000). Hence, this model has been cited by a number of theorists as capturing the essence of preparatory control proposed by TSR and as an account of residual costs (Monsell, 2003; Nieuwenhuis & Monsell, 2002).

The mixture model was tested in a set of studies using the explicit cueing (Dejong, 2000) and alternating runs paradigm (Nieuwenhuis & Monsell, 2002; Rogers & Monsell, 1995). These studies showed that a mixture model was capable of accounting for the residual cost, and further that the fit of the model was not improved by including a baseline shift in RT meant to directly model the residual cost. However, it is important to consider the details of how these approaches fit a mixture model to the observed preparation data.

So that basis and mixture distributions could be defined empirically, Equation 6 was adapted as shown in Equation 7

$$m_{switch\,/\,longcsi}(x) = \alpha f_{repeat\,/\,longcsi}(x) + (1-\alpha) f_{switch\,/\,shortcsi}(x) \qquad \text{(E7)}$$

Hence, subjects were considered to be in an almost entirely unprepared state on Switch trials following only a short CSI and in an almost entirely prepared state for Repeat trials following the longest CSI. These basis RT distributions were then be used to estimate the mixing probability (α) that would produce the mixture distribution; RTs on a switch trial following a long CSI for which a residual cost is commonly evident.

To estimate α, these studies used a maximum likelihood estimation algorithm (Yantis, Meyer, & Smith, 1991) developed to account for not only binary mixtures, such as in the present mixture model, but also for multinomial mixtures for which analytic estimation procedures do not exist. The multinomial maximum likelihood mixture (MMLM) algorithm (Yantis et al., 1991) simultaneously estimates the basis and mixture distributions, as well as the mixture probabilities, based on observed basis and mixture distributions. Furthermore, it produces a goodness-of-fit statistic ($G^2$) that is computed based on the distance between the observed and estimated basis and mixture distributions. $G^2$ takes on a $\chi^2$ distribution with degrees of freedom computed as

$$(K-J)*(I-J) \tag{E8}$$

based on K distributions of which J are basis distributions and I is the number of bins over which observations are distributed to make up the observed distribution. Critically, the null hypothesis of this test is that a mixture distribution exists. So, a reliable $G^2$ statistic indicates that the data are not well fit by a mixture distribution, whereas a null result indicates a failure to disconfirm a mixture distribution.

Previous studies that used the MMLM algorithm demonstrated that residual switch costs at the longest CSI of task switching experiments, including the classic data from Experiment 5 of Rogers and Monsell (1995), could be accounted for using a mixture model by showing that that $G^2$ was not reliable. Furthermore, by experimentally increasing subjects' motivation, residual costs were shown to decrease and the estimate of the mixing probability increased (Nieuwenhuis & Monsell, 2002). This was considered indicative of a reduction in the rate of failure to engage in a preparatory switch.

The intentional-activation and mixture models of task switching are an elegant and influential approach to formalizing reconfiguration theories of task switching.

Furthermore, these models encourage analysis of the entire RT distribution as opposed to only estimates of the first moment. However, the evidence to date may not be sufficient to provide strong support for these models in task switching. MMLM, relying on an EM estimation algorithm, is very powerful and does an excellent job of generating basis and mixture distributions that fit the properties of a true mixture. However, the nature of the statistical test associated with this algorithm, particularly the null hypothesis, may make it difficult to use as a test of mixture models of task switching. In particular, data coming from studies to date may not have derived from a mixture distribution but simply had insufficient power to reveal that a mixture distribution did not give rise to the data. Recognizing this point, Nieuwenhuis and Monsell (2002) noted that without an alternative quantitative hypothesis, such issues of power are difficult to assess. Hence, CAM-TS may be useful is assessing the adequacy of extant empirical attempts to validate intention-activation using MMLM.

CAM-TS is not well characterized by two discrete prepared and unprepared states. Rather, the state of the model at any time step is simply the activity in its units and the strength of its weights. Its progression over a preparation interval represents a continuum of individual states giving rise to a corresponding continuum of distributions rather than two discrete prepared and unprepared distributions. Thus, as an alternative quantitative hypothesis, we may be confident that the switch declines from CAM-TS do not arise from a binary mixture. Indeed, at a high sampling rate (>1000 trials per CSI) sufficient to generate a distribution over 20 bins and using all intermediate switch CSI distributions from the standard simulation, MMLM is capable of disconfirming that the output of CAM-TS comes from a finite mixture distribution ($G^2(54) = 73.4$, $p < .05$).

Hence, it is reasonable, then, to assess whether the $G^2$ test would be capable of disconfirming a binary mixture model for the output of CAM-TS at the sampling rate utilized by extant studies.

Three simulations were run using standard settings in CAM-TS with the following changes. The first simulation followed Dejong (2000) by using 400 trials[3] for each CSI conditions in 8 simulations (e.g., 8 subjects). The second simulation followed Rogers and Monsell (1995) by using 288 trials in 10 simulations. The third simulation simulated the motivation study (Nieuwenhuis & Monsell, 2002) by slightly increasing the preparatory gain parameter ($g$, E2) from 2.0 to 2.3 for 240 trials in 11 simulations. Mixture models were then estimated using MMLM based on the output from each simulation. The procedures were identical to those used by others (Dejong, 2000; Nieuwenhuis & Monsell, 2002).

