The neural bases of distracter-resistant working memory

Tor D. Wager[1]*
Rachel Insler
Edward E. Smith[2]

1 Department of Psychology and Neuroscience, University of Colorado, Boulder
2 Department of Psychology, Columbia University

* Corresponding author information:
Tor D. Wager
Associate Professor of Psychology and Neuroscience
University of Colorado, Boulder
345 UCB, Boulder, CO 80309
tor.wager@colorado.edu
303-492-7487

Abstract

A major difference between humans and other animals is our capacity to maintain information in working memory (WM) while performing secondary tasks, which enables sustained, complex cognition. A common assumption is that the lateral prefrontal cortex (PFC) is critical for WM performance in the presence of distracters, but direct evidence is scarce. We assessed the relationship between fMRI activity and WM performance within-subjects, with performance matched across Distracter and No-distracter conditions. Activity in ventrolateral PFC during WM encoding and maintenance positively predicted performance in both conditions, whereas activity in the pre-supplementary motor area (pre-SMA) predicted performance only under distraction. Other parts of dorsolateral and ventrolateral PFC predicted performance only in the No-distracter condition. These findings challenge a lateral PFC-centered view of distracter-resistance, and suggest that the lateral PFC supports a type of WM representation that is efficient for dealing with task-irrelevant input but is nonetheless easily disrupted by dual-task demands.

Acknowledgements

We would like to thank Matthew Davidson for assistance with data analysis and manuscript preparation. This work was supported in part by grants NSF 0631637 and R01MH076136 (T.D.W).

Working memory (WM) refers to a system that maintains currently relevant goals and information for use in guiding ongoing information processing. Information in WM is thought to guide the deployment of attention (*1*) and the manipulation of information to achieve current task goals (*2, 3*). Probably no factor has more effect on WM performance than the presence of distracters (e.g., *4*). Consider the simple task of remembering a group of three letters—a trigram—over an interval of a few seconds. In a classic article (*5*), subjects rehearsed a trigram for intervals up to 18 sec while counting backwards by threes. With this distracting secondary counting task, what would otherwise have been trivial to remember became virtually impossible, with near-zero recall by 18 sec. The powerful effect of distraction is evident in everyday situations: Memory for the contents of a book or a phone number just memorized may be obliterated by a few moments of idle conversation.

Distracters in WM are of two basic types. They may not require any intentional processing, in which case they are simply perceptual experiences to be ignored while remembering other items, or they may constitute a secondary task that must be performed during maintenance of the memory set. Successful inhibition of merely perceptual distracters may rely on perceptual filtering mechanisms that can operate at an early processing stage (*6, 7*), whereas dual-task interference is likely to rely more heavily on task switching, coordination, and information-selection processes. Performance on tasks involving these two distracter types is essentially uncorrelated (*8*). Crucially, performance on WM tasks with dual-task distracters, rather than perceptual distractors, predicts performance on complex tests of fluid intelligence (*9-14*), correlates with other indices of 'executive function' (*14, 15*), and predicts everyday cognitive failures (*16*).

Though there have been many studies investigating the brain mechanisms of selective attention and perceptual distracter suppression, very few studies have investigated those involved in creating WM representations resistant to dual -task distracters. The prefrontal cortex (PFC) is thought to be crucial for distracter resistance (*17-19*), but much of the key evidence has come from studies of perceptual distracters (*20-28*). In an important study, Sakai et al. (*29*) studied the basis of resistance to dual-task distraction by studying the relationship between PFC activity during WM encoding and maintenance with recall accuracy after performing a distracting secondary task.

Activity during encoding/maintenance in right Brodmann's Area (BA) 46, which spans dorsolateral and ventrolateral prefrontal cortices (DLPFC/VLPFC), was greater for correct than error trials. On the basis of this finding, Saki et al. suggested that this area is critical for distracter-resistant memory.

However, the Sakai et al. study was limited in a way that substantially undermines its conclusion. Though they showed an accuracy effect in the presence of a distracting task, they did not demonstrate that the accuracy effect was *specific* to the distraction condition. They did include a no-distraction condition, but participants made virtually no errors without distraction. Thus, the alternative remains that the lateral PFC plays a general role in WM maintenance with or without distraction. In addition, the accuracy probe on Sakai et al.'s WM task involved a binary "yes or no" recognition decision, precluding the possibility of assessing parametric relationships between brain activity and accuracy, which would help establish that there is a meaningful relationship between PFC activity and accuracy across the normal range of performance.

