

Functional Neuroimaging of Working Memory in Schizophrenia: Task Performance as a Moderating Variable

Jared X. Van Snellenberg
Simon Fraser University

Ivan J. Torres
Simon Fraser University and Riverview Hospital

Allen E. Thornton
Simon Fraser University

Functional neuroimaging studies in schizophrenia have demonstrated abnormal activation of dorsolateral prefrontal cortex (DLPFC) during working memory (WM) performance. However, findings of increased and decreased activity have been reported. The authors used meta-analysis to investigate whether diverging results arise as a function of differential WM task performance between patients and control participants. Results indicate that the magnitude of the group difference in WM performance is a moderator of DLPFC activation differences, and concepts such as hypo- or hyperfrontality do not universally characterize WM findings in schizophrenia. Thus, the variability in the WM activation findings between participants with schizophrenia and control participants reflects the specific conditions under which WM functions are evaluated, not just the WM construct per se.

Keywords: schizophrenia, working memory, neuroimaging, prefrontal cortex, meta-analysis

A substantial body of research implicates pathology, both structural and functional, of the prefrontal cortex in schizophrenia (e.g., Akbarian et al., 1993; Callicott et al., 2003; Selemon, Rajkowska, & Goldman-Rakic, 1995). In addition, functional neuroimaging studies of healthy individuals performing working memory (WM) tasks have demonstrated that dorsolateral prefrontal cortex (DLPFC) is activated during their performance (for review, see D'Esposito, 2001; Smith & Jonides, 1999). A common approach to elucidating the nature of prefrontal deficit in schizophrenia has been to compare prefrontal blood flow between affected individuals and healthy controls by using functional neuroimaging during WM tasks. The conclusions to be drawn from this approach, however, remain unclear, as patients have demonstrated both increased and decreased prefrontal activity relative to controls (for review, see Callicott, 2003; Manoach, 2003).

A nascent explanation for these inconsistent findings is that the direction and magnitude of differences in activation of DLPFC between patient and control participants may depend on the difficulty of a given task, which can be manipulated by changing WM load demands. Specifically, although in healthy participants DLPFC acti-

vation shows monotone increases with increasing load (Jonides et al., 1997; Manoach et al., 1997), it has been proposed that as WM load increases beyond a participant's ability to effectively perform the task, the relationship between DLPFC activation and WM load becomes nonmonotone (Callicott et al., 2003; Manoach, 2002, 2003). Indeed, an inverted-U relationship between DLPFC activation and WM load has been demonstrated in normal volunteers through use of the *n*-back task (Callicott et al., 1999) and in a recent reanalysis (Rypma, 2006) of data from an item-recognition task (Rypma, Berger, & D'Esposito, 2002). Thus, prefrontal deficit in schizophrenia may not manifest as a simple increase or decrease in the degree of DLPFC activation during WM task performance but rather as a shift in a nonmonotone or inverted-U relationship between WM load and DLPFC activation, resulting in greater cortical activation by patients at relatively low loads and lesser cortical activation at relatively high loads (see Figure 1).

We undertook the present meta-analysis to determine whether the literature supports this view. However, because of the considerable heterogeneity in WM task types used in functional imaging studies, a straightforward quantification or categorization of task difficulty based on load was not possible. As a result, we elected to investigate patient-control differences in task performance. Imaging studies in healthy individuals have consistently demonstrated decreases in performance with increasing WM load (e.g., Braver et al., 1997; Cairo, Liddle, Woodward, & Ngan, 2004; Callicott et al., 1999; Jonides et al., 1997; Manoach et al., 1997; Rypma et al., 2002; Rypma, Prabhakaran, Desmond, Glover, & Gabrieli, 1999), presumably as a result of increased maintenance demands, strategy changes, and/or the deployment of executive strategies (e.g., chunking) at different loads (see Rypma, 2006). Thus, although differences in performance are not likely to be related to WM load in a straightforward way, we assumed that patients exhibit differentially poorer performance relative to controls at greater loads, in which case patient-control differences in WM task performance could be viewed as an index of task diffi-

Jared X. Van Snellenberg and Allen E. Thornton, Department of Psychology, Simon Fraser University, Burnaby, British Columbia, Canada; Ivan J. Torres, Department of Psychology, Simon Fraser University, and Department of Research, Riverview Hospital, Coquitlam, British Columbia, Canada.

Additional materials are on the Web at <http://dx.doi.org/10.1037/0894-4105.20.5.497.supp>

We would like to acknowledge Ray Koopman for his helpful comments on the statistical procedures used in our study and Damon Morris for his assistance in obtaining effect size estimates from published figures.

Correspondence concerning this article should be addressed to Ivan J. Torres, Department of Psychology, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia V5A 1S6, Canada. E-mail: itorres@sfu.ca

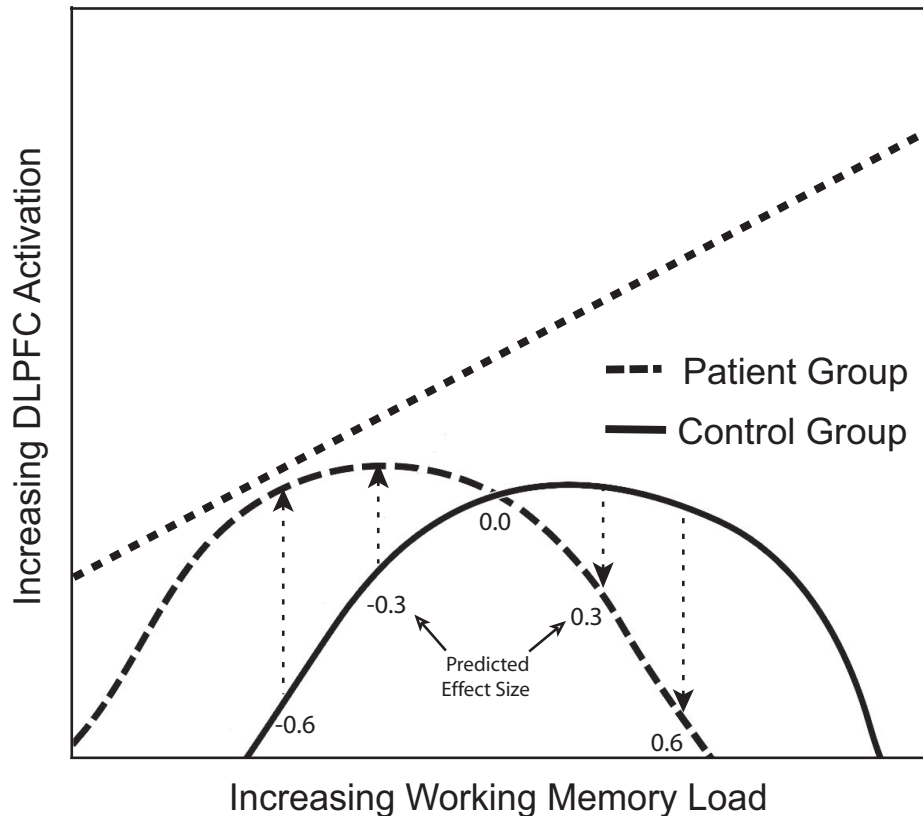


Figure 1. Representation of the hypothesized inverted-U relationship between working memory load and dorsolateral prefrontal cortex (DLPFC) activation for patients and controls. Numbers indicate the predicted control minus patient effect size as a consequence of the left shift in the inverted U for patients. The dashed straight line represents the expected linear relationship between this group difference in activation and working memory load.

culty and an indirect measure of WM load. In this case, the inverted-U hypothesis outlined above would predict an approximately linear relationship between group differences in DLPFC activation and task performance.

Method

Study Ascertainment

An attempt was made to identify all studies that (a) were published between January 1, 1980, and June 30, 2004; (b) utilized functional MRI or positron emission tomography; (c) included both schizophrenic patient and healthy control groups; (d) conducted scanning under conditions other than a resting state, auditory verbal hallucination, or pharmacological challenge; and (e) tested the between-group difference in activation of DLPFC or reported data that permitted an estimation of this difference (see below). We identified candidate studies through a search of the online Medline and PsycINFO databases by using each of the pairs of keywords listed in Appendix A. In addition, a single master's thesis meeting the above criteria was ascertained by other means (Cairo, 2003).

Identified studies were retained for analysis if the task used during imaging or positron emission topography tracer uptake was identical or highly similar to a task considered by the authors of at least one study to require working memory. Of these, studies were included if they used tasks for which it was reasonable to assume that functional images obtained

during their performance would predominantly reflect the active maintenance of information in WM and operations on this information. This resulted in the exclusion of (a) tasks with heavy demands on nonmnemonic executive processing, such as the Wisconsin Card Sorting Task; (b) tasks that predominantly rely on long-term memory; (c) tasks predominantly requiring attentional and stimulus-discrimination processes, such as the standard Continuous Performance Test (CPT); and (d) a click-counting task (Ojeda et al., 2002), as the working memory load was likely minimal and task performance required accurate estimation of the rate of presentation of clicks. Studies using CPT tasks were included if they required WM processing (i.e., if the status of a presented stimulus as a target was contingent on previously occurring stimuli), as is the case for the AX-CPT (e.g., Barch et al., 2001) and CPT tasks in which stimuli were targets whenever they were identical to the stimulus that occurred on the preceding trial (e.g., Salgado-Pineda et al., 2004). Some studies were excluded because of considerable sample overlap with another study included in the analysis; in this case, preference was always given to studies in which all necessary effect sizes were estimable (see below) or to the study with the larger sample. This resulted in 29 studies in the analysis, using the following task types: (a) *n*-back tasks, (b) item recognition tasks (e.g., Sternberg tasks), (c) complex CPT tasks, (d) a heterogeneous group of delayed match-to-sample tasks, and (e) a single mental arithmetic task (Hugdahl et al., 2004). One doctoral dissertation (Mendrek, 2000) reported data for two independent samples of patient and control participants and so was included as two separate studies, resulting in a total of 30 studies in the

analysis (the complete citation list for all studies in the analysis is available on the Web at <http://dx.doi.org/10.1037/0894-4105.20.5.497.suppl>). Four of these studies indicated that a portion of their sample included patients with a diagnosis of schizoaffective disorder.