As depicted in Figure 7, using the same experimental parameters in CAM-TS as those of previous studies produced output that could be modeled using MMLM as a mixture distribution. The Dejong simulation, that had the highest power of the three (400 trials/CSI), resulted in a statistical null ($G^2(24) = 20.6$, $p = .66$), and so failed to disconfirm that a mixture distribution gave rise to the output of CAM-TS. Likewise, the Rogers and Monsell (1995) simulation also resulted in a statistical null ($G^2(30) = 38.3$, $p = .14$). The motivation simulation (Nieuwenhuis & Monsell, 2002) in which the gain was higher ($g$=2.3) during the preparation interval estimated its mixing probability ($\alpha$) at the

---

[3] The number of trials used by Dejong (2000) was difficult to determine from the text. They cite 124 blocks divided into short blocks (12 trials) and long blocks (48 trials) in a 4:1 ratio. However, 124 is not divisible into whole blocks that can take this ratio. If 125 are used, there are an equal number of trials on short and long blocks (400) for which separate mixture models were estimated. Thus, 400 trials were used in the simulation.

longest CSI to be .83, significantly higher than the mixing probability estimated for the Rogers and Monsell (1995) simulation ($\alpha = .75$; $t(19)=2.2$, $p < .05$). If interpreted the same as previous empirical results, this shift would be considered indicative of a larger proportion of" high motivation" RTs coming from a prepared distribution than under standard conditions. Again the goodness-of-fit test resulted in a null ($G^2(33) = 40.8$, $p = .16$) and so failed to disconfirm that a mixture model gave rise to the output of CAM-TS.

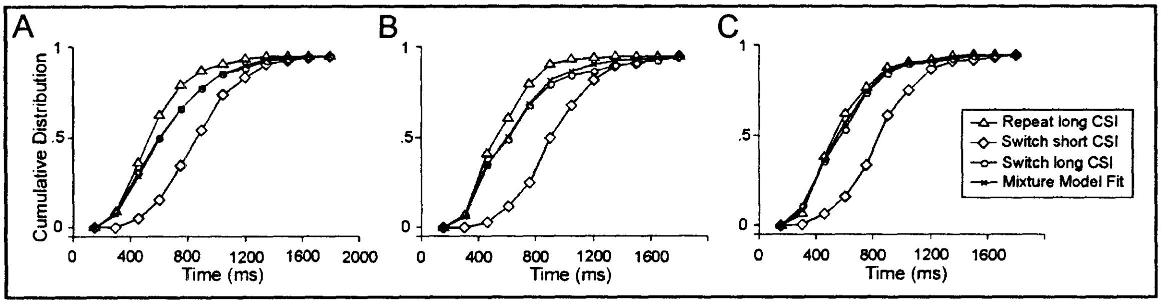

**Figure 7.** Cumulative distribution functions from mixture distribution simulations in CAM-TS. Using the same experimental parameters as (A) Dejong (2000), (B) Rogers and Monsell (1995; Experiment 5), and (C) Nieuwenhuis and Monsell (2002) CAM-TS produced distributions that appear to fit a mixture model, even though CAM-TS is not a binary state model and so is not appropriately characterized by a discrete mixture of two basis distributions. CAM-TS even produced an apparent shift toward the prepared distribution (C) by increasing the gain during the preparation interval from 2.0 to 2.3.

These simulations demonstrate that a quantitative alternative to a two-state mixture model gave rise to output that appears to conform to a two-state mixture model. Indeed, a gain increase in CAM-TS can account for the reduction in the estimated mixing probability under conditions of higher motivation. Though these simulations do not entirely rule out a mixture model of task switching, they do suggest that data arising from an alternative non-mixture structure could be successfully modeled as a mixture distribution using current procedures, and so raise concerns whether the extant evidence for intention activation truly provides an adequate test of the model.

It should be noted that the intention-activation model relies on a binary mixture, which is a special class of mixture distribution for which there are a number of distribution-based (Evritt, 1985) and distribution-free (Thomas, 1969) tests. Hence, rather than relying on MMLM which was designed to solve for the general case of a multinomial mixture, analytic tests can be derived for assessing intention-activation on task switching data. Without such tests, however, this theoretical framework and the regime under which it is being tested appears to be too unconstrained to garner much support.

