Lesion Evidence That Two Distinct Regions within Prefrontal Cortex are Critical for *n*-Back Performance in Humans

Ami Tsuchida and Lesley K. Fellows

Abstract

■ Although prefrontal cortex is clearly important in executive function, the specific processes carried out by particular regions within human prefrontal cortex remain a matter of debate. A rapidly growing corpus of functional imaging work now implicates various areas within prefrontal cortex in a wide range of ''executive'' tasks. Loss-of-function studies can help constrain the interpretation of such evidence by testing to what extent particular brain areas are necessary for a given cognitive process. Here we apply a component process analysis to understand prefrontal contributions to the *n*-back task, a widely used test of working memory, in a cohort of patients with focal prefrontal damage. We investigated letter 2-back task performance in 27 patients with focal damage to various regions within prefrontal cortex, compared to 29 demographically matched control subjects. Both ''behavior-defined'' approaches, using qualitative lesion analyses and voxel-based lesion–symptom mapping methods, and more conventional ''lesion-defined'' groupwise comparisons were undertaken to determine the relationships between specific sites of damage within prefrontal cortex and particular aspects of *n*-back task performance. We confirmed a critical role for left lateral prefrontal cortex in letter 2-back performance. We also identified a critical role for medial prefrontal cortex in this task: Damage to dorsal anterior cingulate cortex and adjacent dorsal fronto-medial cortex led to a pattern of impairment marked by high false alarm rates, distinct from the impairment associated with lateral prefrontal damage. These findings provide converging support for regionally specific models of human prefrontal function. ■

INTRODUCTION

Prefrontal cortex (PFC) is important in ''executive function,'' a term referring to a range of supervisory processes. Working memory is generally viewed as a linchpin of executive function, providing the ability to work with information in the absence of a stimulus. This ability would seem crucial to many higher-order strategic functions, and has long been linked to the frontal lobes (Goldman-Rakic, 1987; Fuster, 1985). The simplest definition of working memory is the ability to maintain information across a delay, an ability canonically captured by delayed match-to-sample tasks. Neurophysiological studies have shown that neurons in dorsolateral PFC (DLPFC) encode stimulus information across delays (Funahashi, Bruce, & Goldman-Rakic, 1989), and disruption of DLPFC function impairs performance of delayed match-to-sample tasks, arguing that this area of the frontal lobes plays a critical role in working memory in nonhuman primates (Butters & Pandya, 1969).

Damage to lateral PFC in humans does not disrupt simple span or delayed match-to-sample tests of working memory as reliably: A systematic review of published reports found that more than half of the experiments in such patients failed to find significant deficits in delayed response tasks (D'Esposito & Postle, 1999). Several of the reviewed studies reported working memory deficits in some patients but not in others, or only under certain conditions (Bechara, Damasio, Tranel, & Anderson, 1998; Ptito, Crane, Leonard, Amsel, & Caramanos, 1995). However, working memory tasks that require more than simple retention of information (e.g., manipulation and monitoring of information, or active retention in the face of distractors during the delay period) seem to be more sensitive to PFC damage in humans (D'Esposito & Postle, 1999; Petrides & Milner, 1982).

The *n*-back task is an example of such a working memory paradigm, requiring active comparison of the current stimulus to information held in memory, and concomitant manipulation or ''updating'' of relevant information in memory when $n > 1$. Participants are presented with a series of stimuli, one at a time, and must continuously judge if the stimulus is the same as the one presented n trials previously. The paradigm was developed to study working memory with functional imaging, and has been widely used for that purpose (Cohen et al., 1993, 1997). This work has yielded a general consensus that the *n*-back task recruits ventrolateral PFC (VLPFC) and DLPFC, with the latter more sensitive to increased load, that is, when $n > 1$ (Hautzel et al., 2002; Nystrom

Montreal Neurological Institute, Canada

© 2008 Massachusetts Institute of Technology

et al., 2000; Courtney, Petit, Haxby, & Ungerleider, 1998; Braver et al., 1997; Jonides et al., 1997). In addition, medial PFC (particularly dorsal anterior cingulate cortex [dACC]) and posterior parietal cortex are commonly activated in the *n*-back task and are also sensitive to parametric changes in load (McMillan, Laird, Witt, & Meyerand, 2007; Owen, McMillan, Laird, & Bullmore, 2005; Cohen et al., 1997; Jonides et al., 1997).

The interpretation of these activations has relied on evidence from functional imaging studies using different tasks, and on various pieces of converging evidence from nonhuman primate work, transcranial magnetic stimulation (TMS), and lesion studies in humans (see below). Lateral PFC is thought to be important for the higher-order control processes within working memory that the task was intended to tap. The roles played by medial PFC and parietal cortex are less well defined. Based primarily on functional imaging results, it has been speculated that posterior parietal cortex activation reflects storage of perceptual attributes (Callicott et al., 1999; Jonides et al., 1997), with dACC somehow important in the allocation of effort or cognitive control (Carter et al., 1998; Paus, Koski, Caramanos, & Westbury, 1998; Barch et al., 1997).

Despite the imaging findings in support of a network of brain areas recruited during *n*-back performance, the task is increasingly being employed as a specific measure of lateral PFC-mediated working memory ability in contexts ranging from schizophrenia (Thermenos et al., 2005; Perlstein, Carter, Noll, & Cohen, 2001; Callicott et al., 2000) to aging (Mattay et al., 2006). The claim that *n*-back performance reflects lateral PFC function is an example of reverse inference, a logic that has several pitfalls, particularly when applied to relatively complex cognitive functions (such as those engaged by the *n*-back task), or to regions of the brain that are implicated in many cognitive processes (such as lateral PFC; Duncan & Owen, 2000) (Poldrack, 2006).

Two questions need to be answered to more directly link *n*-back task performance and its underlying neural substrates, and neither can be definitively addressed with functional imaging methods: First, is lateral PFC critical to *n*-back performance? And second, is *n*-back performance sensitive to dysfunction in other regions of the brain? The answers to these questions have relevance for our basic understanding of the component processes tapped by the *n*-back task and the neural substrates that underlie them. They are also relevant to interpreting studies in which the *n*-back task is used as a neurobehavioral measure. These questions are best addressed with loss-of-function methods, such as lesion or TMS studies, which can test whether a particular brain region is necessary for task performance (Fellows et al., 2005; Rorden & Karnath, 2004; Walsh & Cowey, 2000). A handful of TMS studies support a role for lateral PFC in *n*-back performance, but disagree on whether only left, or both left and right DLPFC are critical, and have not evaluated medial PFC regions (Mottaghy, Gangitano, Krause, & Pascual-Leone, 2003; Mull & Seyal, 2001; Mottaghy et al., 2000).

The neuropsychological evidence is also inconsistent and has also not addressed the potential role of medial PFC. A case study reported number 2-back task impairment in a single patient with right DLPFC damage secondary to traumatic brain injury (Ptak & Schnider, 2004), and a group study reported spatial and object 2-back impairment in four patients with damage affecting broad regions of (primarily left) lateral PFC, but not ventromedial PFC ($n = 5$) or restricted regions in DLPFC ($n = 6$) (Muller, Machado, & Knight, 2002). In contrast, a more recent study reported that left superior frontal gyrus damage ($n = 8$), and not more ventrally located lateral PFC damage ($n = 5$) or parietal damage ($n = 4$), disrupted spatial, face, and letter 2- and 3-back task performance (Boisgueheneuc et al., 2006). The explanation for these conflicting results is not clear. Small sample sizes are a perennial challenge in this work, and differences in the boundaries used in defining subgroups based on location of damage make it difficult to synthesize findings across studies.

In order to clarify the specific contributions of prefrontal regions to performance of the *n*-back task, we administered a letter 2-back task to a larger sample of patients with chronic, focal prefrontal lesions that varied in size and location. Their performance, measured in terms of types of errors made as well as with two signal detection measures, the discriminability index (d') and response criterion (C), was compared to healthy age- and education-matched control subjects.

The relation between performance and lesion location was examined using overlap and voxel-based lesion–symptom mapping (VLSM) methods (Bates et al., 2003). VLSM provides a statistical evaluation of the relationship between behavior and lesion location, does not require patients to be assigned to potentially arbitrary lesion groups, and allows performance to be considered as a continuous variable. The statistical maps generated with this method complement functional imaging data, identifying brain regions that are critical for task performance (Rorden, Karnath, & Bonilha, 2007; Bates et al., 2003). This method has been successfully used to delineate the brain regions necessary for specific component processes of language comprehension (Tyler, Marslen-Wilson, & Stamatakis, 2005; Saygin, Wilson, Dronkers, & Bates, 2004). In addition to these ''behavior-defined'' approaches, a more conventional secondary analysis was carried out, comparing performance in patients classified into subgroups according to lesion location.