Meta-Analytic Procedures

Average effect sizes were estimated with a random-effects procedure (Hedges & Olkin, 1985), as implemented in Comprehensive Meta-Analysis (Borenstein & Rothstein, 1999). A random-effects model was chosen because homogeneity tests (Q tests) indicated that there was substantial heterogeneity in the data (all $ps < .0005$); in such cases, fixed-effects models overestimate the confidence in the average effect estimate (Hedges & Vevea, 1998), resulting in inappropriately small confidence intervals and increased Type I error rates.

The analysis of patient–control differences in task performance and of other continuous variables as moderators of patient–control differences in activation was carried out with Hedges and Olkin's (1985) regression-analog model for continuous moderator variables, implemented in SPSS and MATLAB (The Mathworks). Categorical moderators were tested with the Hedges and Olkin (1985) analysis of variance analog model in Comprehensive Meta-Analysis.

The Hedges and Olkin (1985) regression-analog model permits multivariate analyses, and because of the possibility that identified moderators may not remain significant after controlling for other variables, we carried out multivariate analyses as well. However, because of missing data due to nonreporting of many moderators of interest in most studies, the inclusion of multiple variables in moderator analyses can substantially reduce power and make it difficult to interpret nonsignificant results (e.g., in our sample of 30 studies, 15 reported years of education and 7 reported parental socioeconomic status, but only 4 reported both). As a result, we carried out multivariate analyses only with moderators that were significant in bivariate analyses, and we carried out these analyses only on data sets containing enough observations to have confidence in the results of these analyses (see *Results* below).

Effect Size Estimates

For each study, a standardized effect size (Cohen's d , the mean difference between groups divided by the pooled standard deviation) was estimated for performance (accuracy and, when reported, reaction time [RT]) and DLPFC activation. Because a large majority of included studies reported accuracy only in terms of the total number or proportion of correct or incorrect responses, these measures were used in preference to d' or similar measures. We adjusted all effect sizes by using the Hedges and Olkin (1985) correction for small sample bias. Effect estimates were made from the best available data in each study, in the following priority: (a) reported means and standard deviations for each group or a direct report of Cohen's d ; (b) the results of t or one degree of freedom F tests (test statistics or p values); (c) Z scores of the contrast in activation between groups, typically output from statistical parametric mapping (Wellcome Department of Imaging Neuroscience); (d) means and standard errors or raw data estimated from published figures (bar graphs, line graphs, or scatterplots); and (e) an in-text indication of whether a significant difference between groups was observed: When the difference was not significant, the effect size was estimated as zero, whereas when the difference was significant, the effect size was estimated as the smallest effect required to achieve significance given the sample size and reported significance threshold used in the study in question. If none of the first four types of data were available, we contacted the authors with a request to supply the required data. For one study (Manoach et al., 2000), a reported mean was assumed to have a misplaced decimal point, because it differed from other means of the same type by an order of magnitude and was inconsistent with the p value reported by the authors (moving the decimal place made the p value consistent). All estimation procedures are presented in Appendix B.

Because variations in the quality of reported data across studies meant that a number of our effect size estimates probably reflected poorer estimates of the actual effect that occurred in a study, we carried out a set of restricted analyses that excluded these studies.¹ Consequently, results of all analyses will be reported for both the full set of studies and for a restricted set that excludes those effects discussed above.

In nearly all included studies, multiple effect estimates were available, which were treated differently according to how they arose. Specifically, when data were available for multiple scanning sessions, only data for the first scanning session were included, thereby avoiding the influence of practice effects or treatments (e.g., pharmaceutical) between sessions. For event-related functional MRI studies (three studies), which presented data for multiple scans over a single epoch, effects for each scan were averaged to allow for comparison with blocked designs. Similarly, when the magnitude of an activation was reported by using multiple dependent variables (e.g., maximum voxel change, number of activated voxels, etc.), effects from each were averaged together (two studies).

When data were available for different WM tasks in the same scanning session or for different difficulty levels of the same task, two approaches were taken. The first was to include performance and activation data for all tasks as though they were observations on independent samples, thereby using all of the available data. This approach was not taken in analyzing whether other study characteristics (see below) moderated activation or performance effect sizes because study characteristics do not vary within a study, and this would have resulted in variation in the activation or performance effect size (possibly due to task differences) without the possibility of corresponding variation in the moderator variable. However, because this approach violates the statistical assumption of the independence of observations, we also carried out four analyses in which data from only one task per study were included from each sample. Selection of tasks for each analysis was made pseudorandomly, such that every WM task within a study was included in at least one of the four analyses, and if there were two tasks in a study, each was included in two analyses. In order to simplify data reporting, all statistics reported for these analyses are the mean statistic obtained from the four analyses. Because these two approaches produced roughly equivalent results in all analyses (i.e., in no case was one statistically significant while the other was not), we report data only for the independent analyses (except, as noted, in Figure 2).

Finally, when multiple activations occurred in the DLPFC of a given hemisphere, these were averaged together to form a single effect for that hemisphere. Because 2 studies reported data only for both hemispheres combined, in studies reporting DLPFC activation for each hemisphere separately we averaged effect sizes for both hemispheres to create a composite activation effect size so that all 30 studies could be included in a single analysis. Analyses were carried out on each hemisphere separately as well as on this set of composite effect sizes.

Definition of DLPFC

To determine which activations should be treated as having occurred within DLPFC, we used a broad definition of DLPFC based on the results of a cytoarchitectural study of Brodmann's areas (BAs) 9 and 46. Rajkowska and Goldman-Rakic (1995) localized BAs 9 and 46 on lateral photographs of five normal brains by examining the cytoarchitecture of

¹ Specifically, effects were excluded from this analysis if their estimation could be made only from an indication of whether the effect was or was not significant. In addition, for one study, activation and performance effect estimates could be made only by collapsing across two roughly equivalent tasks in different modalities (Barch, Csermansky, Conturo, & Snyder, 2002), and for another study (Kim et al., 2003) by using means and standard deviations for brain regions within DLPFC that differed in each group (and was thus not a direct comparison of activation in the same DLPFC region).

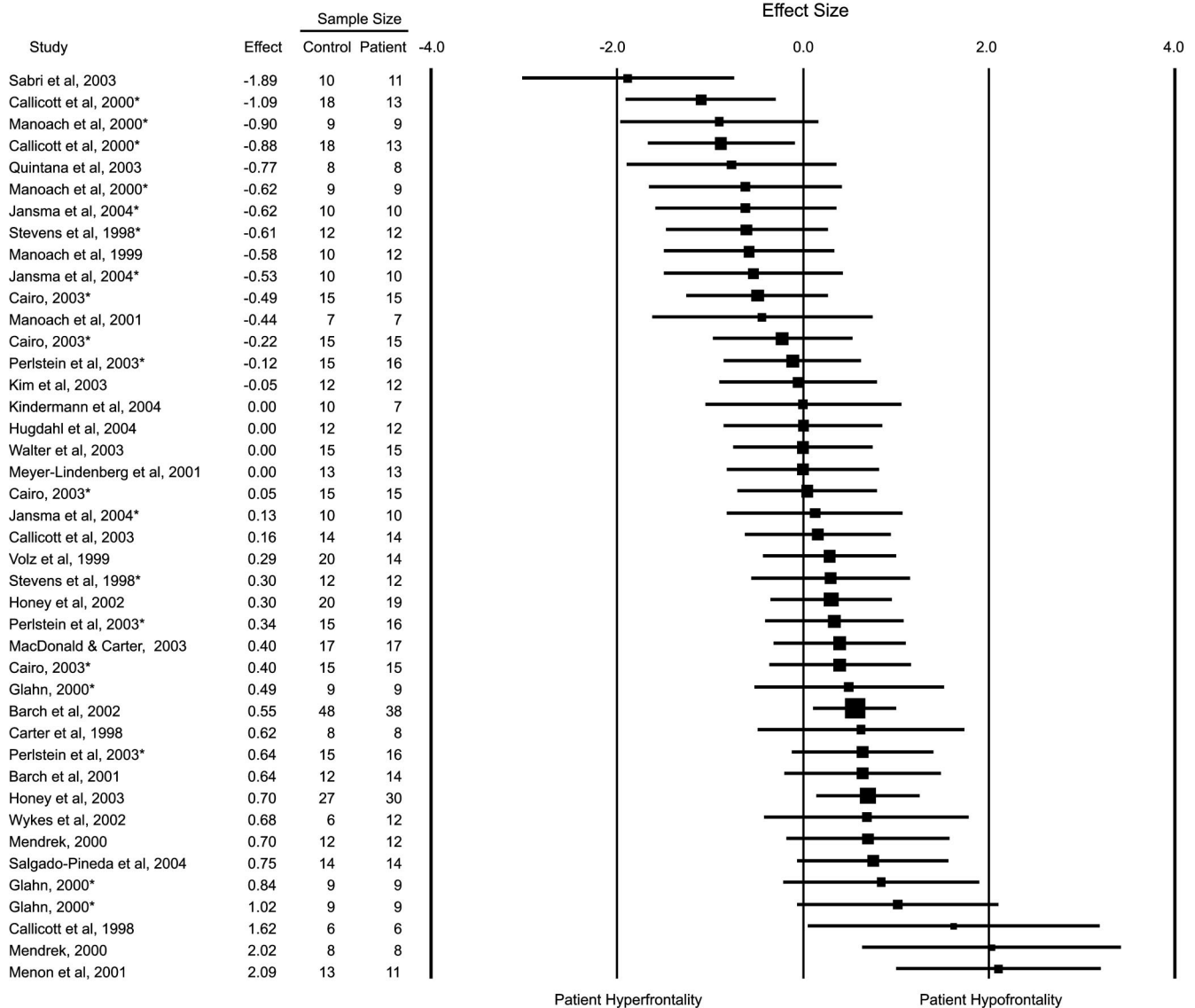


Figure 2. Forest plot of all combined hemisphere effects used in the most inclusive (dependent) analysis. Asterisks indicate effects that are based on the same sample as another effect in the dependent analysis.