### *Logan and Bundesen, 2003*

Recently, Logan and Bundesen (2003; also Logan, 2003; Arrington and Logan, 2005) have argued that the explicit cueing task, such as that used in Experiments 1 and 2 and modeled by CAM-TS, does not require endogenous control at all, but that task switching during the explicit cueing task may be fully accomplished via the conjunction of the task cue and target, and further that preparation effects during the explicit cueing task may be entirely accounted for by facilitated task cue encoding at longer preparation intervals. Logan and Bundesen (2003) proposed three models of task switching RT during repetition (Repeat) and alternation (Switch) trials, one that included a change in RT due to a set shifting control process, a second that included only a cue encoding process that was affected by the rate of comparison within short and long-term memory, and a third that combined the two models. The second model best fit the data from a set of experiments using the explicit cueing procedure, suggesting that no set shifting control process was necessary to obtain preparation effects during explicit cueing. Furthermore, these experiments and theoretical treatment demonstrated that a large portion of the switch cost could be attributed to facilitated encoding of the task cue, as switching the task cue resulted in RT slowing almost as costly as a task switch (Logan & Bundesen, 2003, 2004). Hence, it was argued that cue encoding determines the switch cost in the explicit cueing procedure rather than an endogenous control process that reconfigures the system, and further that facilitation of the task cue during a task repetition rather than interference of an alternation cue is the determining factor.

In several ways, CAM-TS is actually consistent with the perspective of Logan and Bundesen (2003). The implementation of a control process in CAM-TS is not what Logan and Bundesen (2003) argue against as an endogenous set shifting control process that operates to reconfigure the system during a task switch and consumes time during a serial information processing chain. In CAM-TS, the operation of control does not directly translate into a time consuming process. Rather, the switch cost is due to the difference, both facilitative and competitive, of long-term associations from trial to trial. Control is an increase in the gain on the Task layer during the preparation interval that reduces this time demand by biasing relevant representations in advance of interference in the Conceptual layer. This might be considered an enhancement in cue encoding to compensate for the lack of facilitation on alternation trials. Indeed, as with Logan and Bundesen (2003), CAM-TS ultimately suggests that the critical feature in preparation for switching tasks is how effectively the representation of the task cue, in the Task layer, can bias relevant Conceptual representations. This is helped along by the increase in gain.

The difference between CAM-TS and the Logan and Bundesen (2003) model is in the way in which cue encoding and memory retrieval are conceptualized. CAM-TS is based on associative models of memory in which presentation of a cue results in automatic retrieval of associated information to the extent that this information is strongly associated with that cue. Hence, under certain circumstances, other factors such as proactive interference or priming of competing representations can diminish the success of this bottom-up retrieval process. Under such circumstances, control processes are often required to guide retrieval of relevant information and select relevant information from competitors (Badre & Wagner, 2002; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997; Wagner, Paré-Blagoev, Clark, & Poldrack, 2001). Hence, in the Logan and Bundesen (2003) model, the rate at which memory comparison occurs with long-term (and short-term) memory may be complicated in its precise details, particularly with more stimuli than were used in their experiments.

Another difference between the models may concern what cue encoding refers to precisely. In CAM-TS, sufficient time for cue encoding means time for the maintained representation of the cue in working memory to filter down and automatically activate relevant members of the task set. By contrast, in Logan and Bundesen (2003), cue encoding may simply entail identification of the cue. Though, it should be noted that simple identification of the task cue is not associated with long-term carry-over effects (Sohn & Anderson, 2003). Nevertheless, the effect of a cue alternation would be similar between the models. If, for example, CAM-TS were to have four task nodes associated with the two tasks, switching among task nodes without otherwise switching tasks might result in simulated RT effects similar to switching tasks. This is because changing a task cue would not benefit from any priming facilitation among the weights of that task node and associated concepts. Indeed, it is likely that weights between the alternate task node and the conceptual node would be temporarily reduced, contributing to a substantially reduced facilitation on such a task repeat. Hence, cue encoding in CAM-TS might predict the same effects as in Logan and Bundesen (2003), but would emerge from the effectiveness of the nodes in the Task layer in activating or retrieving relevant information. As this is a mnemonic process, control is necessary in the same way it is necessary to regulate memory more generally.

## Conclusions

In this chapter, the operation and performance of CAM-TS was described, detailing how changes in the model's associative structure could give rise to task switching effects and further how control, implemented as an increase in gain in the Task layer, can result in preparation to overcome this interference. With this structure established, the performance of CAM-TS was characterized in a series of simulations that demonstrated its ability to capture a wide range of task switching phenomena. Interestingly, the model appeared unable to model those contexts in which task switching manipulations increased demands on goal setting or task specification, supporting the existence of an independent set of processes devoted to these control demands (Rubinstein et al., 2001). Finally, the relationship to three prominent models of task switching was considered and the similarities and differences considered.

An important test of any formal or informal model is that it generates novel predictions. CAM-TS provides an opportunity to make predictions about both the behavioral and neural impact of task switching. One potentially important class of predictions arising from CAM-TS concern the temporal dynamics of processing during task switching. CAM-TS predicts a specific temporal pattern of control and further suggests that selectivity demands shift at longer preparation intervals from Conceptual to

Response level demands. Based on the imaging data, one might further hypothesize this shift to occur between frontal and parietal cortices. To broadly test this prediction, data collection has recently been completed in a magnetoencephalogram (MEG) study using a similar paradigm to that described in Chapter 4. MEG permits precise temporal characterization of neural signals arising during task switching. Furthermore, using current adaptive-beamformer techniques of MEG data analysis should provide a relatively accurate and precise source localization capability (Sekihara, Nagarajan, Poeppel, Marantz, & Miyashita, 2002). Hence, this experiment will permit testing of the temporal assumptions of CAM-TS in addition to the hypotheses regarding the frontal-parietal dynamics emerging from the fMRI results.