Two plausible alternatives for the role of the lateral PFC provide the basis for hypotheses in the current study. First, the lateral PFC may be important for WM maintenance with or without distractors. In several studies (*30, 31*), patients with lesions of Brodmann's Area 46 were impaired on delay tasks with or without distraction. Similarly, monkeys with prefrontal lesions have shown impaired performance on working memory tasks with and without distracters (*32-34*), though it has been argued that monkeys are distractible even without explicit distracters included in the task. Secondly, the lateral PFC may be particularly important for filtering out and reducing interference from perceptual distractors—as evidenced by neuroimaging studies (*1, 25, 27, 28, 35*), humans and monkeys with prefrontal damage (*20-22*), and primate electrophysiology (*23, 24, 36*)—but this region may not be crucial for resistance to dual-task distraction. This idea is consistent with evidence for a broad role for the VLPFC in particular in the selection of task-relevant sensory stimuli (*37-40*). Thus, resisting perceptual distracters may load heavily on stimulus selection, whereas resisting dual-task distraction may load heavily on coordination and scheduling of WM processes, and thus place demands on different cortical areas.

We studied resistance to dual-task distraction using a standard verbal WM task with several novel features (Figure 1). First, like Sakai et al., we included both distracter and no-distracter conditions. In the distracter condition, there was a delay before onset of the secondary distracter task so that we could assess brain activity *before* distraction and relate it to subsequent accuracy. Second, to provide a parametric assessment of accuracy on each trial, we included a series of four yes/no recognition probes after each trial, resulting in five levels of accuracy (0-4 correct). Finally, to permit a sensitive analysis of the brain-accuracy relationship for *both* distracter and no-distracter conditions, we used an adaptive staircase procedure to select a memory set size for each subject and each condition that resulted in approximately 75% accuracy (3/4 probes correct on average) in both conditions. This allowed us to compare brain-accuracy relationships for distracter vs. no-distracter conditions, testing whether the lateral PFC and/or other areas show monotonic increases in activity with performance only under dual-task distracter conditions.

Results

Encoding/Maintenance **vs. Baseline**. Standard working-memory related regions were strongly activated during the Encoding/Maintenance period (prior to distracter presentation; $p < .01$ FDR). As shown in Figure 2A (hot colors, yellow/orange/red), activated regions included bilateral middle and inferior frontal gyri, intraparietal sulcus, inferior temporal cortices, and pre-supplementary motor cortices (pre-SMA) and anterior cingulate (ACC). These results correspond closely with those from a recent meta-analysis of working memory and related executive tasks (*41*); see Figure 2B). De-activation, shown in purple/blue in Figure 2A, was found in "Default mode" regions associated with internal monitoring and task-unrelated thought during rest (*42, 43*), including anterior medial prefrontal cortex (aMPFC), posterior cingulate (PCC) and precuneus, inferior parietal cortex (IPC), and superior temporal sulcus (STS). In addition, de-activations were found in areas associated with sensory visual and somatosensory processing, including V1/V2, extrastriate cortex, and dorsal posterior insula (dpINS) and SII. The combined activated and deactivated regions were used as a mask for subsequent

analyses of accuracy.

Predictors of subsequent accuracy. We searched within a subset of functionally defined [Encoding/Maintenance – Baseline] ROIs of *a priori* interest (see Methods for region definition details) for subsequent accuracy effects. We expected activated areas in lateral and dorsomedial PFC to show positive accuracy-performance relationships—i.e., greater activity should predict parametric increases in subsequent accuracy—and deactivated "default mode" and somatosensory regions [dpINS and SII]) to show negative accuracy-performance relationships.

We found monotonic relationships between brain activity increases and accuracy in three ROIs: the pars opercularis of the inferior frontal gyrus (IFG) bilaterally (Brodmann's Area 44, based on the SPM Anatomy Toolbox v. 1.6 (*44*)), and the pre-SMA; both are shown in yellow in Figure 3. Detailed plots of responses for the IFG and pre-SMA are shown in Figures 4 and 5, respectively. Further analysis in all areas that showed significant [Encoding/Maintenance – Baseline] responses revealed positive relationships with accuracy in other standard working memory regions (Figure 3; see Table 1 for statistics), including the premotor cortex, right anterior insula, inferior parietal lobule, left premotor cortex, and left inferior occipital/temporal cortex. These results are consistent with the hypothesis that increases in multiple regions with a distributed network support robust working memory encoding and maintenance (cf. *45*).