In summary, we aimed to identify the regions within PFC that are necessary for intact performance of this widely used working memory paradigm. Further, we examined the behavioral profile of those who were found to be impaired, in order to specify the component processes subserved by the regions we found to be critical.

METHODS

Subjects

Patients with focal damage to the frontal lobes ($n = 27$) were identified through the research databases of the Center for Cognitive Neuroscience at the University of Pennsylvania and McGill University. The group included 12 patients with ischemic or hemorrhagic stroke, 9 who had undergone resection of low-grade tumors, and 6 with damage due to cerebral aneurysm rupture. Thirteen subjects were taking psychoactive medications. These were most commonly anticonvulsants or antidepressants. Subjects were tested at least 6 months (mean = 5.9 years, range = 10 months to 16 years) after brain injury had occurred.

Age- and education-matched control subjects ($n = 29$) were recruited by local advertisement in Montreal. Normal controls were not taking psychoactive medication and had no history of neurological or psychiatric disease. They were excluded if they scored less than 28/30 on the Mini-Mental Status Examination (MMSE) (Folstein, Folstein, & McHugh, 1975) or less than 26/30 on the Montreal Cognitive Assessment (MoCA) (Nasreddine et al., 2005). IQ was estimated with the American version of National Adult Reading Test (ANART). All subjects provided written, informed consent prior to participation in the study, in accordance with the Declaration of Helsinki, and were paid a nominal fee for their time. The study protocol was approved by the Institutional Review Boards of the University of Pennsylvania and McGill University.

Table 1 provides demographic information for all participants. Unpaired t test revealed no significant difference between frontal patients and controls for age or education, as intended ($p > .05$). The frontal group as a whole had significantly lower estimated premorbid IQ, as well as higher BDI scores compared to control subjects ($p < .05$).

All participants completed a brief screening neuropsychological evaluation, to verify that cognitive abilities of no interest in this study but important to task performance, such as language comprehension, were not contributing to the effects of interest, and to characterize potentially relevant frontally mediated abilities. Tasks included letter and category verbal fluency tests, sentence comprehension, backward digit span, and a modified, computerized version of the backward Corsi block tapping task. Table 2 summarizes the results of these tests. There were no significant differences between the frontal patients and control subjects for scores on sentence comprehension, phonologic fluency, or the backward Corsi block tapping task (Mann–Whitney test, $p > .05$), but the frontal group did perform more poorly on semantic fluency and backward digit-span tasks (Mann–Whitney test, $p < .05$).

Image Analysis

Individual lesions were traced from the most recent clinical imaging onto the standard Montreal Neurological Institute brain using MRIcro software (Rorden & Brett, 2000) by a neurologist experienced in image analysis and blind to task performance. The same software was used to estimate lesion volumes as well as to generate overlap images for groups defined according to task performance.

2-Back Task

A computerized letter 2-back task was created using E-Prime (www.pstnet.com). A fixed pseudorandom sequence of letters (any letter from A to Z) was presented centrally on the computer screen, one letter at a time. Each letter was presented for 500 msec, with an interstimulus interval of 1500 msec. The target was any letter repeated after one intervening letter. Participants were asked to respond only to targets by pressing the space bar as quickly as possible. There were two types of possible errors: misses (i.e., missed targets) and false alarms (i.e., incorrect responses to nontargets). A brief practice sequence of 10 letters (with two targets) was given

Table 1. Demographic and Background Information for Control (CTL) and Frontal Patients (FP) (Mean (*SD*))

Group	*Age (years)*	*Education (years)*	*Estimated IQ*	*Beck Depression Inventory*	*Lesion Laterality (R/L/B)*	*Frontal Lesion Volume (cc)*
CTL ($n = 29$)	55.1 (10.1)	14.7 (3.1)	122.6 (7.7)	6.2 (5.5)	–	–
FP ($n = 27$)	57.1 (11.3)	13.6 (3.4)	115.1 (9.8)*	11.4 (9.3)*	8/8/11	33.7 (30.4)
LL ($n = 4$)	52.5 (7.3)	13.7 (1.0)	105.8 (5.4)	6.5 (5.1)	0/4/0	44.3 (19.5)
RL ($n = 5$)	61.0 (13.4)	14.0 (2.6)	116.4 (10.6)	16.0 (10.0)	4/0/1	48.8 (31.5)
MF ($n = 11$)	55.1 (12.2)	13.9 (3.5)	119.5 (8.0)	9.8 (9.1)	3/3/5	28.5 (36.5)
OF ($n = 7$)	60.7 (11.2)	12.6 (4.7)	111.7 (10.2)	13.6 (10.5)	1/1/5	25.1 (23.1)

This information is also shown for subgroups of FP, divided according to main site of damage: left lateral (LL), right lateral (RL), medial frontal (MF), and orbitofrontal (OF). See text for details.

Not all subjects completed the ANART.

*Denotes significant differences (ANOVA; $p < .05$).

Table 2. Performance on Neuropsychological Screening Tests for CTL and FP Groups

Group	*Sentence Comprehension Accuracy*	*F′ Fluency*	*Animal Fluency*	*Digit Span Backward*	*Backward Corsi Score*
CTL	0.98 (0.06)	12.7 (3.8)	20.5 (4.2)	4.9 (1.3)	7.5 (2.9)
FP	0.97 (0.04)	10.4 (5.1)	16.3 (5.2)*	4.2 (1.7)*	5.9 (3.3)
LL	0.95 (0.05)	6.3 (3.9)	13.0 (6.1)	3.0 (0.8)	7.3 (1.1)
RL	1.00 (0.00)	13.0 (2.6)	18.6 (4.6)	4.6 (0.9)	5.2 (3.7)
MF	0.97 (0.05)	10.4 (5.6)	17.0 (5.0)	4.6 (1.6)	6.3 (3.8)
OF	0.96 (0.05)	11.1 (5.3)	15.3 (5.4)	3.9 (0.7)	5.0 (3.1)

This information is also shown for subgroups of FP, divided according to main site of damage.

*Denotes significant differences (ANOVA; $p < .05$).

before the actual test. Feedback was given for these trials, and the practice sequence was repeated until the participant correctly identified both targets in order to ensure that the instructions had been understood. The test sequence consisted of 122 letters, with infrequent target letters (20/122 trials). There were no 1-back or 3-back lures.

Signal Detection Analysis

The frequency of omission and commission errors, transformed to adjust for unequal numbers of target and nontargets (Corwin, 1994), was used to calculate two signal detection measures: a logistic equivalent discriminability measure (d'_L) and a response bias measure (C_L). Logistic rather than normal distributions were used to calculate these measures because it simplifies the calculation without affecting how these measures behave (Corwin, 1994). The response bias measure C, or response criterion, was chosen over another commonly used response bias measure β because C is less affected by changes in d' (Stanislaw & Todorov, 1999). Briefly, these signal detection measures describe two aspects of performance in tasks involving discrimination of two stimulus types. The discriminability measure captures the ability of a subject to discriminate between the two stimulus types, in this case, the 2-back target letter and nontarget. When subjects fail to keep track of, or to update, relevant information (letters being presented and the order of presentation) in working memory, it leads to a low discriminability score. This measure takes into account the fact that poor ability to discriminate target from nontarget can be manifested behaviorally as missing a target (omission error) or as mistaking a nontarget for a target (commission error). Faced with a similar difficulty in identifying targets, one subject may take a chance and respond to the targets with the risk of responding to nontargets as well, whereas another may prefer to miss uncertain targets than make erroneous responses to nontargets. This difference is captured by the response bias measure (C). This measures the response tendency of the subject for a given degree of uncertainty. Thus, the two hypothetical subjects in the example may have similar discriminability (d'), but show very different response bias (C).

Initial analysis of data from control subjects revealed significant effects of age on both of these measures (Pearson's $r = .43$ and $.55$, $p = .02$ and $.002$ for d'_L and C_L, respectively). In order to avoid confounding from age in the performance-based analyses, we therefore generated age-adjusted z-scores for the d'_L and C_L measures.