Many of the behavioral phenomena associated with task switching may arise from performance-dependent changes in long-term associative memory. These changes result in conflict during retrieval of task-relevant information and so must be overcome for effective flexible performance. From the present perspective, control is required to overcome conflict in task switching and is expressed as an increased top-down bias signal that permits relevant representations to win out over competitors. Moreover, this process can be deployed strategically and in advance of a task stimulus, and so can permit preparation for an upcoming task switch. The theoretical perspective advanced here can account for many behavioral phenomena associated with task switching and provides a broad resolution to the current debate over the source and resolution of task switch costs. Furthermore, the control of associative memory perspective may be considered outside the specific domain of task switching, and makes contact with the broader literature surrounding cognitive control and the regulation of memory.

## References

Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and Performance XV: Conscious and nonconscious information processing* (pp. 421-452). Cambridge, MA: MIT Press.

Allport, A., & Wylie, G. (2000). 'Task-switching', stimulus-response bindings and negative priming. In S. Monsell & J. Driver (Eds.), *Control of Cognitive Processes: Attention and Performance XVIII* (pp. 35-70). Cambridge, MA: MIT Press.

Badre, D., & Wagner, A. D. (2002). Semantic retrieval, mnemonic control, and prefrontal cortex. *Behavioral and Cognitive Neuroscience Reviews, 1*(3), 206-218.

Berlyne, D. E. (1957). Uncertainty and conflict: A point of contact between information-theory and behavior-theory concepts. *Psychological Review, 64*, 329-339.

Bogacz, R., & Cohen, J. D. (2004). Parameterization of connectionist models. *Behavior Research Methods, Instruments, and Computers, 36*(4), 732-741.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol Rev, 108*(3), 624-652.

Boynton, G. M., Engel, S. A., Glover, G. H., & Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. *Journal of Neuroscience, 16*, 4207-4221.

Brass, M., & von Cramon, D. Y. (2004). Selection for cognitive control: a functional magnetic resonance imaging study on the selection of task-relevant information. *J Neurosci, 24*(40), 8847-8852.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97*, 332-361.

Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: A connectionist approach to behavior and biology in schizophrenia. *Psychological Review, 99*, 45-77.

Dejong, R. (2000). An intention-activation account of residual switch costs. In S. Monsell & J. Driver (Eds.), *Control of Cognitive Processes: Attention and Performance XVIII* (pp. 357-376). Cambridge, MA: MIT Press.

Evritt, B. S. (1985). Mixture distributions. In S. Kotz & N. L. Johnson (Eds.), *Encyclopedia of Statistical Sciences* (Vol. 5, pp. 559-569). New York: Wiley.

Gilbert, S. J., & Shallice, T. (2002). Task switching: a PDP model. *Cognit Psychol, 44*(3), 297-337.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc Natl Acad Sci U S A, 79*, 2554-2558.

Jersild, A. T. (1927). Mental set and shift. *Archives of Psychology*(Whole No. 89).

Koch, I., Prinz, W., & Allport, A. (2005). Involuntary retrieval in alphabet-arithmetic tasks: Task-mixing and task-switching costs. *Psychol Res, 69*(4), 252-261.

Logan, G. D., & Bundesen, C. (2003). Clever homunculus: is there an endogenous act of control in the explicit task-cuing procedure? *J Exp Psychol Hum Percept Perform, 29*(3), 575-599.

Logan, G. D., & Bundesen, C. (2004). Very clever homunculus: compound stimulus strategies for the explicit task-cuing procedure. *Psychon Bull Rev, 11*(5), 832-840.

Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychol Rev, 108*(2), 393-434.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychol Bull, 109*(2), 163-203.

Mayr, U., & Kliegl, R. (2000). Task-set switching and long-term memory retrieval. *J Exp Psychol Learn Mem Cogn, 26*(5), 1124-1140.

Mayr, U., & Kliegl, R. (2003). Differential effects of cue changes and task changes on task-set selection costs. *J Exp Psychol Learn Mem Cogn, 29*(3), 362-372.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review, 88*, 375-407.

Meiran, N. (2000). Modeling cognitive control in task-switching. *Psychological Research, 63*, 234-249.

Meiran, N., Chorev, Z., & Sapir, A. (2000). Component processes in task switching. *Cognit Psychol, 41*(3), 211-253.

Monsell, S. (2003). Task switching. *Trends Cogn Sci, 7*(3), 134-140.

Monsell, S., Sumner, P., & Waters, H. (2003). Task-set reconfiguration with predictable and unpredictable task switches. *Mem Cognit, 31*(3), 327-342.