prefrontal coronal slices from these brains, transposed the area they occupied onto a lateral rendition of the left hemisphere from the stereotaxic atlas of Talairach and Tournoux (1988), and reported the range of coordinates in the anterior-posterior (AP) and dorsal-ventral (DV) dimensions that they occupied. Because five individuals probably do not adequately characterize the range of variability in the population, we elected to use the limits of variability observed in the location of these regions rather than the more conservative area in which BAs 9 and 46 overlapped in all five brains. These limits were from +64 mm to +14 mm in the AP dimension and from +57 mm to -10 mm in the DV dimension.

In order to establish which BAs, according to the Talairach and Tournoux (1988) atlas, were included in this coordinate range, we localized each mm³ point in the volume described by the above coordinate range (within the full extent of the lateral dimension for one hemisphere) by using the Talairach Daemon (Lancaster et al., 2000), excluding axial slice +35 because of an error in the reporting of prefrontal BAs on this slice (Maldjian, Laurienti, & Burdette, 2004). The results of this localization are

presented in Table 1, along with activation foci from studies reporting stereotaxic coordinates included in the analysis. Data from studies not reporting stereotaxic coordinates were included if the authors indicated that the data pertained to DLPFC. Anterior cingulate and temporal lobe activations were excluded, irrespective of whether they were within the Rajkowska and Goldman-Rakic (1995) coordinate range (which did not include the lateral dimension). To compare activations reported in the coordinate system of the Montreal Neurological Institute (MNI) templates to these criteria, we transformed MNI coordinates to Talairach coordinates by using the nonlinear transformation provided by Brett (2002; for discussion, see Brett, Johnsrude, & Owen, 2002). When authors analyzed their data by using statistical parametric mapping and did not explicitly state that coordinates were reported according to the Talairach atlas, we assumed that MNI coordinates were reported.

One problem with this approach is that because of the considerable variability observed in the Rajkowska and Goldman-Rakic (1995) sample, the selected coordinate range would likely include activations in a considerable

Table 1
Prefrontal Cortical Regions in Rajkowska and Goldman-Rakic (1995) Coordinate Range (64 > y > 14; 57 > z > -10)

Region (BA)	% of all 1-mm ³ cortical regions in range	No. included activations from studies reporting MNI or Talairach coordinates (%) ^a
Included regions		
6	4.8	3 (8.6)
8	17.0	3 (8.6)
9	18.1	9 (25.7)
10	23.5	6 (17.1)
11	0.4	3 (8.6)
44	1.3	0 (0.0)
45	4.6	1 (2.9)
46	9.3	7 (20.0)
47	6.1	3 (8.6)
Excluded regions		
13	2.1	
24	2.8	
25	0.3	
32	9.1	
33	0.1	
38	2.9	

Note. BA = Brodmann's area; MNI = Montreal Neurological Institute.
^a Nearest BA.

expanse of the lateral cortical surface, which almost certainly have a different functional significance than activations occurring in a more strictly defined DLPFC region of interest. However, using a more stringent coordinate range for DLPFC—that is, the region in which DLPFC overlapped in all five brains in the Rajkowska and Goldman-Rakic (1995) sample, from +53 mm to +26 mm AP and +50 mm to +14 mm DV—would result in a substantial loss of data, as a large number of studies reporting Talairach coordinates did not find group differences in activation in this restricted coordinate range (and thus did not publish data allowing an estimate of the magnitude of the difference in this region). Consequently, we elected to use the broader coordinate range for our primary analysis (at the expense of specificity with regard to DLPFC as opposed to prefrontal function in general), although we also report on an analysis that followed the exact same procedures but used the smaller coordinate range described above.

Moderator Analyses

Because meta-analysis is particularly well suited to identifying variables that moderate the magnitude of an effect size, a number of additional variables were coded across studies and analyzed as potential moderators (see Table 2 for summary data; the entire data set, along with effect sizes for each analysis, is available on the Web at <http://dx.doi.org/10.1037/0894-4105.20.5.497.supp>). Only variables reported in five or more studies (or five at each of two levels of a categorical variable) were tested. Because of the number of analyses this entailed, the Dunn-Šidák correction for multiple comparisons was used to hold family-wise error at 0.10 (with separate adjustments for moderators of activation, accuracy, and RT; in addition, overall, positive, and negative symptom scores were treated as a single test for the purpose of family-wise error correction).

To derive estimates of symptom severity in patient samples that were comparable across different symptom assessment scales, we divided the mean scale score reported for a sample by the maximum possible score on the scale (with an adjustment for scales in which the lowest possible value on any item was 1, rather than 0), thereby producing a value for each study

on a scale from 0 to 1, with 0 indicating that all participants received the minimum score on every item on the scale and with 1 indicating that all participants received the maximum score on every item on the scale. Scores were calculated separately for positive symptoms (Positive and Negative Syndrome Scale [PANSS; Kay, Fiszbein, & Olper, 1987] positive score or Scale for the Assessment of Positive Symptoms [Andreasen, 1984] score), negative symptoms (PANSS negative score or Scale for the Assessment of Negative Symptoms [Andreasen, 1984] score), and overall pathology (Brief Psychiatric Rating Scale [Overall & Gorham, 1962]; PANSS total score; Signs and Symptoms of Psychotic Illness [Liddle, Ngan, Duffield, Kho, & Warren, 2002]; Psychiatric Symptom Assessment Scale [Bigelow & Berthot, 1989]; or, if none of these were available, from the average of the Scale for the Assessment of Positive Symptoms and Scale for the Assessment of Negative Symptoms or PANSS positive and PANSS negative scores).

Finally, if authors did not report the voxel size used in statistical analyses directly, we calculated voxel size from the in-plane resolution and slice thickness (plus the gap between slices, if any) of the functional imaging sequence whenever these data were reported.

Results

Evaluation of the Relationship Between Performance Differences and WM Load

To evaluate whether between-group differences in task performance are related to WM load, we carried out an additional,

Table 2
Variables Coded for Moderator Analysis

Variable	No. of studies reporting	Descriptives	
		<i>M</i>	<i>SD</i>
Imaging technology used	30		
PET	4		
fMRI	26		
Voxel size (mm ³) ^a	26	57.34	27.57
Material type ^a	30		
Verbal	18		
Nonverbal	5		
Mixed	7		
Reward for task performance	30		
Yes	4		
No	26		
Patient status ^{a,b}	21		
Inpatient	9		
Outpatient	12		
Symptom severity ^a	21		
Overall	21	0.21	0.09
Positive	16	0.20	0.14
Negative	16	0.29	0.10
Age at illness onset ^a	5	22.66	1.19
Length of illness (years) ^a	14	12.39	6.43
Patient gender (% male) ^a	27	70.73	16.93
Age ^{a,c}	24	0.23	0.44
Education (years) ^{a,c}	15	1.19	1.08
Parental education (years) ^{a,c}	7	0.27	0.15
Parental SES ^{a,c}	7	0.12	0.37
National Adult Reading Test score ^{a,c}	7	0.91	1.09

Note. PET = positron emission tomography; fMRI = functional MRI; SES = socioeconomic status.

^a Variables included in moderator analysis. ^b Studies reporting on mixed samples were excluded from moderator analyses. ^c Values reflect a standardized mean difference between patients and controls; positive values indicate lower values for patients, except age.

limited, analysis on performance effect sizes in a subset of the included studies. This analysis included all studies in the primary analysis that reported performance data on the same set of patient and control participants for two or more loads of the same working memory task (e.g., 0-back and 2-back loads of the *n*-back task). A paired-sample *t* test of the difference in performance effect sizes between the lowest and highest WM loads used in each study indicated significantly larger performance effects (indicating better performance by controls than patients) at higher WM loads for accuracy, $d = 0.68$, $t(14) = 2.68$, $k = 15$, $p = .018$, thereby supporting the existence of a relationship between group differences in accuracy and WM task load. The results for RT generally revealed a similar pattern despite the fact that fewer studies reported RT as a performance measure. Excluding a single outlier (two standard deviations away from the mean), patients tended to show greater deficits in RT relative to controls at higher WM loads, $d = 0.62$, $t(9) = 1.96$, $k = 10$, $p = .082$, although the effect was diminished in magnitude when the outlier was included, $d = 0.30$, $t(10) = 1.01$, $k = 11$, $p = .336$.