Behavior-driven Lesion Analyses

Lesion overlap images were constructed to identify brain areas affected in common in subjects who were impaired on the task. The resulting overlap image was compared to the overlap image of those who performed within the normal range. These qualitative lesion–symptom analyses were followed up with VLSM. In VLSM, the performance measure of interest is entered as a continuous variable, and statistical comparisons are made for each eligible voxel, comparing the performance of subjects with a lesion affecting a given voxel to that of subjects with lesions sparing that voxel. We used the nonparametric Brunner-Munzel (BM) test (Brunner & Munzel, 2000) to perform statistical comparisons on a voxel-wise basis, as implemented in the freely available NPM (version 29 October 2008) and MRIcron (version 15 October 2008) software (Rorden et al., 2007; www.mricro.com/npm/ for NPM and www.mricro.com/mricron for MRIcron). Note that this version of NPM appropriately corrects for small group size. For this analysis, only voxels affected in at least three cases were included. Figure 1 shows the lesion coverage in our sample, indicating those regions where VLSM could test for potential lesion–behavior relationships. BM tests were performed at each eligible voxel with the performance measure as the dependent variable. In order to control for multiple comparisons, permutation tests (Nichols & Holmes, 2002) were applied to generate thresholds for the statistical maps, which were then represented graphically using MRIcron. Briefly, FWER permutation tests are a nonparametric resampling approach to control family-wise error rate that is assumption-free and more powerful compared to other commonly used ap-

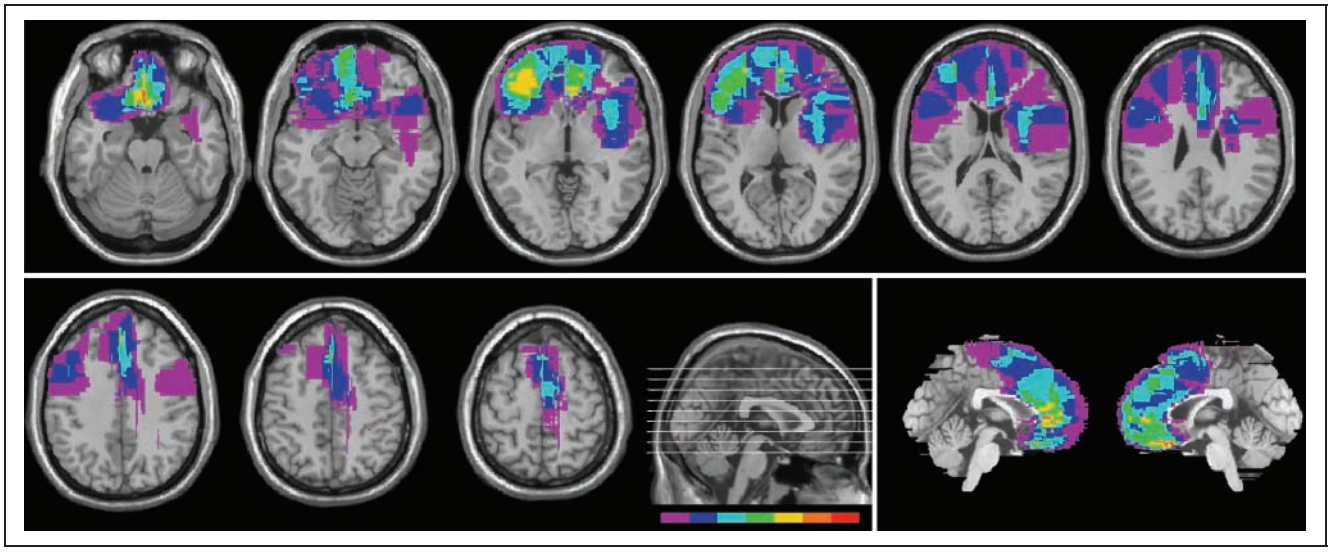

Figure 1. Representative axial slices, and left and right mid-sagittal views of the MNI brain, showing the degree of lesion coverage in our sample (n = 27); the levels of the axial slices are shown on a sagittal view. Colors indicate the degree of overlap across subjects. Only voxels damaged in at least three subjects (i.e., light blue and warmer colors) were entered into the VLSM analysis. Axial slices in this and subsequent figures are oriented according to radiological convention (right is left).

proaches, such as Bonferroni correction (Kimberg, Coslett, & Schwartz, 2007).

Region-of-Interest Groupwise Comparison

In addition to these behavior-based lesion analysis methods, conventional lesion-based subgroup comparisons were performed to verify the VLSM findings and to allow the results to be more readily compared to the existing neuropsychological literature. This anatomical subgroup analysis also allowed potential demographic confounds to be evaluated, at least qualitatively, as only age was adjusted for in the VLSM analysis. The patients were therefore divided into four groups based on the location of damage as determined by review of their most recent clinical brain imaging, by a neurologist experienced in interpretation of brain imaging and blind to task performance. These divisions conform to widely used regional PFC boundaries (e.g., Alexander, Stuss, Shallice, Picton, & Gillingham, 2005; Stuss et al., 2005) and resulted in subgroups which were of interest, given the VLSM results. Subjects were assigned to right or left lateral PFC groups (RL and LL, n = 4 and 5, respectively) when damage principally involved the middle and/or inferior frontal gyri in the right or left hemisphere, to the medial PFC group (MF, n = 11) when dorsal medial PFC, including dACC and the medial portion of the superior frontal gyrus, was the main site of damage, or to the orbito-frontal group (OF, n = 7) when damage primarily affected ventromedial and/or orbito-frontal cortex (OFC). When damage extended across these subdivisions, the subject was assigned to the group that was of particular interest, given the VLSM results, that is, patients with both dorsomedial and ventromedial/OFC damage were included in the MF group, and patients with both OFC and lateral PFC damage were included in the lateral groups. There were no patients with both MF and lateral PFC damage. Figure 2 shows the lesion locations, as overlap images, for these four anatomically defined groups. Welch's ANOVA (with post hoc Games–Howell tests when $p < .05$) was used to compare the performance of these subgroups with each other, and with the substantially larger control sample. Where parametric tests were inappropriate, the nonparametric Kruskall–Wallis test was applied.

RESULTS

In order to identify areas of damage common to poorly performing individuals, a lesion overlap image was first constructed for those who showed low d'_L, defined as age-adjusted z-scores of less than −1.5 (i.e., performing at more than 1.5 standard deviations below control levels) (Figure 3A). Out of 27 subjects with frontal lobe damage, seven patients fell below this cutoff. The overlap image revealed two patterns of damage associated with poor target discrimination: two out of the seven had damage involving left lateral PFC, variably involving DLPFC and VLPFC, and five patients had damage to the medial wall of PFC, with the highest overlap located in dACC. All of the subjects whose damage was restricted to right lateral PFC or OFC performed within the control range.

VLSM analysis allows statistical testing of these qualitative observations and considers the behavioral measures as continuous variables, avoiding the need for an arbitrary cutoff. VLSM with an FWER threshold of 0.05 identified dACC as significantly associated with poor n-back performance as captured by d'_L. Consistent with