Nieuwenhuis, S., & Monsell, S. (2002). Residual costs in task switching: Testing the 'failure to engage' hypothesis. *Psychon Bull Rev, 9*, 86-92.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124*(2), 207-231.

Rogers, R. D., Sahakian, B. J., Hodges, J. R., Polkey, C. E., Kennard, C., & Robbins, T. W. (1998). Dissociating executive mechanisms of task control following frontal lobe damage and Parkinson's disease. *Brain, 121 ( Pt 5)*, 815-842.

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *J Exp Psychol Hum Percept Perform, 27*(4), 763-797.

Sekihara, K., Nagarajan, S. S., Poeppel, D., Marantz, A., & Miyashita, Y. (2002). Application of an MEG eigenspace beamformer to recontructing spatio-temporal activities ofneural sources. *Hum Brain Mapp, 15*(4), 199-215.

Sohn, M.-H., & Anderson, J. R. (2003). Stimulus-related priming during task-switching. *Memory and Cognition, 31*(5), 775-780.

Spector, A., & Biederman, I. (1976). Mental set and mental shift revisited. *American Journal of Psychology, 89*, 669-679.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643-662.

Thomas, E. A. C. (1969). Distribution free tests for mixed probability distributions. *Biometrika, 56*(3), 475-484.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proc Natl Acad Sci U S A, 94*(26), 14792-14797.

Wagner, A. D., Paré-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning: Left prefrontal cortex guides controlled semantic retrieval. *Neuron, 31*, 329-338.

Wylie, G. R., Javitt, D. C., & Foxe, J. J. (2003). Cognitive control processes during an anticipated switch of task. *Eur J Neurosci, 17*(3), 667-672.

Yantis, S., Meyer, D. E., & Smith, J. E. (1991). Analyses of multinomial mixture distributions: new tests for stochastic models of cognition and action. *Psychol Bull, 110*(2), 350-374.

Yeung, N., Cohen, J. D., & Botvinick, M. M. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychol Rev, 111*(4), 931-959.

# Chapter 6

## Conclusions

Switching from performance of a given task to a new task, as the context demands, is an important instance of the flexible behavior, both in our remarkable capacity to accomplish the shift and in the difficulty we experience when doing so. The experiments and theoretical framework presented in this thesis argue from strong evidence that task switching may be understood as an act of memory, and that PFC control mechanisms, particularly those in left mid-VLPFC, are critical in overcoming interference arising from automatically retrieved competing representations during a task switch. In reviewing the cognitive and neurobiological literature on task switching, Chapter 1 introduced the respective theoretical positions that have framed the debate over the sources of task switch costs and the putative mechanisms that are critical in resolving these costs. From this review, it was argued that memory places important constraints on task switching. However, different theoretical perspectives emphasize different aspects of mnemonic performance as the determining factor in task switching. From the perspective of TSR, reconfiguration might entail a process that accesses long-term memory in order to retrieve relevant representations for the upcoming task (Monsell, Sumner, & Waters, 2003; Rogers & Monsell, 1995). Such a process might consume additional time prior to a task switch, thus giving rise to a switch cost. By contrast, the TSI and TSP hypotheses note that active representations from a competing task may interfere with performance of the presently relevant task, giving rise to the task switch cost (Allport, Styles, & Hsieh, 1994; Allport & Wylie, 2000). This interference can arise either transiently because of carry-over from recent activation of these representations during performance of the competing task (TSI) or because they were automatically retrieved by cues encountered during performance of a competing task (TSP). Interestingly, evidence from neuroimaging studies highlighted left VLPFC as a region of PFC commonly engaged differentially during a task switch (e.g. Brass & von Cramon, 2004; Dove, Pollmann, Schubert, Wiggins, & von Cramon, 2000; Dreher, Koechlin, Ali, & Grafman, 2002). Given the hypothesized role of VLPFC in the cognitive control of memory, formal investigation of the contributions of VLPFC to task switching promised to provide leverage on the debate over the sources and mechanisms of task switching. In particular, it was hypothesized that mnemonic control processes enacted by left VLPFC may be particularly important under conditions of interference from memory, akin to those proposed by TSP. The experiments described in the remaining chapters sought to test this hypothesis directly.

Chapter 2 provided an initial test of the hypothesis that left VLPFC is critical to resolving a form of proactive interference similar to that proposed to occur during task switching. However, in this experiment, this proactive interference effect was tested during simple short-term item recognition, and so did not involve task switching, *per se*. On critical trials of this experiment, a target recently encoded in the memory set of a previous trial was encountered again as a target for the current trial (i.e. *recent* target). Proactive interference was evident in a longer RT for these trials, particularly when the target was not a member of the current set and so required a negative response despite its familiarity. From one perspective, proactive interference occurs in this task because presentation of the recent target can cue retrieval of contextual information associated with this target on the preceding trial. This retrieved information results in conflict that must be overcome to select the relevant information needed to generate a response. Critically, this automatic retrieval, and so conflict, will occur on both positive and negative trials, a hypothesis consistent with data from fMRI showing increases in left mid-VLPFC on both positive and negative recent trials (Fig. 1C). However, potentially due to the relative weakness of the relevant information during a negative trial, this conflict is greater during negative trials, evident in an interaction in left VLPFC. These results provided an initial rationale that left mid-VLPFC may be required to control memory under conditions of proactive interference, similar to those proposed for task switching.