Effect Size Estimates

Effect size estimates and confidence intervals for patient–control differences in DLPFC activation are presented in Table 3 (descriptive data for unweighted effect estimates are available on the Web at <http://dx.doi.org/10.1037/0894-4105.20.5.497.supp>). The average effect size was not significantly different from zero in any analysis. Homogeneity tests in all analyses indicated that the set of activation effect sizes was heterogeneous (all $ps < .0005$). A forest plot of all combined hemisphere activation effect sizes is shown in Figure 2.

Analyses of performance differences between patients and controls indicated that patient samples were significantly impaired in terms of both accuracy ($d = 0.88$; 95% CI = 0.66–1.09; $k = 30$; $p < .0005$) and RT ($d = 0.88$; 95% CI = 0.60–1.15; $k = 21$; $p < .0005$). One study was an extreme outlier (>4 standard deviations from the mean) on accuracy. Consequently, we repeated all of our analyses excluding this outlier, with the same results. All results reported below involving accuracy exclude the outlying study.

Moderator Analyses

Analyses of accuracy and RT as moderators of patient–control differences in DLPFC activation are shown in Table 4. Broadly

speaking, these analyses showed that patients demonstrated increasing hypofrontality when they exhibited greater performance deficits relative to controls, although this was true of accuracy only for differences in activation of the right hemisphere. Scatterplots of DLPFC activation differences and performance differences are shown by hemisphere for accuracy and RT in Figure 3. As shown in Table 5, the analysis on the limited definition of DLPFC produced roughly the same results as the primary analysis, although many of the effects were only marginally significant.

The only other significant moderators (from Table 2) of patient–control differences in DLPFC activation were voxel size and the group difference in age. Greater spatial resolution (smaller voxel size) was associated with findings of greater DLPFC activation in patients relative to controls in all analyses (for the inclusive analysis of both hemispheres, $r = .47$; $r^2 = .22$; $k = 26$; $p = .0002$). This relationship is shown in Figure 4. In addition, as the effect size of the age difference between patients and controls increased (indicating greater patient age relative to controls), patients demonstrated greater activation relative to controls in the left hemisphere (for the inclusive analysis, $r = .30$; $r^2 = .09$; $k = 22$; $p = .007$; although the restricted analysis was not significant after correction for multiple comparisons, this probably reflected low power, as there were only 13 studies in this analysis), but not the right hemisphere (for the inclusive analysis, $r = .04$; $r^2 = .00$; $k = 22$; $p = .693$). This relationship is shown in Figure 5. No variable was a significant moderator of either accuracy or RT.

Multivariate Moderator Analyses

To evaluate the relative effects of performance variables (accuracy vs. RT) on patient–control differences in DLPFC activation during WM tasks, we carried out an analysis in which both accuracy and RT were entered as moderators together; following this, we analyzed a model in which the other significant moderators of group differences in DLPFC activation were also included (standardized group difference in age as well as voxel size). The results of these analyses are presented in Table 6 for all of our analysis sets, excluding those for the restricted data sets because of concerns about loss of sample size (and thus power) and the consequent unreliability of parameter estimates in these models. These analyses indicate that relative to accuracy, RT remained a strong predictor of group

Table 3
Results of DLPFC (Large Definition) Activation Analyses

Analysis set	Average <i>d</i>	95% CI	<i>p</i>	<i>k</i>	<i>N</i>
Inclusive analyses					
Combined	0.20	−0.05, 0.44	.13	30	808
Left hemisphere	0.23	−0.05, 0.51	.11	28	776
Right hemisphere	0.15	−0.13, 0.42	.34	28	776
Restricted analyses					
Combined	0.21	−0.25, 0.67	.38	17	345
Left hemisphere	0.38	−0.08, 0.83	.12	17	415
Right hemisphere	0.04	−0.50, 0.58	.88	14	344

Note. Positive values of *d* indicate patient hypofrontality. DLPFC = dorsolateral prefrontal cortex; CI = confidence interval; *k* = number of studies; *N* = number of participants in all included studies.

Table 4
Analyses of Performance as a Moderator of DLPFC (Large Definition) Activation

Analysis set	Accuracy					Reaction time				
	<i>r</i>	<i>r</i> ²	β	ρ	<i>k</i>	<i>r</i>	<i>r</i> ²	β	ρ	<i>k</i>
Inclusive analyses										
Combined	.19	.04	.28	.092	29	.43	.19	.50	.002	20
Left hemisphere	.04	.00	.04	.736	27	.43	.18	.55	.001	18
Right hemisphere	.34	.12	.53	.003	27	.41	.16	.43	.006	18
Restricted analyses										
Combined	.33	.11	.60	.042	14	.66	.43	.84	.001	12
Left hemisphere	.10	.01	.15	.487	16	.49	.24	.75	.001	13
Right hemisphere	.40	.16	.85	.006	13	.70	.49	.83	.002	10

Note. Positive values of *r* indicate a positive correlation between the effect sizes of patient performance deficits and patient hypofrontality. DLPFC = dorsolateral prefrontal cortex; *k* = number of studies.

differences in DLPFC activation. Moreover, RT retains its predictive power regardless of which other variables are included in the model, particularly for the left and composite (left and right) DLPFC analyses (Table 6).

Discussion

Overview of Results

Primary analyses. The main finding in this study was that the magnitude of WM impairment in schizophrenia was associated with differences in DLPFC activation between patients and controls. This finding indicates that the nature of DLPFC dysfunction in patients with schizophrenia is less straightforward than has been generally appreciated. Furthermore, the fact that the average effect size of patient-control differences in DLPFC activation was not significantly different from zero challenges the notion of hypofrontality as a general characteristic of prefrontal dysfunction in schizophrenia, at least in WM tasks. Indeed, given the role of WM task performance in differential DLPFC activation, characterizing functional pathology in higher cortical regions as either too much or too little may be overly simplistic. Finally, the identification of RT as a significant moderator of differences in DLPFC activation is consistent with Manoach's (2002, 2003) and Callicott et al.'s (2003) hypothesis of a shifted inverted-U relationship between DLPFC activation and WM performance in schizophrenia, which, as shown in Figure 1 and discussed in the introduction, predicts the results of the present study.

Given that DLPFC is thought to play a number of superordinate executive roles in WM in normal subjects (e.g., strategic organization or chunking of to-be-remembered material, manipulation and ordering of the contents of the memory store, and strategy shifting; see Rypma, 2006; Smith & Jonides, 1999), and given that at least some of these have been shown to be dissociable (Rypma, 2006), it is difficult to arrive at an unequivocal interpretation of our results. Reduced activation of DLPFC could result from intentional or unintentional failures to use some or all of these strategies. Thus, our results may reflect relatively intentional failures to perform the task because of a lack of motivation or other causes, structural or functional deficits in DLPFC that make the deploy-

ment of normal WM strategies nonoptimal or impossible for patients to use (and whose exact implications for brain function and task performance may depend critically on the nature and extent of the task load), and/or deficits in related brain regions other than DLPFC that have similar consequences for strategy deployment. We now consider each of these interpretations.

One possibility is that reduced DLPFC activation is a direct consequence of patients failing to perform the task normally; that is, at higher difficulty levels, poorly performing patients are not sufficiently activating DLPFC because they are not (willfully or otherwise) actually engaging WM. The fact that patients achieve greater than chance performance (as they did in all studies in our analysis), however, suggests that patients are at least minimally engaged in WM tasks. Thus, if present, task disengagement is likely partial in nature and could reflect a host of responses, including frustration, fatigue, guessing, or poor effort. It is also possible that task disengagement could be present in a subset of the patient sample or on a subset of study trials. This may be explained by the subsequent-memory effect described by Rypma and D'Esposito (2003), who demonstrated greater DLPFC activation in normal participants during successful encoding than during unsuccessful encoding in a WM task, which could potentially reflect task disengagement on a trial-by-trial basis. This interpretation, however, does not address the finding of greater DLPFC activation by patients in some studies under conditions of near-comparable task performance.