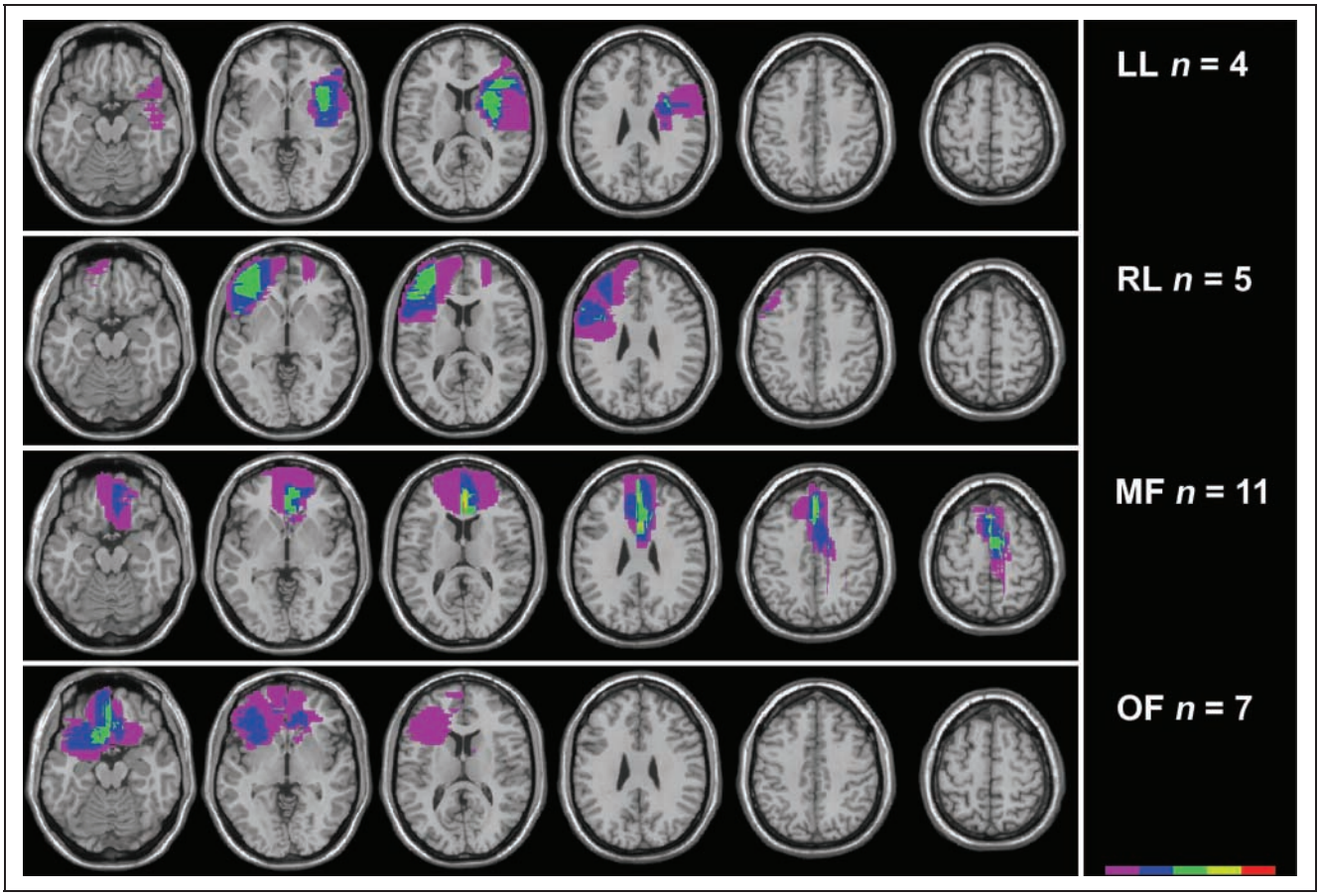

Figure 2. Representative axial slices showing the degree of lesion overlap for subjects with damage to left (LL group, $n = 4$, top row) or right (RL group, $n = 5$, second row) DLPFC and/or VLPFC, damage to the medial surface of PFC (MF group, $n = 11$, third row), or damage involving OFC (OF group, $n = 7$, bottom row). Colors indicate the degree of overlap across subjects, as shown in the legend.

the qualitative results from the overlap image, a region of left lateral PFC was also associated with poor performance, albeit to an extent that did not survive strict correction for multiple comparisons (Figure 3B).

These lesion–performance relationships were verified with a more conventional lesion-defined subgroup analysis. A between-group Welch's ANOVA revealed a significant main effect of group [mean age-adjusted d'_L *z*-score (*SD*) CTL 0.00 (0.96), LL −1.51 (0.25), RL 0.8 (1.64), MF −1.54 (1.44), OF 0.00 (0.79), $p < .01$] on age-adjusted d'_L. Pair-wise comparisons showed that both the LL group and the MF group had significantly lower d'_L compared to healthy control subjects (Games–Howell test, $p < .05$). We had several motivations for carrying out this planned analysis: First, to check the correspondence between voxel-based and conventional approaches, thereby allowing these findings to be confidently related to the existing neuropsychological literature: Second, to allow data regarding potential demographic confounders (other than age, which we did adjust for in the VLSM analysis) to be presented in an interpretable way. Finally, a subgroup analysis allowed a more fine-grained examination of the component processes that might be driving poor performance as captured by d'_L.

The demographic information for the anatomically defined subgroups is shown in Table 1. As expected, there were differences in the etiologies of damage across groups: Medial PFC damage followed ischemic or hemorrhagic stroke in five cases, resection of meningioma or low-grade glioma in five, and ruptured aneurysm in one. Damage to orbito-frontal PFC was due to ruptured aneurysm in five cases and meningioma resection in two cases. Damage to lateral PFC was caused by ischemic stroke in seven cases and tumor resection in two; all LL patients had suffered strokes.

Lesion volumes did not differ significantly between groups, and there was no effect of lesion volume on d'_L. Other demographic and neuropsychological factors did not seem to contribute to the key findings: There were no significant differences between the frontal subgroups in any of the demographic variables (Table 1), or the neuropsychological test scores (Table 2) [Kruskal–Wallis tests, all $p > .05$]. Furthermore, the significant differences in performance as captured by d'_L in the LL and MF groups remained when the demographic and neuropsychological variables that were significantly different between frontal patients and control subjects (premorbid IQ, BDI, category fluency, and backward

digit span) were separately entered as covariates in the analysis.

The performance difference between subjects with left and right-sided lateral PFC damage was striking: The d'_L z-scores of those with left lateral PFC damage ranged from −1.2 to −1.8, with a mean of −1.5, whereas subjects with similar damage to right lateral PFC had scores within the control range (−1.0 to 3.2, mean = 0.8). The right and left subgroups did not differ in lesion size [$t(7) = 0.25$, $p = .8$], nor was lesion size in these subjects or the frontal group as a whole a predictor of d'_L, arguing that location rather than lesion volume was the main determinant of performance. The performance difference also remained when the analysis was confined to subjects with ischemic stroke by excluding the two RL patients with tumor resection, arguing that etiology of damage was not a confound.

Although both left lateral and medial PFC damage were associated with a similarly impaired ability to discriminate the 2-back targets, further examination of

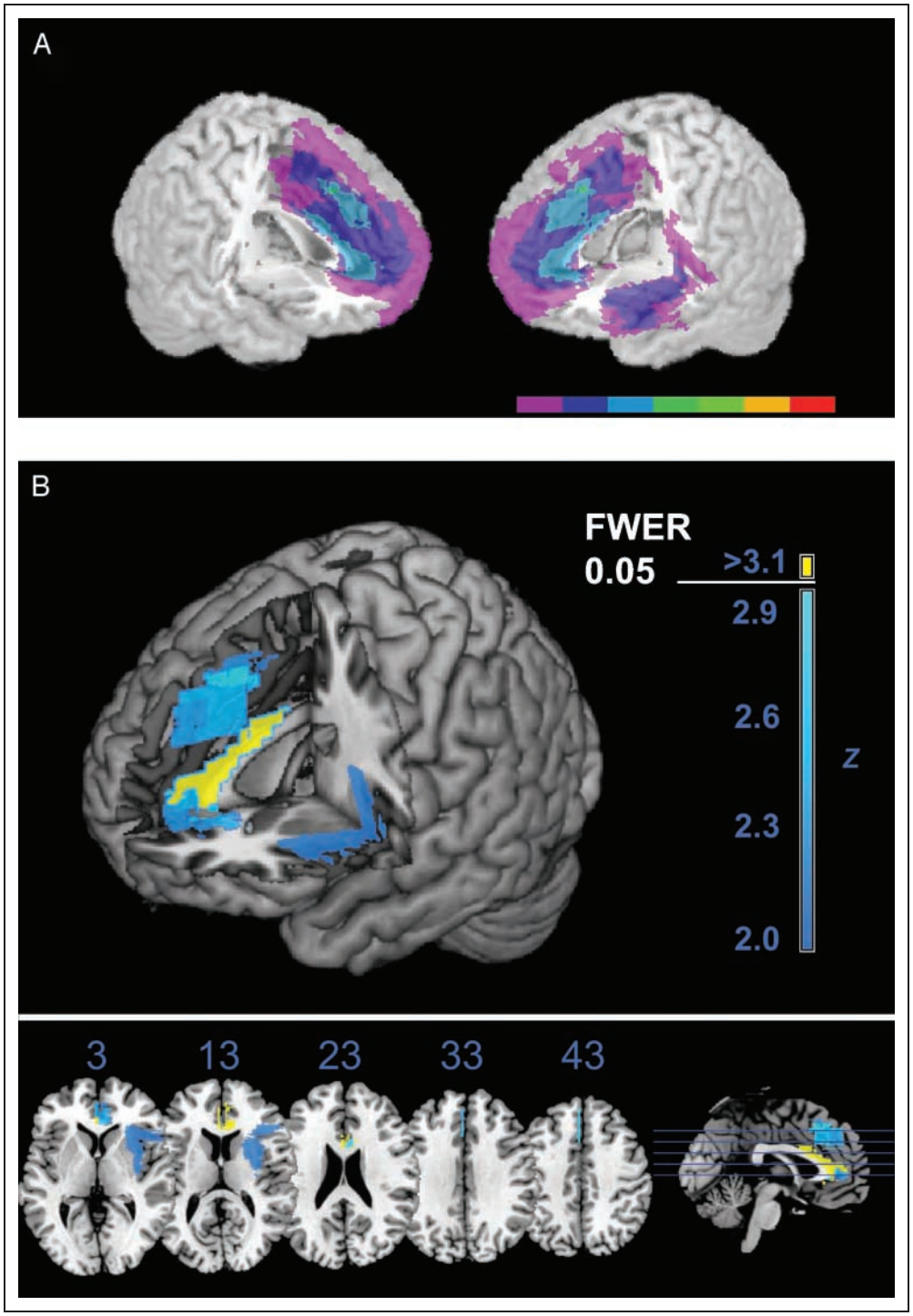

Figure 3. (A) Degree and location of lesion overlap for subjects with an age-adjusted d'_L z-score < −1.5. (B) VLSM statistical map computed for age-adjusted d'_L z-score overlaid on a three-dimensional view of the MNI brain in the top panel, with representative axial slices below. The color scale depicts BM test results comparing d'_L z-score on a voxel-by-voxel basis, with those voxels associated with poor performance as determined by permutation testing (1000 iterations) above a FWER threshold of 0.05 highlighted in yellow.