In Chapter 3, the processes by which left VLPFC controls memory, and so may be engaged during task switching, were further specified and dissociated. In particular, the results reported in this chapter provide evidence for a controlled retrieval process in left anterior VLPFC (inferior frontal gyrus *pars orbitalis*) that guides retrieval, and so operates to bias activation of long-term representations themselves (Figure 1E). Furthermore, this controlled retrieval process was dissociated from a post-retrieval selection process in left mid-VLPFC (inferior frontal gyrus *pars triangularis* and *pars opercularis*) that biases selection of relevant retrieved representations over interference from competitors (Figure 1D & 1E). Distinguishing between these processes was important, as both could potentially contribute to successful performance of a task switch. More specifically, controlled retrieval processes might be required to guide retrieval of relevant representations during task set reconfiguration (i.e. TSR). By contrast, retrieval of the relevant task set may proceed fairly automatically upon encoding of the relevant task cue. However, automatic retrieval of competing representations may cause interference (i.e. TSP) and so necessitates post-retrieval selection of the relevant representations en route to a response. Interestingly, the association of post-retrieval selection with left mid-VLPFC, the region also highlighted in Chapter 2, motivated the specific hypothesis that left mid-VLPFC acts to select relevant representations over retrieved competitors during task switching.

Chapter 4 sought to test this hypothesis directly. To be theoretically explicit, this chapter introduced a computational model, CAM-TS, from which task switching costs arise due to (1) conflict arising from feedback-derived activation of nodes unassociated with the relevant task and (2) to priming of these irrelevant associations due to small performance induced changes in their connection weights. Control in CAM-TS acts in manner similar to the post-retrieval selection process hypothesized for left mid-VLPFC by generally increasing top-down control during the preparation interval and so biasing selection of the relevant retrieved representations over the competitors. Input to this control process is thought to be conflict among retrieved conceptual representations. Hence, the model provided a signature of conceptual conflict, showing a decline in the conflict value over longer preparation intervals, and also an enhancement due to proactive interference from a repeated response. Critically, left mid-VLPFC (Figure 1A & 1B) showed a similar decline in activation over CSI and also showed evidence of an enhancement from response repetition, linking this region to a control process sensitive to conceptual conflict. Furthermore, this pattern of data dissociated left mid-VLPFC from SMA and inferior parietal cortices. Indeed, an unanticipated, but intriguing, finding from this experiment was the relationship between inferior parietal cortex and conflict in the Response layer of the model, with both showing an increase over longer CSI intervals.

Considered together, the results from these experiments provide a strong empirical grounding for the CAM-TS model and for a resolution of the debate over task switching. As depicted in Figure 1, there is a high degree of correspondence in the region of left mid-VLPFC highlighted in each of these independent data sets. Indeed, if switch differences from the task switching experiment (Chapter 4) are computed from an ROI defined from the Selection-Component region defined from Chapter 3, a decline in the Switch versus Repeat difference is evident (Figure 1). By contrast, an ROI also defined from Chapter 3 but in anterior VLPFC (*pars orbitalis*), associated with controlled retrieval, reveals little evidence of such a CSI-dependent decline. Beyond providing