Another possibility is that hypofrontality in poorly performing patients reflects the use of alternative, non- (or less) DLPFC mediated, strategies to perform the task, either as a result of poor strategy choice or prefrontal pathology that makes alternative strategies optimal for these patients. In this case, patients would attempt to compensate through activation of alternative circuits not used by controls (Stern, 2002). Indeed, this interpretation could also account for findings of hyperfrontality when patients perform relatively well in a way analogous to the inverted-U hypothesis. That is, a disruption of the structure or connectivity of DLPFC or other regions involved in WM could result in normal but inefficient strategy usage (resulting in greater activation of DLPFC) on less difficult tasks, as well as shifts in strategy and engagement of

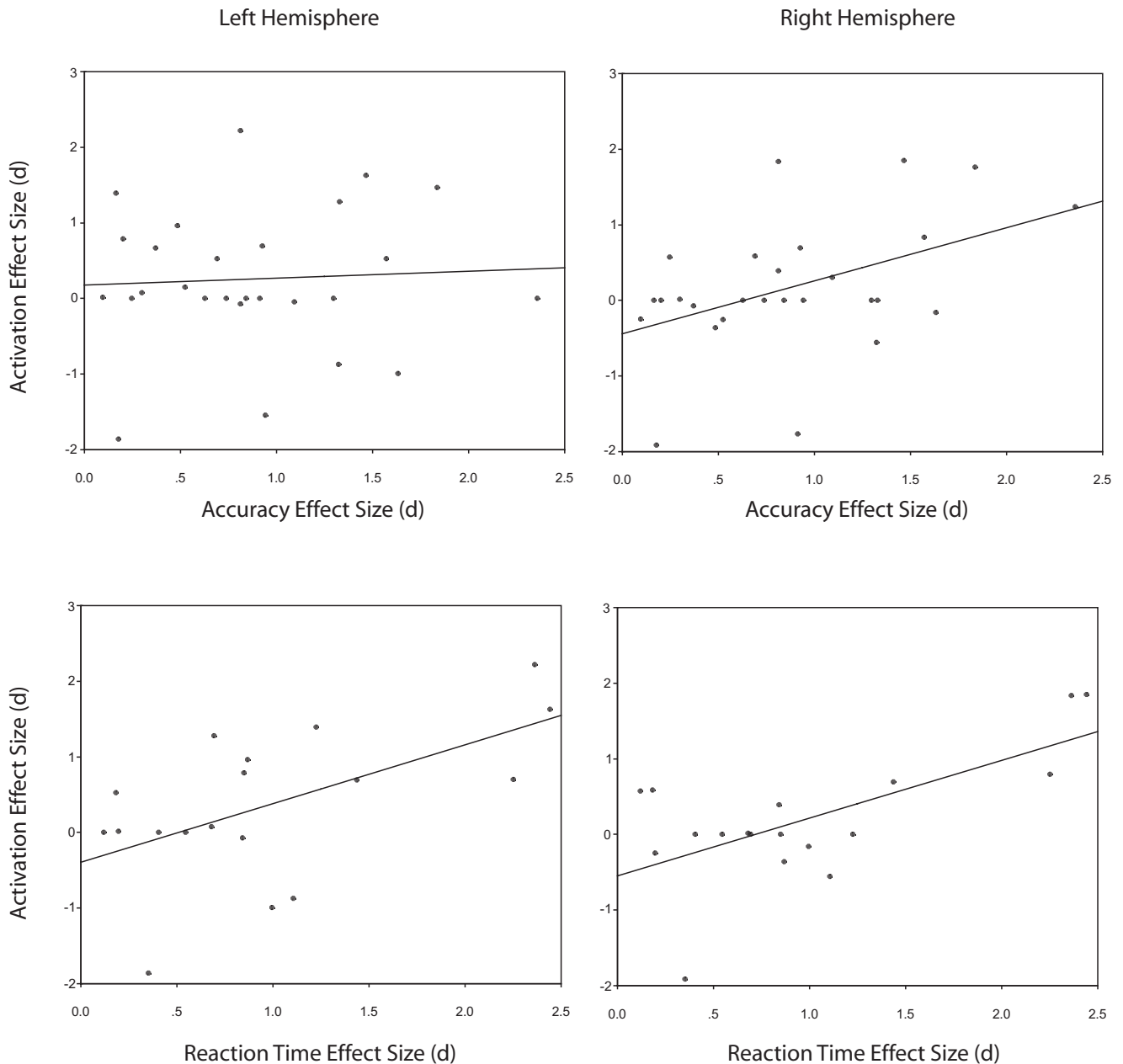


Figure 3. Scatterplots and (unweighted) regression lines for the inclusive analysis of group differences in performance and activation, from one of the independent analyses. Larger effect sizes indicate less activation or poorer performance by patients relative to controls.

neurocircuitry that result in less DLPFC activity under conditions of high difficulty and hence poor performance.

In addition, our results should be considered in light of current theories regarding disturbances of cortico-cortical functional connectivity in schizophrenia (e.g., Glahn et al., 2005; Meyer-Lindenberg et al., 2005; Schlösser et al., 2003). Although we examined group differences in activation only in lateral prefrontal cortex and so can not speak directly to the possibility of disturbances in functional connectivity, it is clear that disturbances in DLPFC activation in schizophrenia bear a complex relationship with WM

performance and that this relationship is probably best viewed in light of interrelationships between DLPFC activation and activation of other cortical and subcortical areas involved in WM, as alluded to above. Recent research has identified several brain regions that are associated with activation changes in DLPFC, including (but not limited to) medial frontal cortex (Glahn et al., 2005), thalamus (Andreasen, Paradiso, & O'Leary, 1998), medial temporal lobe regions (Meyer-Lindenberg et al., 2005), and posterior parietal cortex (Kim et al., 2003; Quintana et al., 2003). In order to further understand the nature of dysfunctional WM net-

Table 5
Analyses of Performance as a Moderator of DLPFC (Limited Definition) Activation

Analysis set	Accuracy					Reaction time				
	<i>r</i>	<i>r</i> ²	β	ρ	<i>k</i>	<i>r</i>	<i>r</i> ²	β	ρ	<i>k</i>
Inclusive analyses										
Combined	.13	.02	.13	.395	29	.34	.11	.25	.067	20
Left hemisphere	-.06	.00	-.05	.744	27	.29	.09	.23	.098	18
Right hemisphere	.22	.05	.31	.060	27	.26	.07	.23	.121	18
Restricted analyses										
Combined	-.10	.01	-.02	.676	9	.63	.40	.99	.034	8
Left hemisphere	-.24	.06	-.45	.287	7	.54	.30	1.12	.045	6
Right hemisphere	.27	.07	.60	.067	10	.71	.50	.83	.013	8

Note. Positive values of *r* indicate a positive correlation between the effect sizes of patient performance deficits and patient hypofrontality. DLPFC = dorsolateral prefrontal cortex; *k* = number of studies.

works in schizophrenia, it will be important to evaluate patterns of coactivation in these candidate network structures in reference to the performance-dependent DLPFC activity we describe. It may be particularly informative to assess these changes when patient performance is comparable to controls, as it is under these conditions that DLPFC is engaged in patients, albeit abnormally (at least in some cases)—as evidenced by hyperfrontality. Thus, relative to DLPFC activity, the observation of corresponding coactivation (or inactivation) in other target structures is more likely to reflect dysfunction involving other nodes within the larger WM network.

Restricted analysis. It is worth noting that group differences in performance tended to account for substantially more of the variability in group differences in activation in the restricted analyses (particularly for RT), suggesting that poorer estimates of DLPFC activation tended to obscure its relationship with WM task performance. Indeed, in the restricted analysis, RT accounted for a substantial amount, ranging from 24% to 50%, of the variability in group differences in DLPFC activation.

Performance analyses. It was unexpected that group differences in RT would better predict group differences in activation in

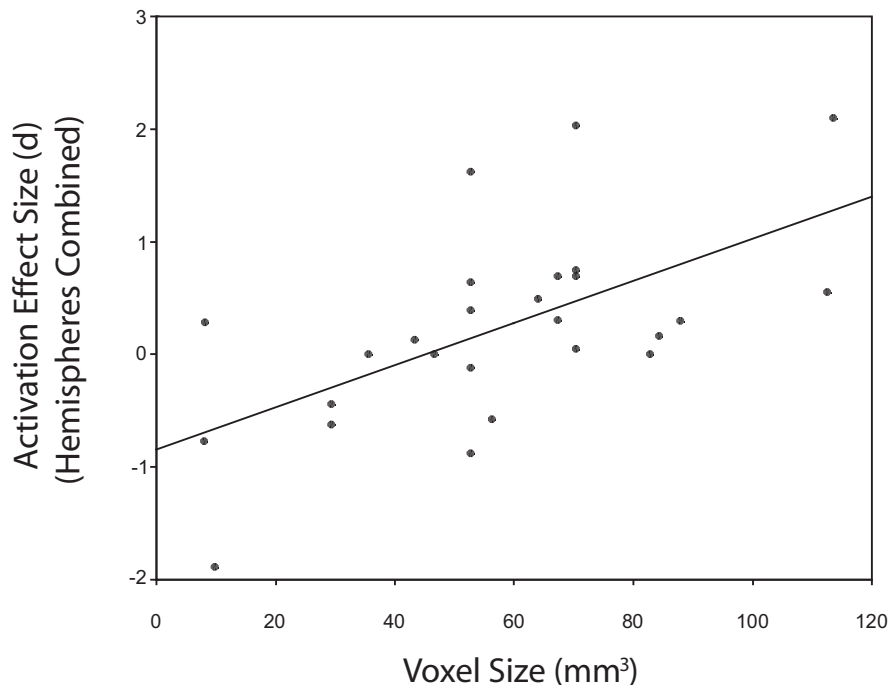


Figure 4. Scatterplot and (unweighted) regression line for combined hemisphere effects and voxel size, from one of the independent analyses. Larger effect sizes indicate less activation by patients relative to controls.