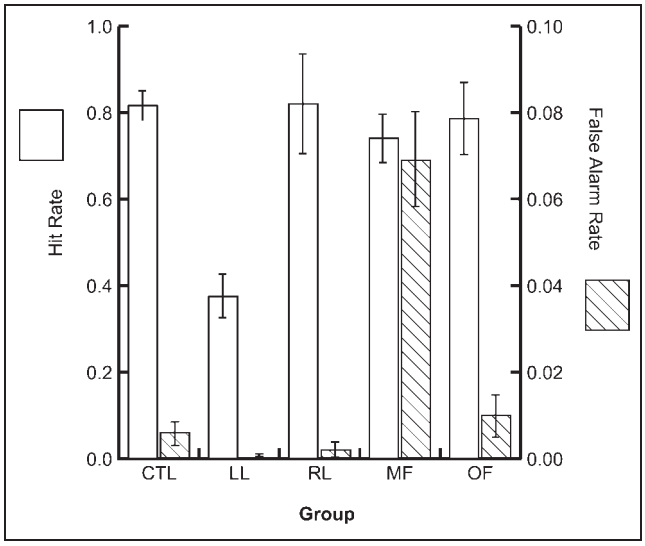

Figure 4. Hit and false alarm rates in each group. Error bars indicate *SEM*.

performance revealed that this impairment was qualitatively different in the two groups. Patients with MF damage tended to have a more liberal response bias (as captured by the signal detection measure C_L), compared to those with LL damage [$t(13) = 2.783, p = .016$]. In other words, for a given ability to discriminate targets from nontargets, subjects in the LL group were less likely to make a button press, whereas those in the MF group were more likely to respond to stimuli, whether target or nontarget. This tendency is even more starkly evident when the rate of false alarms is examined directly. The between-group comparison of false alarm rates revealed a significant effect of group (Kruskal–Wallis test, $p < .01$), with subjects in the MF group making more false alarms compared to the control group or to the other lesion groups (Mann–Whitney test, $p < .01$; Figure 4). On the other hand, although the hit rates were also significantly affected by group membership (Kruskal–Wallis test, $p = .03$), the MF group did not differ from controls on this measure (Mann–Whitney test, $p = .23$). In contrast, the LL group had significantly lower hit rates compared to control subjects (Mann–Whitney test, $p < .01$).

In the control group, examined at the individual level, the pattern is similar to the LL group: Individuals with poor discrimination typically showed a more conservative response bias, missing the targets but almost never making false alarms. The task was not designed to promote false alarms: Targets were infrequent (~16% of trials), presumably leading to a bias toward not responding, and there were no 1- or 3-back lures. These features likely minimized false alarms in the control group, and make the high rate of false alarms in the MF group particularly striking.

To insure that this observation was not an artifact of the particular boundaries we had applied for the subgroup analysis, the relation of false alarm rate to lesion location was tested using VLSM. We first compared the lesion overlap image for the patients with frontal damage with false alarm rates more than 1.5*SD* from the control group mean (>2.78%, or more than two false alarms; $n = 7$) to those whose false alarm rates were within the

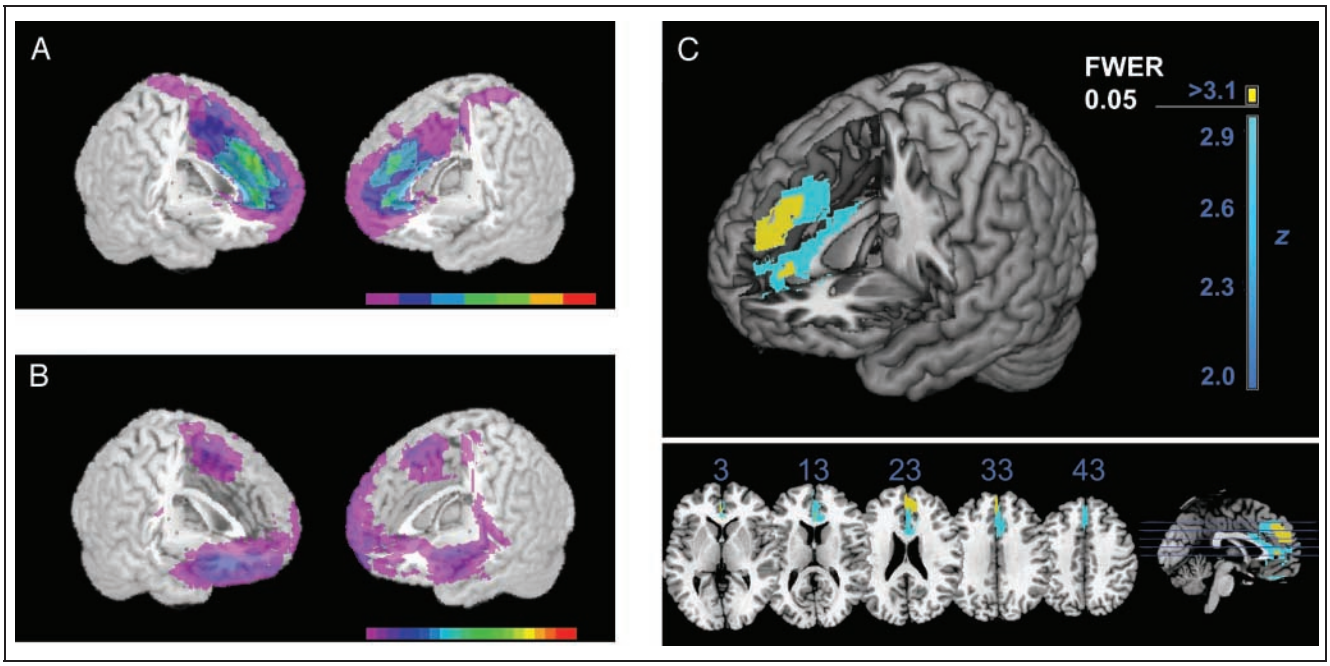

Figure 5. Location and degree of lesion overlap for the MF subjects (A) with false alarm rates exceeding that of the control group and (B) with false alarm rates falling within the control range. The areas damaged in those who showed excessive false alarms were also identified in the VLSM analysis (C). The color scale depicts BM test results comparing false alarm rates on a voxel-by-voxel basis, with those voxels associated with poor performance as determined by permutation testing (1000 iterations) above a FWER threshold of 0.05 highlighted in yellow.

control range (n = 20). As can be seen in Figure 5A and B, subjects who made many false alarms had lesions that overlapped in pregenual and dACC, whereas those with low false alarm rates had lateral or ventral PFC damage. Damage restricted to more dorsal and posterior medial PFC regions, affecting either the right or left supplementary motor area, was not associated with elevated false alarm rates (Figure 5B). The VLSM analysis, with false alarm rate as the dependent variable, revealed significant lesion–performance associations for bilateral dACC and adjacent fronto-median cortex, confirming the qualitative overlap analysis (Figure 5C). No other frontal region was associated with higher false alarm rates, even with more liberal thresholding.

Within the MF group, lesion size tended to be larger for subjects who made more false alarms (n = 6, mean lesion volume = 47 cc) than those with few such errors (n = 5, mean lesion volume = 9 cc). However, even very small lesions affecting the critical region indicated by the VLSM analysis (e.g., 8 cc, 9 cc) could be associated with high false alarm rates, arguing that site rather than volume of damage was the critical determinant. In support of this view, neither false alarm rate nor the response criterion measure C_L showed any correlation with lesion volume in the MF group.