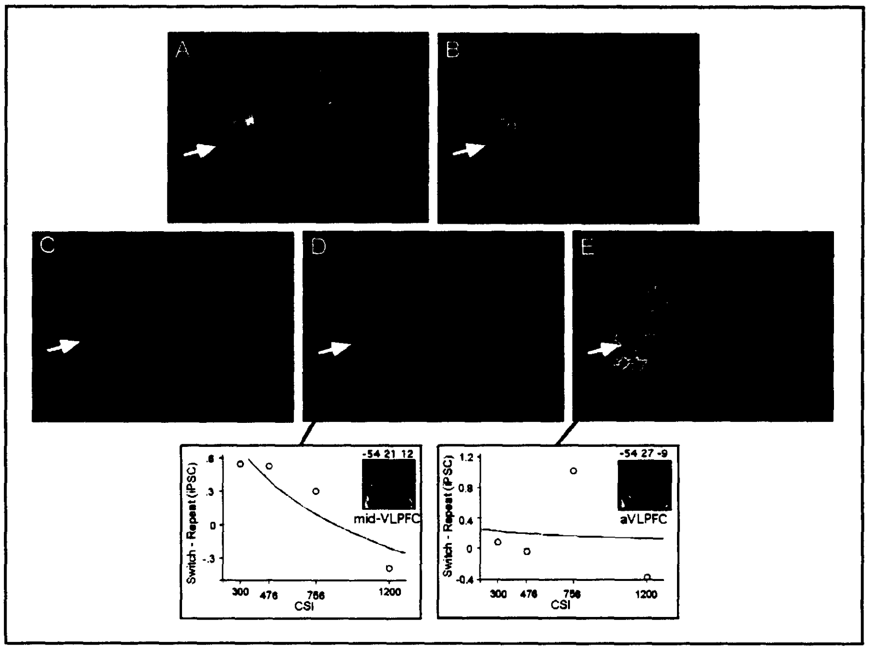

**Figure 1.** Surface rendered results from all experiments demonstrate the convergence of activation in left mid-VLPFC in response to selection demands (arrow in same location for reference). Activation maps are rendered on an inflated MNI canonical for (A) Switch versus Repeat (Chapter 4), (B) Switch versus Repeat Short-CSIs (Chapter 4), (C) Negative Recent versus Non-Recent (Chapter 2), (D) Selection-Component (Chapter 3), and (E) overlap (purple) of Feature versus Related (red) and Weak versus Strong Associative Strength (blue; Chapter 3). Also charted (bottom) are the fitted declines in Switch versus Repeat activation over CSI for mid-VLPFC and anterior VLPFC ROIs defined from Chapter 3. In addition to underscoring the dissociation between these subregions, the differential decline in left mid-VLPFC supports the contribution of a post-retrieval selection process to overcome interference in task switching rather than controlled retrieval of a task set.

further support for the dissociation between these subregions of VLPFC as argued in Chapter 3, this difference provides additional evidence in favor of selection due to TSP rather than guided retrieval during TSR. Rather than placing demands on retrieval itself, which would also modulate activation in left anterior VLPFC, task switching results in retrieval of irrelevant information that must be selected against, and so requires processing in mid-VLPFC.

Finally, the power of this present theoretical perspective, made explicit by CAM-TS, to account for a large portion of task switching phenomena was demonstrated in Chapter 5. In general, any task switching effect that could be accounted for as (1) an increase in the efficacy of a stimulus to cue automatic retrieval of irrelevant information and so increase conflict or (2) a decrease in the efficacy of the task cue to provide a top down signal selecting relevant representations over competitors, was modeled well by CAM-TS. A class of phenomena not modeled by CAM-TS were those relating to a goal setting process that determines what task is appropriate to perform next, in essence determining the input to the Task layer of CAM-TS. Further investigation may provide important insights into this additional component of task switching and so further refinement of theories like CAM-TS. However, derived from its strong empirical support, CAM-TS formalizes a powerful theory of interference and cognitive control in task switching, which emphasizes the processes that control memory to overcome interference during flexible performance.

## Controlling Memory and Resolving Interference: Prefrontal Contributions to Flexible Behavior

Broadly construed, the theoretical framework for understanding task switching supported in this thesis also provides insights more generally into the centrality of mnemonic control in enabling flexible behavior. Beyond providing evidence of a two-stage model of mnemonic control and further demonstrating that this control is the central explanatory construct in at least one important instance of human performance, these experiments and associated theory may encourage a more general understanding of the demands to which prefrontal cortex and cognitive control are responsive.

In particular, these findings highlight the significant impact of conflict from even weak, non-viable competitors; an instance of conflict not often emphasized in experimental settings or even in theoretical models in which it is implicit (e.g., Botvinick, Braver, Barch, Carter, & Cohen, 2001; Yeung, Cohen, & Botvinick, 2004). Traditional examples of conflict cited in the literature arise from competitors that, if not kept in check, will come to dominate behavior. For example, the oft-cited Stroop task (MacLeod, 1991; Stroop, 1935) is one in which the dominant word stimulus may drive behavior should the subject not maintain vigilant determination to perform the color task. Likewise, the most compelling examples from patients with frontal lobe dysfunction are those in which dominant representations appear to control behavior and run roughshod over less viable but relevant representations or cues. For example, perseveration in the Wisconsin Card Sorting Task—wherein the patient is unable to disengage from a previously endorsed response pattern—or utilization behavior —by which the patient is compelled to perform a strongly associated behavior with an item regardless of how contextually inappropriate it is—are commonly cited as evidence for the necessity of prefrontal cortex to guide behavior lest performance become entirely governed by the tendencies of habit and strong association (e.g., "environmental dependency syndrome" Lhermitte, 1983; Lhermitte, Pillon, & Serdaru, 1986).

However, many findings from patients with frontal damage are not nearly so dramatic. Often patients are entirely competent to perform the tasks required of them, but they simply demonstrate considerable difficulty in doing so. For example, patients with left lateral PFC lesions performing task switching do not typically perseverate on one task regardless of the task cue, rather they simply show greatly enhanced switch costs relative to controls (Rogers et al., 1998). Likewise, a patient with damage inclusive of left mid-VLPFC was not compelled to indicate that a presented target was contained in the current target set because it was familiar, having been presented in the previous target set. Rather, this patient simply showed a greatly enhanced RT interference effect relative to controls for negative non-recent targets (Thompson-Schill et al., 2002). Patients performing semantic tasks in which they must select relevant semantic information from amidst competitors are not unable to perform the task at all nor are they compelled to repetitively generate the most associated representation. Rather they are much slower (Swick & Knight, 1996) and generate more errors (Thompson-Schill et al., 1998) than controls. What these studies demonstrate, in experimental contexts similar to those tested in the present thesis, is that cognitive control is not simply required in cases in which, were it to be absent, the system would be in danger of being entirely controlled by a dominant, but irrelevant, competitor. Rather control is also important in contexts, likely more common, in which the system would successfully converge on the appropriate response in the absence of prefrontal control, but it is encumbered by noise, in the form of weak but active competitors.