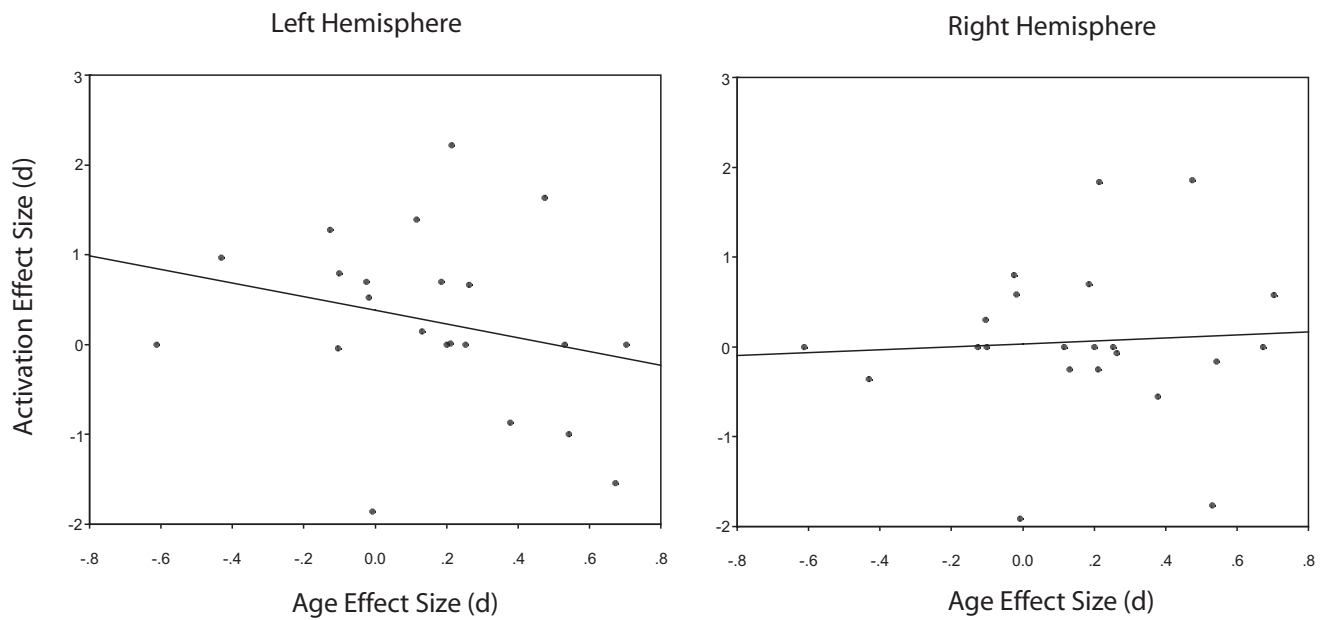


Figure 5. Scatterplots and (unweighted) regression lines for group differences in activation and age, from one of the independent analyses. Larger effect sizes indicate less activation by, or greater age of, patients relative to controls.

WM tasks than did accuracy; in fact, the effect of accuracy generally did not remain significant when RT was simultaneously included as a moderator. A potential interpretation is that because control participants were performing with very high accuracy in many studies (typically greater than 90%), some of the relationship between accuracy and DLPFC activation may have been obscured by ceiling effects. Thus, RT may be more sensitive than accuracy to the differences across studies that are driving the relationship between WM performance and DLPFC activation.

It is also important to note that in no study of patient-control differences in DLPFC activation have patients exhibited WM task

performance (accuracy or RT) that was superior to that of control patients. Although several studies found no significant differences in performance, a nonsignificant difference is not the same as no difference, particularly given the small sample sizes generally used in this literature. Given the moderating influence of WM task performance on DLPFC activation, it will be necessary to carry out studies in which patient and control samples are matched on performance of the task to be used during scanning in order to understand functional DLPFC dysfunction in schizophrenia. Although this is problematic in that it will tend to select high-functioning patients (and/or low-functioning controls), it may be

Table 6
Results of Multivariate Moderator Analyses of Patient-Control Differences in DLPFC Activation Using All Significant Variables in Bivariate Analyses

Analysis	r^2 and no. of included studies					
	Large DLPFC definition			Limited DLPFC definition		
	Composite	Left	Right	Composite	Left	Right
Analysis 1	19	17	17	19	17	17
Accuracy	.00	.02	.03	.02	.13*	.01
RT	.17*	.22*	.08	.21*	.25*	.09
Analysis 2	17	15	15	17	15	15
Accuracy	.02	.06	.01	.05	.11	.02
RT	.16*	.22*	.06	.20*	.28*	.06
Age	.01	.00	.20*	.00	.08	.18*
Voxel size	.20*	.12*	.31*	.08	.00	.20*

Note. Integer values indicate the number of studies included in the analysis, whereas all other numbers are r^2 for the moderator variable. All significant effects are in the same direction as in the bivariate analysis. DLPFC = dorsolateral prefrontal cortex; RT = reaction time.

* $p < .05$.

necessary to discriminate WM activation differences due to schizophrenic illness from those due to poor task performance.

Other Moderators

Voxel size. The significance of voxel size as a moderator of patient–control differences in DLPFC activation is surprising, interpretively challenging, and potentially of considerable importance. Voxel size accounted for a substantial amount (22%) of the total variability in DLPFC activation differences between patient and control groups. This finding demonstrates that greater attention needs to be paid to variables related to functional image acquisition. The fact that findings of greater DLPFC activation by patients are associated with superior spatial resolution in functional images suggests that hyperfrontality may be underestimated in schizophrenia. Although it is likely significant that patient hyperfrontality was never observed with voxel sizes greater than 60 mm³ (see Figure 3), what this means in terms of the nature of the hemodynamic response in patient and control participants in DLPFC is unclear. One possibility is that patients may exhibit a widespread reduction in activation, but with circumscribed regions of hyperactivity. If this were the case, only studies with high spatial resolution would be expected to find significantly greater activation in prefrontal regions by patients.

It should be noted that the effect of voxel size declined uniformly across the combined-, right-, and left-hemisphere multivariate analyses under the limited definition of DLPFC (with accuracy, RT, age, and voxel size as predictors). This may be an artifact of the considerable loss of variability in the activation effect sizes used in these multivariate limited DLPFC analyses. Specifically, these analyses set a large number of the included effects to zero (because of the lack of activation differences within the highly circumscribed region considered to be DLPFC in this analysis), thereby artificially restricting the variability of the predicted variable. Despite this problem, the effect of other variables in the model (except voxel size) remained comparable for the large and limited DLPFC definitions (Table 6).

Age. The significance of group differences in age as a moderator of DLPFC activation underscores the importance of using carefully matched samples in neuroimaging investigations of schizophrenia and other patient populations. The impact of aging on brain function is a literature unto itself, and unless a researcher is investigating the possibility of an interaction between age and illness, it will be important to avoid confounding the influence of age with diagnosis. It may also be the case that other demographic variables have confounding influences as well: Given that in the present study age was the demographic variable for which we had the greatest power (apart from the gender composition of patient samples) to detect an effect on group differences in DLPFC activation, it remains possible that other variables would have emerged as significant moderators had they been reported more frequently in the literature.

Alternative Explanations

It should also be recognized that although our findings are predicted by a model in which patients exhibit a left shift in an inverted-U relationship between DLPFC activation and task difficulty or WM load, it is not necessary to posit an inverted-U

relationship to explain the present findings. These results would also be obtained if both patients and controls showed monotone increases in DLPFC activation at greater WM loads, but with patients tending to activate DLPFC more than controls at low WM loads and reaching asymptotic levels of activation sooner (perhaps due to a physiological constraint on maximal DLPFC activation). However, there is evidence that argues against this interpretation. Specifically, an inverted-U relationship between activation of several brain regions, including DLPFC, and the load level of a version of the *n*-back task has been observed in normal participants in one study (Callicott et al., 1999), and the only study thus far to use the 3-back level of an *n*-back task with schizophrenic patients (Jansma, Ramsey, van der Wee, & Kahn, 2004) found that patients exhibited significant reductions in DLPFC activation from 2-back to 3-back loads that were not observed in control participants. More definitive support for a left-shifted inverted-U relationship between WM load and DLPFC activation in patients with schizophrenia would be obtained if the results of both of these studies could be replicated by using a sufficient range of task loads (e.g., up to a 4-back load) to show reductions in DLPFC activation for both patients and controls in a single study, but with controls showing reductions in activation only at higher load levels than patients.

Results of this study are also consistent with an alternative inverted-U model proposed by Callicott et al. (2003), which does not involve a left shift in the curve for patients with schizophrenia, but rather a shorter distal arm of the curve for patients such that they show equivalent activation to controls at low WM loads, but their activation of DLPFC begins to decrease sooner. Although Callicott et al.'s (2003) data did not support this alternative, it cannot be ruled out in the present study, particularly as Figure 3 does not provide definitive evidence of overall hyperfrontality by patients at the lowest WM loads used in the literature. However, it is possible that the literature available for inclusion in this study simply does not contain tasks with sufficiently low WM loads to reliably demonstrate hyperfrontality in patients with schizophrenia. Although the present study is not equipped to distinguish between these two possibilities, it nonetheless underscores that DLPFC activation differences between patient and control participants are dependent on the WM load of the task used in a given study.

Study Limitations

Sample characteristics. Some caveats should be considered in interpreting our results. First, it is conceivable that the findings could reflect differences in patient characteristics across studies (e.g., high-functioning versus low-functioning patient samples), rather than difference in WM load, which we assumed was indexed by patient–control differences in task performance. However, we were unable to identify any patient characteristics that moderated group differences in performance, and our analysis of performance effect sizes suggested that they adequately characterized task difficulty based on WM load, at least within a given task type. In addition, to the extent that poorly performing patients are effectively under a greater WM load than better performing patients or healthy participants performing the same task, our results would still support an inverted-U model. That is, one can consider task difficulty or WM load to be a function not only of task parameters

but of the WM capabilities of individual participants, in which case task performance may be an ideal means of quantifying task difficulty across studies with substantially different patient samples. This is because, given constant task demands, individuals with reduced WM capacity would effectively be performing under greater WM load, which should result in both poorer performance and a greater left shift in the hypothesized inverted U.