A potential explanation for the high rate of false alarms after MF damage is that these patients were responding too quickly (i.e., making more ''fast guesses''). However, there was no significant difference in correct response reaction times (RTs) between groups [mean RT (*SD*): CTL, 741.1 (185.0) msec; LL, 765.5 (128.9) msec; RL, 783.8 (128.9) msec; MF, 863.9 (256.6) msec; OF, 815.7 (223.0) msec; Welch's ANOVA, p = .71], or between MF subjects with high [961.0 (286.9) msec] compared to those with low [747.4 (174.9) msec, $t(9)$ = 1.448, p = .18] false alarm rates. If anything, MF subjects making many false alarms tended to have slower RTs on correct trials. RTs were generally slower on incorrect trials than correct trials in all groups [$F(1, 2)$ = 6.010, p = .03], although there was individual variation in this tendency, with some individuals showing faster RTs for incorrect trials, regardless of group membership. Inspection of false alarms in MF subjects did not reveal any consistent pattern, except for one subject with the highest number of false alarms who tended to make two or three consecutive false alarm responses. The majority of the stimuli that induced false alarms had not been presented in the previous 10 trials, and there were even a few trials where subjects made false alarms to a letter that had never been presented, arguing against simple recency effects as an explanation.

DISCUSSION

The n-back task is a widely used task in functional imaging studies of working memory, and is being increasingly used as a behavioral probe of lateral PFC function. However, the existing evidence that task performance relies critically on lateral PFC is conflicting. Furthermore, whether other regions within PFC are necessary for performance of this task, and if so, in what way, is unclear.

The present study examined the effects of focal frontal lobe damage on a letter version of the 2-back task, aiming to determine which areas within PFC are necessary for particular components of task performance. We found that two specific prefrontal regions play a critical role in performance of the 2-back task. A comparison of performance across pre-defined anatomical subgroups revealed significant impairment in LL and MF groups, but not in RL and OF groups, compared to control subjects. The specific areas driving these group effects can be gleaned from the VLSM analysis, which showed clear effects of dACC damage on n-back performance, and a weaker effect of damage to an area within left lateral PFC.

These findings provide important converging support for functional imaging studies of the n-back task, identifying critical nodes within the network of frontal regions activated during this task. The finding that LL damage is associated with poor performance argues that the activations in dorsolateral (BA9/46) and ventrolateral (BA45,47) PFC [Talairach coordinates (x = −44, y = 18, z = 22) and (x = −30, y = 18, z = 6), respectively] that have been identified by a meta-analysis of fMRI studies using the n-back task are indeed task-relevant (Owen et al., 2005). Previous efforts to demonstrate a critical role for dACC in executive functions in patients with damage to the area have provided mixed results (di Pellegrino, Ciaramelli, & Ladavas, 2007; Baird et al., 2006; Fellows & Farah, 2005; Critchley et al., 2003; Swick & Jovanovic, 2002; Stuss, Floden, Alexander, Levine, & Katz, 2001), making the finding of a necessary role for MF, and specifically dACC, in letter 2-back task performance particularly notable.

A clue as to the specific processes subserved by these two PFC regions came from more detailed analysis of the performance deficits in these two groups of patients. Although damage to either medial or left lateral PFC led to similarly poor discriminability, response bias was affected differently by these two sites of damage: Subjects with left lateral PFC damage showed a normal, relatively conservative response bias, here meaning that when there was uncertainty as to whether a stimulus was a target or nontarget, the tendency was *not* to respond, whereas subjects with medial PFC damage were more likely to respond under the same conditions. These distinct behavioral patterns argue for distinct roles for these two regions: Left lateral PFC appears to be necessary for the working memory-based identification of correct 2-back targets, but not for adjusting response bias. In contrast, dACC appears to be critical for optimizing response bias in the face of uncertainty.

Our findings also indicate that intact right lateral PFC is not critical for task performance, despite the frequent

finding of bilateral lateral PFC activations in fMRI studies (Owen et al., 2005). This discrepancy is not without precedent: Bilateral activations in homologous areas are a common finding in functional imaging studies of other cognitive processes, such as language, where loss-of-function evidence argues that only one hemisphere is critically involved (Van Lancker-Sidtis, 2006). TMS studies in healthy adults have shown effects of stimulation over right or left lateral PFC on letter 2-back accuracy (Mottaghy, Pascual-Leone, et al., 2003; Mottaghy et al., 2000). However, only stimulation over the left hemisphere affected both accuracy and RT, and the effective timing of stimulation was earlier in the right than left lateral PFC, leading the authors to suggest that information was converging in left lateral PFC (Mottaghy, Gangitano, et al., 2003). Our findings support a privileged role for left lateral PFC in *n*-back performance, at least in this letter 2-back version of the task.

The apparent contradiction between the two prior neuropsychological group studies focusing on lateral PFC may be explained by the laterality effect observed in our study: In the study by Muller et al. (2002), three out of four patients with impaired *n*-back performance had left lateral PFC damage (DLPFC and VLPFC lesion groups), whereas four of six patients with intact performance had right lateral PFC damage (dorsal PFC lesion group). Similarly, in the study by Boisgueheneuc et al. (2006), all the subjects in the impaired group had damage to the left hemisphere, albeit involving the superior frontal gyrus, whereas the control group of subjects with lateral prefrontal damage sparing the superior frontal gyrus consisted of four subjects with right hemisphere damage and only one subject with left hemisphere damage.

In aggregate, our study and prior work suggest a privileged role of left lateral PFC in *n*-back performance. More work will be needed, however, to clarify the specific areas within that subregion that are critical. Although our findings agree with those of Muller et al., our sample had limited coverage of left lateral PFC, and in particular, provided little information about the more superior lateral region examined in the Boisgueheneuc et al. study.

In accordance with much of the research on working memory in humans, the 2-back task was more specifically related to prefrontal damage than simpler span measures (D'Esposito & Postle, 1999). Performance on the Corsi block test, a spatial span task, did not distinguish frontal patients from controls. Frontal patients as a group did perform worse than control subjects on the backwards digit span test, but performance of this task did not differ across frontal subgroups, nor did it predict 2-back performance. The latter finding agrees with a report that found little relationship between working memory span and *n*-back performance in healthy subjects (Kane, Conway, Miura, & Colflesh, 2007).

The *n*-back task confounds several control aspects of working memory, including selection of relevant information in the face of distractors, updating representations in working memory and maintaining temporal order in working memory. Left lateral PFC has been implicated in each of these component processes, both in imaging (Petrides, 2005; Fletcher & Henson, 2001; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997) and human lesion work (Thompson-Schill et al., 1998, 2002; Petrides & Milner, 1982). Disruption of one or more of these abilities presumably explains the deficits we observed in 2-back performance after left lateral PFC damage.

In our study, participants with left lateral PFC damage showed poor target discriminability, but had a response bias similar to control subjects: that is, they missed targets, but did not make excessive false alarms. This pattern is perhaps surprising because frontal lobe injury is associated rather generally with a tendency to false alarms (Verfaellie, Rapcsak, Keane, & Alexander, 2004; Melo, Winocur, & Moscovitch, 1999; Swick & Knight, 1999). Furthermore, several studies of episodic memory probed by recognition tests have found that lateral PFC damage can increase false alarm rates. However, these false alarms are typically made to ''lure'' stimuli, such as recently presented or semantically related stimuli (Hamilton & Martin, 2005; Alexander, Stuss, & Fansabedian, 2003; Thompson-Schill et al., 2002). Although differences in specific patient characteristics or task parameters may be important in explaining the false alarms in those studies, the absence of false alarms in the present work may relate to the explicit absence of ''lures'' in the 2-back task used here. If so, it would suggest that lateral PFC damage leads to false alarms only when there are particular target (or task) features that ''draw'' a response.

In this context, the increase in false alarm rates observed in subjects with MF damage is striking. The VLSM analysis indicated that this tendency to false alarms was related to damage to dACC. The function of dACC has been the subject of much debate, spurred by the frequent finding of activation of this area in functional imaging studies using a variety of tasks. One prominent theory casts dACC as a monitor of response conflict in the service of cognitive control (Botvinick, Cohen, & Carter, 2004; Kerns et al., 2004). However, loss-of-function evidence has not definitively supported this view, with several studies of patients with damage to dACC reporting intact performance in tasks requiring cognitive control, and even the intact ability to flexibly adjust cognitive control (Baird et al., 2006; Fellows & Farah, 2005; Critchley et al., 2003; but see di Pellegrino et al., 2007; Swick & Jovanovic, 2002). Other theories of dACC function have suggested a role in integration of action–outcome history for the long-term adjustment of goal-directed behavior (Rushworth, Walton, Kennerley, & Bannerman, 2004), ''energization of attention'' (Alexander et al., 2005; Stuss et al., 2005), effort (Paus, 2001), or error prediction (Brown & Braver, 2005), hypotheses that are not necessarily mutually exclusive.