In CAM-TS, the activation of competing representations through feedback is far too weak to drive behavior. Rather, the mild activation in these nodes simply has the effect of weakening the activation of the relevant, and dominant, representations with which they hold mutually inhibitory connections. This slight weakening will be expressed in behavior as a longer RT, but will not lead to domination by the irrelevant representation, even in the absence of a prefrontal cortex. Hence, the critical insight is that any activated competitor, whether viable and dominant or weak and insignificant, will increase conflict and will impede behavior. However, this effect is not trivial. As indicated by the task switch cost, when engaged in the complex behaviors required in everyday life, such subtle effects of conflict can accumulate to produce the severe impediments to normal function of which frontal patients frequently complain. Thus, control can operate, and likely is operating most of the time, to dynamically maintain a balance of activity in response to even small increases in conflict due to weakly activated representations, and so facilitates efficient selection of the relevant representations.

The source of weak competitors in task switching is automatic retrieval from long-term associative memory. Similarly, any instance of flexible performance may require access to memory to inform responses, interpret cues, and comprehend stimuli. However, the structure of memory is such that activation of a given representation will always result in some obligatory activation of other representations, be they relevant or irrelevant. Activation of this irrelevant information, however weak, will then produce interference, and so will demand control. Indeed, interference of this kind is likely a pervasive control demand during performance of any task. Thus, to the extent that access to memory is necessitated for performance, control of memory by prefrontal cortex is likely a fundamental requirement of flexible behavior.

## References

Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and Performance XV: Conscious and nonconscious information processing* (pp. 421-452). Cambridge, MA: MIT Press.

Allport, A., & Wylie, G. (2000). 'Task-switching', stimulus-response bindings and negative priming. In S. Monsell & J. Driver (Eds.), *Control of Cognitive Processes: Attention and Performance XVIII* (pp. 35-70). Cambridge, MA: MIT Press.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol Rev, 108*(3), 624-652.

Brass, M., & von Cramon, D. Y. (2004). Decomposing components of task preparation with functional magnetic resonance imaging. *J Cogn Neurosci, 16*(4), 609-620.

Dove, A., Pollmann, S., Schubert, T., Wiggins, C. J., & von Cramon, D. Y. (2000). Prefrontal cortex activation in task switching: an event-related fMRI study. *Brain Res Cogn Brain Res, 9*(1), 103-109.

Dreher, J. C., Koechlin, E., Ali, S. O., & Grafman, J. (2002). The roles of timing and task order during task switching. *Neuroimage, 17*(1), 95-109.

Lhermitte, F. (1983). 'Utilization behaviour' and its relation to lesions of the frontal lobes. *Brain, 106 (Pt 2)*, 237-255.

Lhermitte, F., Pillon, B., & Serdaru, M. (1986). Human autonomy and the frontal lobes. Part I: Imitation and utilization behavior: a neuropsychological study of 75 patients. *Ann Neurol, 19*(4), 326-334.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychol Bull, 109*(2), 163-203.

Monsell, S., Sumner, P., & Waters, H. (2003). Task-set reconfiguration with predictable and unpredictable task switches. *Mem Cognit, 31*(3), 327-342.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124*(2), 207-231.

Rogers, R. D., Sahakian, B. J., Hodges, J. R., Polkey, C. E., Kennard, C., & Robbins, T. W. (1998). Dissociating executive mechanisms of task control following frontal lobe damage and Parkinson's disease. *Brain, 121 ( Pt 5)*, 815-842.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643-662.

Swick, D., & Knight, R. T. (1996). Is prefrontal cortex involved in cued recall? A neuropsychological test of PET findings. *Neuropsychologia, 34*(10), 1019-1028.

Thompson-Schill, S. L., Jonides, J., Marshuetz, C., Smith, E. E., D'Esposito, M., Kan, I. P., et al. (2002). Effects of frontal lobe damage on interference effects in working memory. *Cogn Affect Behav Neurosci, 2*(2), 109-120.

Thompson-Schill, S. L., Swick, D., Farah, M. J., D'Esposito, M., Kan, I. P., & Knight, R. T. (1998). Verb generation in patients with focal frontal lesions: a neuropsychological test of neuroimaging findings. *Proc Natl Acad Sci U S A, 95*(26), 15855-15860.

Yeung, N., Cohen, J. D., & Botvinick, M. M. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychol Rev, 111*(4), 931-959.