Definition of DLPFC. There are also limitations in the present study with regard to the definition of DLPFC used. Our primary analysis used a definition of DLPFC large enough that it may be more appropriate to consider the results as general characteristics of lateral prefrontal cortex function rather than of DLPFC function in particular. The additional analysis that was limited to a region that is unequivocally DLPFC did, however, produce broadly the same results as the primary analysis. Although the relationship between RT and DLPFC activation remained significant only in the restricted analysis, all of those relationships that were significant in our primary analysis remained marginally significant under the more stringent definition of DLPFC. Given the considerable loss of data involved in this limited definition of DLPFC, these results nonetheless indicate that the relationships identified in the primary analysis are not driven solely by lateral prefrontal regions outside of DLPFC. Consequently, although it is possible—even likely—that the relationship between group differences in performance and activation applies to cortical prefrontal regions contiguous with DLPFC, our results do indicate that this relationship exists within DLPFC in particular.

Material type. Another limitation on the conclusions that can be drawn from our study is the fact that studies in our analysis primarily used verbal task materials, and so our results may be specific to verbal WM. It is known that verbal, spatial, and object WM recruit different regions of prefrontal cortex (Smith & Jonides, 1999), although each of them do activate DLPFC. Although we did not detect any difference in DLPFC activation differences in verbal versus nonverbal tasks, there were only five studies that used clearly nonverbal tasks that could be included in this analysis, and therefore, we may simply have lacked the power to detect an effect of material type. Indeed, the preponderance of verbal materials in included studies may explain why RT effects on the group difference in DLPFC activation were strongest in left- or combined-hemisphere analyses when voxel size and group differences in accuracy, RT, and age were included as moderators.

Conclusion

The present study reveals a number of important features of patient-control differences in DLPFC activation during the performance of WM tasks. First, it appears that it is inappropriate to characterize prefrontal dysfunction in schizophrenia as being a pattern of generalized hypofrontality, at least in the context of WM task performance. Our results, taken in combination with the experimental results of Callicott et al. (1999) and Jansma et al. (2004) discussed above, indicate that the hypothesis advanced by Manoach (2002, 2003) and Callicott et al. (2003) that patients exhibit a left-shifted inverted-U function between task difficulty and DLPFC activation can be empirically supported, although a direct experimental demonstration is still needed. This also raises several broader issues. It may be that a similar relationship exists in other brain regions that respond in a load-dependent fashion to

task difficulty, that nonmonotone changes in the hemodynamic response of prefrontal or other brain regions may not be restricted to WM tasks, and that a shift in a nonlinear relationship between activation of a brain region and task parameters may not be specific to schizophrenia. Thus, in functional imaging studies, it appears important to recognize that emergent activation patterns may reflect specific task parameters, as opposed to differences solely in the putative cognitive functions being studied.

References

- Akbarian, S., Bunney, W. E., Potkin, S. G., Wigal, S. B., Hagman, J. O., Sandman, C. A., et al. (1993). Altered distribution of nicotinamide-adenine dinucleotide phosphate-diaphorase cells in frontal lobes of schizophrenics implies disturbances of cortical development. *Archives of General Psychiatry*, *50*, 169–177.
- Andreasen, N. C. (1984). *Scale for the Assessment of Negative Symptoms/Scale for the Assessment of Positive Symptoms* [Manual]. Iowa City: University of Iowa Press.
- Andreasen, N. C., Paradiso, S., & O'Leary, D. S. (1998). Cognitive dysmetria as an integrative theory of schizophrenia: A dysfunction in cortical-cerebellar circuitry? *Schizophrenia Bulletin*, *24*, 203–218.
- Barch, D. M., Carter, C. S., Braver, T. S., Sabb, F. W., MacDonald, A., III, Noll, D. C., et al. (2001). Selective deficits in prefrontal cortex function in medication-naïve patients with schizophrenia. *Archives of General Psychiatry*, *58*, 280–288.
- Barch, D. M., Csernansky, J. G., Conturo, T., & Snyder, A. (2002). Working and long-term memory deficits in schizophrenia: Is there a common prefrontal mechanism? *Journal of Abnormal Psychology*, *111*, 478–494.
- Bigelow, L. B., & Berthot, B. D. (1989). The Psychiatric Symptom Assessment Scale. *Psychopharmacology Bulletin*, *25*, 168–173.
- Borenstein, M., & Rothstein, H. (1999). *Comprehensive meta-analysis* [Computer software]. Englewood, NJ: Biostat.
- 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.
- Brett, M. (2002). *The MNI brain and the Talairach atlas*. Retrieved July 11, 2006, from Cambridge University, Cognition and Brain Sciences Web site: <http://www.mrc-cbu.cam.ac.uk/Imaging/Common/mninspace.shtml>
- Brett, M., Johnsrude, I. S., & Owen, A. M. (2002). The problem of functional localization in the human brain. *Nature Reviews: Neuroscience*, *3*, 243–249.
- Cairo, T. A. (2003). *Delineating the effects of load on temporally dissociable working memory components in schizophrenia*. Unpublished master's thesis, University of British Columbia, Vancouver, Canada.
- Cairo, T. A., Liddle, P. F., Woodward, T. S., & Ngan, E. T. C. (2004). The influence of working memory load on phase specific patterns of cortical activity. *Cognitive Brain Research*, *21*, 377–387.
- Callicott, J. H. (2003). An expanded role for functional neuroimaging in schizophrenia. *Current Opinion in Neurobiology*, *13*, 256–260.
- Callicott, J. H., Bertolino, A., Mattay, V. S., Langheim, F. J., Duan, 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.
- Callicott, J. H., Mattay, V. S., Verchinski, B. A., Marenco, S., Egan, M. F., & Weinberger, D. R. (2003). Complexity of prefrontal cortical dysfunction in schizophrenia: More than up or down. *American Journal of Psychiatry*, *160*, 2209–2215.
- Callicott, J. H., Ramsey, N. F., Tallent, K., Bertolino, A., Knable, M. B., Coppola, R., et al. (1998). Functional magnetic resonance imaging brain