The results of the present study in isolation do not allow us to adjudicate definitively between these

possibilities. Generalized slowing has been observed in patients with similar damage performing simple RT and choice RT tasks (although such slowing was not a prominent feature in our sample), and interpreted as a deficit in ''energizing'' attention to initiate a response (Alexander et al., 2005; Stuss et al., 2005). Although waning of attention could explain why those with dACC damage occasionally missed 2-back targets, this account does not readily explain why they made excessive responses to nontargets. These patients were also not particularly fast in responding when they made false alarms, which might be expected if the false alarms resulted from the impulsivity sometimes observed in patients with ventromedial damage (Berlin, Rolls, & Kischka, 2004). Three of the subjects in the MF group who had high false alarm rates here are the same individuals who were able to adjust their performance in response to changing demands for cognitive control in Stroop and simple go/no-go tasks in a prior study (Fellows & Farah, 2005), making it unlikely that the elevated false alarms are related to a generic inability to detect conflict or inhibit prepotent responses.

Although further work will be needed to clarify the critical task demands that lead to deficits in patients with dACC damage, the combination of a ''go/no-go'' decision requirement and of ''on-line'' uncertainty about whether a stimulus is a target, are potentially relevant task requirements here. Consistent with error prediction and action–outcome integration accounts of dACC function, our findings suggest that this region is involved in adjusting response bias based on a signal about the likelihood of error determined just prior to the response. dACC damage may impair the ability to stop a potential response judged on-line to be either wrong or too uncertain, or the ability to make this judgment prior to responding. In support of this view, when a subset of the patients with dACC damage was subsequently tested on a version of the *n*-back task that asked for confidence ratings at the time of response, the patients reported high confidence prior to false alarms, whereas controls were more likely to rate themselves as ''uncertain'' when making false alarms (Modirrousta & Fellows, 2008b). This ''predicted performance''–''actual performance'' mismatch was also evident prior to attempted recognition in a separate set of experiments testing episodic memory in the same patients (Modirrousta & Fellows, 2008a).

In summary, our findings suggest that there are at least two component processes required for normal performance of the letter 2-back task, and that these rely critically on two distinct areas within PFC. This study supports the ''fractionated'' view of PFC function (Stuss & Alexander, 2007), and extends our understanding of the distinct functions for which two of these PFC regions are necessary. This work also has implications for interpreting behavioral studies that use the *n*-back paradigm as a probe for regionally specific prefrontal–executive functions. Overall task performance can reflect either left lateral or medial PFC dysfunction, or (presumably) both. Specific behavioral measures, such as false alarm rate, may allow more specific inferences to be drawn about the neuroanatomical substrates of impaired performance.

Acknowledgments

We thank the patient database managers at both the University of Pennsylvania (Marianna Stark) and McGill University (Arlene Berg) for their help with patient recruitment, and the patients and their families for their generous participation. Alisa Padon provided expert technical assistance. The work was supported by the National Institutes of Health (National Institute of Neurological Disorders and Stroke) NS045074 and Canadian Institutes of Health Research (MOP 77583). L. K. F. is a Killam Scholar and CIHR Clinician–Scientist. A. T. was supported by the David Stewart Memorial Fellowship and Frederick Banting and Charles Best Canada Graduate Scholarships-Doctoral Award.

Reprint requests should be sent to Ami Tsuchida, Montreal Neurological Institute, Rm 276, 3801 University St, Montréal, QC, Canada H3A 2B4, or via e-mail: ami.tsuchida@mail.mcgill.ca.

REFERENCES

Alexander, M. P., Stuss, D. T., & Fansabedian, N. (2003). California Verbal Learning Test: Performance by patients with focal frontal and non-frontal lesions. *Brain, 126,* 1493–1503.

Alexander, M. P., Stuss, D. T., Shallice, T., Picton, T. W., & Gillingham, S. (2005). Impaired concentration due to frontal lobe damage from two distinct lesion sites. *Neurology, 65,* 572–579.

Baird, A., Dewar, B.-K., Critchley, H., Gilbert, S. J., Dolan, R. J., & Cipolotti, L. (2006). Cognitive functioning after medial frontal lobe damage including the anterior cingulate cortex: A preliminary investigation. *Brain and Cognition, 60,* 166–175.

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

Bates, E., Wilson, S. M., Saygin, A. P., Dick, F., Sereno, M. I., Knight, R. T., et al. (2003). Voxel-based lesion–symptom mapping. *Nature Neuroscience, 6,* 448–450.

Bechara, A., Damasio, H., Tranel, D., & Anderson, S. W. (1998). Dissociation of working memory from decision making within the human prefrontal cortex. *Journal of Neuroscience, 18,* 428–437.

Berlin, H. A., Rolls, E. T., & Kischka, U. (2004). Impulsivity, time perception, emotion and reinforcement sensitivity in patients with orbitofrontal cortex lesions. *Brain, 127,* 1108–1126.

Boisgueheneuc, F. d., Levy, R., Volle, E., Seassau, M., Duffau, H., Kinkingnehun, S., et al. (2006). Functions of the left superior frontal gyrus in humans: A lesion study. *Brain, 129,* 3315–3328.

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences, 8,* 539–546.

Braver, T. S., Cohen, J. D., Nystrom, L. E., Jonides, J., Smith, E. E., & Noll, D. C. (1997). A parametric study of prefrontal cortex involvement in human working memory. *Neuroimage, 5,* 49–62.

Brown, J. W., & Braver, T. S. (2005). Learned predictions of error likelihood in the anterior cingulate cortex. *Science, 307,* 1118–1121.

Brunner, E., & Munzel, U. (2000). The nonparametric Behrens-Fisher problem: Asymptotic theory and a small-sample approximation. *Biometrical Journal, 42,* 17–25.
Butters, N., & Pandya, D. (1969). Retention of delayed-alternation: Effect of selective lesions of sulcus principalis. *Science, 165,* 1271–1273.
Callicott, J. H., Bertolino, A., Mattay, V. S., Langheim, F. J. P., Duyn, J., Coppola, R., et al. (2000). Physiological dysfunction of the dorsolateral prefrontal cortex in schizophrenia revisited. *Cerebral Cortex, 10,* 1078–1092.
Callicott, J. H., Mattay, V. S., Bertolino, A., Finn, K., Coppola, R., Frank, J. A., et al. (1999). Physiological characteristics of capacity constraints in working memory as revealed by functional MRI. *Cerebral Cortex, 9,* 20–26.
Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science, 280,* 747–749.
Cohen, J. D., Forman, S. D., Braver, T. S., Casey, B. J., Servan-Schreiber, D., & Noll, D. C. (1993). Activation of the prefrontal cortex in a nonspatial working memory task with functional MRI. *Human Brain Mapping, 1,* 293–304.
Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., et al. (1997). Temporal dynamics of brain activation during a working memory task. *Nature, 386,* 604–608.
Corwin, J. (1994). On measuring discrimination and response bias: Unequal numbers of targets and distractors and two classes of distractors. *Neuropsychology, 8,* 110–117.
Courtney, S. M., Petit, L., Haxby, J. V., & Ungerleider, L. G. (1998). The role of prefrontal cortex in working memory: Examining the contents of consciousness. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 353,* 1819–1828.
Critchley, H. D., Mathias, C. J., Josephs, O., O'Doherty, J., Zanini, S., Dewar, B.-K., et al. (2003). Human cingulate cortex and autonomic control: Converging neuroimaging and clinical evidence. *Brain, 126,* 2139–2152.
D'Esposito, M., & Postle, B. R. (1999). The dependence of span and delayed-response performance on prefrontal cortex. *Neuropsychologia, 37,* 1303–1315.
di Pellegrino, G., Ciaramelli, E., & Ladavas, E. (2007). The regulation of cognitive control following rostral anterior cingulate cortex lesion in humans. *Journal of Cognitive Neuroscience, 19,* 275–286.
Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences, 23,* 475–483.
Fellows, L. K., & Farah, M. J. (2005). Is anterior cingulate cortex necessary for cognitive control? *Brain, 128,* 788–796.
Fellows, L. K., Heberlein, A. S., Morales, D. A., Shivde, G., Waller, S., & Wu, D. H. (2005). Method matters: An empirical study of impact in cognitive neuroscience. *Journal of Cognitive Neuroscience, 17,* 850–858.
Fletcher, P. C., & Henson, R. N. (2001). Frontal lobes and human memory: Insights from functional neuroimaging. *Brain, 124,* 849–881.
Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). ''Mini-mental state''. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198.
Funahashi, S., Bruce, C. J., & Goldman-Rakic, P. S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology, 61,* 331–349.
Fuster, J. M. (1985). The prefrontal cortex, mediator of cross-temporal contingencies. *Human Neurobiology, 4,* 169–179.
Goldman-Rakic, P. S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In F. Plum & V. B. Mountcastle (Eds.), *Handbook of physiology, the nervous system, higher functions of the brain* (Vol. 5, pp. 373–417). Bethesda, MD: American Physiological Society.
Hamilton, A. C., & Martin, R. C. (2005). Dissociations among tasks involving inhibition: A single-case study. *Cognitive, Affective, & Behavioral Neuroscience, 5,* 1–13.
Hautzel, H., Mottaghy, F. M., Schmidt, D., Zemb, M., Shah, N. J., Muller-Gartner, H. W., et al. (2002). Topographic segregation and convergence of verbal, object, shape and spatial working memory in humans. *Neuroscience Letters, 323,* 156–160.
Jonides, J., Schumacher, E. H., Smith, E. E., Lauber, E. J., Awh, E., Minoshima, S., et al. (1997). Verbal working memory load affects regional brain activation as measured by PET. *Journal of Cognitive Neuroscience, 9,* 462–475.
Kane, M. J., Conway, A. R., Miura, T. K., & Colflesh, G. J. (2007). Working memory, attention control, and the N-back task: A question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 615–622.
Kerns, J. G., Cohen, J. D., MacDonald, A. W., III, Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science, 303,* 1023–1026.
Kimberg, D. Y., Coslett, H. B., & Schwartz, M. F. (2007). Power in voxel-based lesion–symptom mapping. *Journal of Cognitive Neuroscience, 19,* 1067–1080.
Mattay, V. S., Fera, F., Tessitore, A., Hariri, A. R., Berman, K. F., Das, S., et al. (2006). Neurophysiological correlates of age-related changes in working memory capacity. *Neuroscience Letters, 392,* 32–37.
McMillan, K. M., Laird, A. R., Witt, S. T., & Meyerand, M. E. (2007). Self-paced working memory: Validation of verbal variations of the n-back paradigm. *Brain Research, 1139,* 133–142.
Melo, B., Winocur, G., & Moscovitch, M. (1999). False recall and false recognition: An examination of the effects of selective and combined lesions to the medial temporal lobe/diencephalon and frontal lobe structures. *Cognitive Neuropsychology, 16,* 343–359.
Modirrousta, M., & Fellows, L. K. (2008a). Medial prefrontal cortex plays a critical and selective role in ''feeling of knowing'' meta-memory judgments. *Neuropsychologia, 46,* 2958–2965.
Modirrousta, M., & Fellows, L. K. (2008b). Dorsal medial prefrontal cortex plays a necessary role in rapid error prediction in humans. *Journal of Neuroscience.* DOI: 10.1523/JNEUROSCI.4450-08.2008.
Mottaghy, F. M., Gangitano, M., Krause, B. J., & Pascual-Leone, A. (2003). Chronometry of parietal and prefrontal activations in verbal working memory revealed by transcranial magnetic stimulation. *Neuroimage, 18,* 565–575.
Mottaghy, F. M., Krause, B. J., Kemna, L. J., Topper, R., Tellmann, L., Beu, M., et al. (2000). Modulation of the neuronal circuitry subserving working memory in healthy human subjects by repetitive transcranial magnetic stimulation. *Neuroscience Letters, 280,* 167–170.
Mottaghy, F. M., Pascual-Leone, A., Kemna, L. J., Topper, R., Herzog, H., Muller-Gartner, H. W., et al. (2003). Modulation of a brain–behavior relationship in verbal working memory by rTMS. *Brain Research, Cognitive Brain Research, 15,* 241–249.