- mapping in psychiatry: Methodological issues illustrated in a study of working memory in schizophrenia. *Neuropsychopharmacology*, 18, 186–196.
- Carter, C. S., Perlstein, W., Ganguli, R., Brar, J., Mintun, M., & Cohen, J. D. (1998). Functional hypofrontality and working memory dysfunction in schizophrenia. *American Journal of Psychiatry*, 155, 1285–1287.
- D'Esposito, M. (2001). Functional neuroimaging of working memory. In R. Cabeza & A. Kingstone (Eds.), *Handbook of functional neuroimaging of cognition* (pp. 293–327). Cambridge, MA: MIT Press.
- Glahn, D. C. (2000). Behavioral and neural correlates of spatial working memory dysfunction in schizophrenia. *Dissertation Abstracts International: B. The Physical Sciences and Engineering*, 61, 3275.
- Glahn, D. C., Ragland, J. D., Abramaoff, A., Barrett, J., Laird, A. R., Bearden, C. E., et al. (2005). Beyond hypofrontality: A quantitative meta-analysis of functional neuroimaging studies of working memory in schizophrenia. *Human Brain Mapping*, 25, 60–69.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
- Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods*, 3, 486–504.
- Honey, G. D., Bullmore, E. T., & Sharma, T. (2002). De-coupling of cognitive performance and cerebral functional response during working memory in schizophrenia. *Schizophrenia Research*, 53, 45–56.
- Honey, G. D., Sharma, T., Suckling, J., Giampietro, V., Soni, W., Williams, S. C., et al. (2003). The functional neuroanatomy of schizophrenic subsyndromes. *Psychological Medicine*, 33, 1007–1018.
- Hugdahl, K., Rund, B. R., Lund, A., Asbjørnsen, A., Egeland, J., Erslund, L., et al. (2004). Brain activation measured with fMRI during a mental arithmetic task in schizophrenia and major depression. *American Journal of Psychiatry*, 161, 286–293.
- Jansma, J. M., Ramsey, N. F., van der Wee, N. J., & Kahn, R. S. (2004). Working memory capacity in schizophrenia: A parametric fMRI study. *Schizophrenia Research*, 68, 159–171.
- 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.
- Kay, S. R., Fiszbein, A., & Olper, L. A. (1987). The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. *Schizophrenia Bulletin*, 13, 261–276.
- Kim, J. J., Kwon, J. S., Park, H. J., Youn, T., Kang, d. H., Kim, M. S., et al. (2003). Functional disconnection between the prefrontal and parietal cortices during working memory processing in schizophrenia: A [¹⁵O]H₂O PET study. *American Journal of Psychiatry*, 160, 919–923.
- Kindermann, S. S., Brown, G. G., Zorrilla, L. E., Olsen, R. K., & Jeste, D. V. (2004). Spatial working memory among middle-aged and older patients with schizophrenia and volunteers using fMRI. *Schizophrenia Research*, 68, 203–216.
- Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, L. R., Kochunov, P. V., et al. (2000). Automated Talairach atlas labels for functional brain mapping. *Human Brain Mapping*, 10, 120–131.
- Liddle, P. F., Ngan, E. T. C., Duffield, G., Kho, K., & Warren, A. J. (2002). Signs and Symptoms of Psychotic Illness (SSPI): A rating scale. *British Journal of Psychiatry*, 180, 45–50.
- MacDonald, A. W., III, & Carter, C. S. (2003). Event-related fMRI study of context processing in dorsolateral prefrontal cortex of patients with schizophrenia. *Journal of Abnormal Psychology*, 112, 689–697.
- Maldjian, J. A., Laurienti, P. J., & Burdette, J. H. (2004). Precentral gyrus discrepancy in electronic versions of the Talairach atlas. *NeuroImage*, 21, 450–455.
- Manoach, D. S. (2002). Functional neuroimaging investigations of working memory deficits in schizophrenia: Reconciling discrepant findings. In M. F. Lenzenweger & J. M. Hooley (Eds.), *Principles of experimental psychopathology: Essays in honor of Brendan A. Maher* (pp. 119–134). Washington, DC: American Psychological Association.
- Manoach, D. S. (2003). Prefrontal cortex dysfunction during working memory performance in schizophrenia: Reconciling discrepant findings. *Schizophrenia Research*, 60, 285–295.
- Manoach, D. S., Gollub, R. L., Benson, E. S., Searl, M. M., Goff, D. C., Halpern, E., et al. (2000). Schizophrenic subjects show aberrant fMRI activation of dorsolateral prefrontal cortex and basal ganglia during working memory performance. *Biological Psychiatry*, 48, 99–109.
- Manoach, D. S., Halpern, E. F., Kramer, T. S., Chang, Y., Goff, D. C., Rauch, S. L., et al. (2001). Test–retest reliability of a functional MRI working memory paradigm in normal and schizophrenic subjects. *American Journal of Psychiatry*, 158, 955–958.
- Manoach, D. S., Press, D. Z., Thangaraj, V., Searl, M. M., Goff, D. C., Halpern, E., et al. (1999). Schizophrenic subjects activate dorsolateral prefrontal cortex during a working memory task, as measured by fMRI. *Biological Psychiatry*, 45, 1128–1137.
- Manoach, D. S., Schlaug, G., Siewert, B., Darby, D. G., Bly, B. M., Benfield, A., et al. (1997). Prefrontal cortex fMRI signal changes are correlated with working memory load. *NeuroReport*, 8, 545–549.
- Mendrek, A. (2000). An fMRI investigation of trait and state markers in schizophrenia. *Dissertation Abstracts International: B. The Physical Sciences and Engineering*, 62, 3505.
- Menon, V., Anagnoson, R. T., Mathalon, D. H., Glover, G. H., & Pfefferbaum, A. (2001). Functional neuroanatomy of auditory working memory in schizophrenia: Relation to positive and negative symptoms. *NeuroImage*, 13, 433–446.
- Meyer-Lindenberg, A. S., Olsen, R. K., Kohn, P. D., Brown, T., Egan, M. F., Wienberger, D. F., et al. (2005). Regionally specific disturbance of dorsolateral prefrontal-hippocampal functional connectivity in schizophrenia. *Archives of General Psychiatry*, 62, 379–386.
- Meyer-Lindenberg, A., Poline, J.-B., Kohn, P. D., Holt, J. L., Egan, M. F., Weinberger, D. R., et al. (2001). Evidence for abnormal cortical functional connectivity during working memory in schizophrenia. *American Journal of Psychiatry*, 158, 1809–1817.
- Ojeda, N., Ortuño, F., Arbizu, J., López, P., Martí-Climent, J. M., Peñuelas, I., et al. (2002). Functional neuroanatomy of sustained attention in schizophrenia: Contribution of parietal cortices. *Human Brain Mapping*, 17, 116–130.
- Overall, J. E., & Gorham, D. R. (1962). The Brief Psychiatric Rating Scale. *Psychological Reports*, 10, 799–812.
- Perlstein, W. M., Dixit, N. K., Carter, C. S., Noll, D. C., & Cohen, J. D. (2003). Prefrontal cortex dysfunction mediates deficits in working memory and prepotent responding in schizophrenia. *Biological Psychiatry*, 53, 25–38.
- Quintana, J., Wong, T., Ortiz-Portillo, E., Kovalik, E., Davidson, T., Marder, S. R., & Mazzotta, J. C. (2003). Prefrontal-posterior parietal networks in schizophrenia: Primary dysfunctions and secondary compensations. *Biological Psychiatry*, 53, 12–24.
- Rajkowska, G., & Goldman-Rakic, P. S. (1995). Cytoarchitectonic definition of prefrontal areas in the normal human cortex: II. Variability in locations of areas 9 and 46 and relationship to the Talairach coordinate system. *Cerebral Cortex*, 5, 323–337.
- Rypma, B. (2006). Factors controlling neural activity during delayed-response task performance: Testing a memory organization hypothesis of prefrontal function. *Neuroscience*, 139, 223–235.
- Rypma, B., Berger, J. S., & D'Esposito, M. (2002). The influence of working-memory demand and subject performance on prefrontal cortical activity. *Journal of Cognitive Neuroscience*, 14, 721–731.
- Rypma, B., & D'Esposito, M. (2003). A subsequent-memory effect in dorsolateral prefrontal cortex. *Cognitive Brain Research*, 16, 162–166.
- Rypma, B., Prabhakaran, V., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1999). Load-dependent roles of frontal brain regions in the maintenance of working memory. *NeuroImage*, 9, 216–226.

- Sabri, O., Owega, A., Schreckenberger, M., Sturz, L., Fimm, B., Kunert, P., et al. (2003). A truly simultaneous combination of functional transcranial Doppler sonography and H₂¹⁵O PET adds fundamental new information on differences in cognitive activation between schizophrenics and healthy control subjects. *Journal of Nuclear Medicine*, *44*, 671–681.
- Salgado-Pineda, P., Junqué, C., Vendrell, P., Baeza, I., Bargalló, N., Falcón, C., et al. (2004). Decreased cerebral activation during CPT performance: Structural and functional deficits in schizophrenic patients. *NeuroImage*, *21*, 840–847.
- Schlösser, R., Gesierich, T., Kaufmann, B., Vucurevic, G., Hunsche, S., Gawehn, J., et al. (2003). Altered effective connectivity during working memory performance in schizophrenia: A study with fMRI and structural equation modeling. *NeuroImage*, *19*, 751–763.
- Selemon, L. D., Rajkowska, G., & Goldman-Rakic, P. S. (1995). Abnormally high neuronal density in the schizophrenic cortex: A morphometric analysis of prefrontal area 9 and occipital area 17. *Archives of General Psychiatry*, *52*, 805–818.
- Smith, E. E., & Jonides, J. (1999, March 12). Storage and executive processes in the frontal lobes. *Science*, *283*, 1657–1661.
- Stern, Y. (2002). What is cognitive reserve? Theory and research application of the reserve concept. *Journal of the International Neuropsychological Society*, *8*, 448–460.
- Stevens, A. A., Goldman-Rakic, P. S., Gore, J. C., Fulbright, R. K., & Wexler, B. E. (1998). Cortical dysfunction in schizophrenia during auditory word and tone working memory demonstrated by functional magnetic resonance imaging. *Archives of General Psychiatry*, *55*, 1097–1103.
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain: An approach to medical cerebral imaging*. New York: Thieme Medical Publishers.
- Volz, H. P., Gaser, C., Haeger, F., Rzanny, R., Poenisch, J., Mentzel, H. J., et al. (1999). Decreased frontal activation in schizophrenics during stimulation with the Continuous Performance Test: A functional magnetic resonance imaging study. *European Psychiatry*, *14*, 17–24.
- Walter, H., Wunderlich, A. P., Blankenhorn, M., Schaefer, S., Tomczak, R., Spitzer, M., et al. (2003). No hypofrontality, but absence of prefrontal lateralization comparing verbal and spatial working memory in schizophrenia. *Schizophrenia Research*, *61*, 175–184.
- Wykes, T., Brammer, M., Mellers, J., Bray, P., Reeder, C., Williams, C., et al. (2002). Effects on the brain of a psychological treatment: Cognitive remediation therapy: Functional magnetic resonance imaging in schizophrenia. *British Journal of Psychiatry*, *181*, 144–152.

Appendix A

Literature Search Terms

Asterisks indicate truncated terms.

Schizophrenia and functional neuroimag*

Schizophrenia and functional imag*

Schizophrenia and fMR*

Schizophrenia and functional MR*

Schizophrenia and functional magnetic

Schizophrenia and PET

Schizophrenia and positron

Schizophrenia and brain activation

Schizophrenia and cognitive activation

Schizophrenia and prefrontal activation

Schizophrenia and frontal activation

Appendix B

$$d = \frac{\bar{X}_c - \bar{X}_s}{S_p} = \frac{t}{\sqrt{\frac{n_c n_s}{n_c + n_s}}} = \frac{\sqrt{F}}{\sqrt{\frac{n_c n_s}{n_c + n_s}}}$$

and

$$S_p = \sqrt{\frac{S_c^2(n_c - 1) + S_s^2(n_s - 1)}{n_c + n_s - 2}},$$

where subscripts *c* and *s* refers to control and patient samples, respectively; \bar{X} = sample mean; *s* = standard deviation; *n* = sample size; *t*, *F* = inferential sample statistics; *F* refers only to *F* statistics with 1 numerator

degree of freedom. Statistical parametric mapping *Z* scores of the maximum activation difference between patients and controls were converted to *p* values.

Estimates from *p* values were made by obtaining the *t* statistic corresponding to the reported *p* on a *t* distribution with $n_c + n_s - 2$ degrees of freedom. Estimates from published figures were made from high-resolution (1,280 × 1,024 screen resolution) screenshots of figures in PDF versions of articles. A 1-pixel² grid was overlaid, and a straight line was used to measure the precise height of means and standard errors in the figure.

Received August 30, 2005

Revision received March 27, 2006

Accepted April 10, 2006 ■