Mull, B. R., & Seyal, M. (2001). Transcranial magnetic stimulation of left prefrontal cortex impairs working memory. *Clinical Neurophysiology, 112,* 1672–1675.
Muller, N. G., Machado, L., & Knight, R. T. (2002). Contributions of subregions of the prefrontal cortex to working memory: Evidence from brain Lesions in humans. *Journal of Cognitive Neuroscience, 14,* 673–686.
Nasreddine, Z. S., Phillips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., et al. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 53,* 695–699.
Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping, 15,* 1–25.
Nystrom, L. E., Braver, T. S., Sabb, F. W., Delgado, M. R., Noll, D. C., & Cohen, J. D. (2000). Working memory for letters, shapes, and locations: fMRI evidence against stimulus-based regional organization in human prefrontal cortex. *Neuroimage, 11,* 424–446.
Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping, 25,* 46–59.
Paus, T. (2001). Primate anterior cingulate cortex: Where motor control, drive and cognition interface. *Nature Reviews Neuroscience, 2,* 417–424.
Paus, T., Koski, L., Caramanos, Z., & Westbury, C. (1998). Regional differences in the effects of task difficulty and motor output on blood flow response in the human anterior cingulate cortex: A review of 107 PET activation studies. *NeuroReport, 9,* R37–R47.
Perlstein, W. M., Carter, C. S., Noll, D. C., & Cohen, J. D. (2001). Relation of prefrontal cortex dysfunction to working memory and symptoms in schizophrenia. *American Journal of Psychiatry, 158,* 1105–1113.
Petrides, M. (2005). Lateral prefrontal cortex: Architectonic and functional organization. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 360,* 781–795.
Petrides, M., & Milner, B. (1982). Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia, 20,* 249–262.
Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences, 10,* 59–63.
Ptak, R., & Schnider, A. (2004). Disorganised memory after right dorsolateral prefrontal damage. *Neurocase: Case Studies in Neuropsychology, Neuropsychiatry, and Behavioural Neurology, 10,* 52–59.
Ptito, A., Crane, J., Leonard, G., Amsel, R., & Caramanos, Z. (1995). Visual–spatial localization by patients with frontal-lobe lesions invading or sparing area 46. *NeuroReport, 6,* 1781–1784.
Rorden, C., & Brett, M. (2000). Stereotaxic display of brain lesions. *Behavioural Neurology, 12,* 191–200.
Rorden, C., & Karnath, H.-O. (2004). Using human brain lesions to infer function: A relic from a past era in the fMRI age? *Nature Reviews Neuroscience, 5,* 812–819.
Rorden, C., Karnath, H.-O., & Bonilha, L. (2007). Improving lesion–symptom mapping. *Journal of Cognitive Neuroscience, 19,* 1081–1088.
Rushworth, M. F., Walton, M. E., Kennerley, S. W., & Bannerman, D. M. (2004). Action sets and decisions in the medial frontal cortex. *Trends in Cognitive Sciences, 8,* 410–417.
Saygin, A. P., Wilson, S. M., Dronkers, N. F., & Bates, E. (2004). Action comprehension in aphasia: Linguistic and non-linguistic deficits and their lesion correlates. *Neuropsychologia, 42,* 1788–1804.
Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers: A Journal of the Psychonomic Society, 31,* 137–149.
Stuss, D. T., & Alexander, M. P. (2007). Is there a dysexecutive syndrome? *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 362,* 901–915.
Stuss, D. T., Alexander, M. P., Shallice, T., Picton, T. W., Binns, M. A., Macdonald, R., et al. (2005). Multiple frontal systems controlling response speed. *Neuropsychologia, 43,* 396–417.
Stuss, D. T., Floden, D., Alexander, M. P., Levine, B., & Katz, D. (2001). Stroop performance in focal lesion patients: Dissociation of processes and frontal lobe lesion location. *Neuropsychologia, 39,* 771–786.
Swick, D., & Jovanovic, J. (2002). Anterior cingulate cortex and the Stroop task: Neuropsychological evidence for topographic specificity. *Neuropsychologia, 40,* 1240–1253.
Swick, D., & Knight, R. T. (1999). Contributions of prefrontal cortex to recognition memory: Electrophysiological and behavioral evidence. *Neuropsychology, 13,* 155–170.
Thermenos, H. W., Goldstein, J. M., Buka, S. L., Poldrack, R. A., Koch, J. K., Tsuang, M. T., et al. (2005). The effect of working memory performance on functional MRI in schizophrenia. *Schizophrenia Research, 74,* 179–194.
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. *Proceedings of the National Academy of Sciences, U.S.A., 94,* 14792–14797.
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. *Cognitive, Affective, & Behavioral Neuroscience, 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. *Proceedings of the National Academy of Sciences, U.S.A., 95,* 15855–15860.
Tyler, L. K., Marslen-Wilson, W., & Stamatakis, E. A. (2005). Dissociating neuron-cognitive component processes: Voxel-based correlational methodology. *Neuropsychologia, 43,* 771–778.
Van Lancker-Sidtis, D. (2006). Does functional neuroimaging solve the questions of neurolinguistics? *Brain and Language, 98,* 276–290.
Verfaellie, M., Rapcsak, S. Z., Keane, M. M., & Alexander, M. P. (2004). Elevated false recognition in patients with frontal lobe damage is neither a general nor a unitary phenomenon. *Neuropsychology, 18,* 94–103.
Walsh, V., & Cowey, A. (2000). Transcranial magnetic stimulation and cognitive neuroscience. *Nature Reviews Neuroscience, 1,* 73–80.