Deactivation in several "default mode" and somatosensory ROIs also predicted subsequent WM accuracy (Table 1; negative relationships, shown in blue in Figure 3), including the pregenual cingulate/aMPFC, right STS, hippocampus, and bilateral IPC. All of these areas were de-activated during Encoding/Maintenance overall, and greater de-activation predicted increasing subsequent WM accuracy. In addition, right SII showed the same pattern of stronger de-activation with increasing accuracy. In the broader search within all [Encoding/Maintenance – Baseline] areas, negative relationships with accuracy were also found in lateral occipital cortex and several motor/premotor regions. These findings are broadly consistent with the idea that accurate WM encoding and maintenance relies on reduced activation in brain systems that encode task-unrelated perceptual and motor processes (e.g., *42*).

Predictors of distracter-resistant memory. Regions in which activation during Encoding/Maintenance supports distracter-resistant memory should (1) be activated during Encoding/Maintenance, (2) *positively* predict accuracy overall, and (3) show a *positive* interaction with Distraction, indicating a stronger accuracy effect on D trials. Thus, to assess regions that differentially predicted accuracy under Distracter and No-distracter conditions, we tested the Distraction x Accuracy interaction in the average signal within accuracy-predictive regions discussed above (i.e., regions listed in Table 1).

Only one region, pre-SMA, fit all of the criteria (Figure 4). Positive Accuracy and Accuracy x Distracter effects indicated that activity was more positively related to subsequent memory for Distracter trials than No-distracter trials ($p < 0.05$; see Table 1). Futhermore, post hoc tests revealed that the accuracy effect was significant for Distracter trials ($t(16) = 2.86$, $p = 0.01$ two-tailed) but not for No-distracter trials ($t(16) = -.33$, $p = 0.74$), suggesting that pre-SMA activity is important for establishing distracter-resistant memory. In contrast, the bilateral IFG regions that predicted accuracy were equally predictive for Distracter and No-distracter trials, as shown in Figure 5 (see Table 1 for statistics).

It is possible that distracter-resistant memories are encoded in regions of the prefrontal cortex other than our stringently defined *a priori* ROIs, which required significant [Encoding/Maintenance – Baseline] activity and positive accuracy prediction across all trials. To test this possibility, we searched for voxels that showed Accuracy x Distracter effects at a threshold of $p < .001$ across the frontal cortex. The results revealed both positive and negative Accuracy x Distracter interactions. Only one region of right middle frontal gyrus (MFG; BA 9/46; orange in Figure 6) showed a positive relationship with performance only in the Distracter condition, as predicted by the distracter-resistance account of lateral prefrontal function. It showed [Encoding/Maintenance – Baseline] activation overall, but at a lower threshold ($t(16) = 3.04$, $p = .008$), and was not predictive of accuracy across all trial types. Other prefrontal regions, including bilateral MFG (BA 45) and premotor cortex (BA 6) and left ventrolateral PFC (BA 47/12), showed a negative interaction (blue in Figure 6). Examining the pattern of these interactions revealed that in right premotor and VLPFC, the interaction was driven

mainly by an association between greater activity and lower levels of accuracy on Distracter trials, contrary to the distracter-resistance account of lateral PFC function. In MFG (BA 45), the interaction was driven mainly by stronger positive brain-accuracy relationships on No-distracter trials. These findings are inconsistent with the idea that lateral PFC facilitates distracter-resistant working memory.

In addition to activated areas in which the magnitude of increases predicts subsequent memory, we expected that the magnitude of decreases in de-activated areas may predict subsequent memory specifically on Distracter trials. Results for "default mode" regions of interest are shown in Figure 7. Among areas in which de-activation magnitude predicted accuracy across all trials (Figure 3 and Figure 7A, left), three regions showed interactions with Distracter status (Figure 7A, right): Right SII and dpINS, associated with somatic sensory perception, and STS, broadly associated with spontaneous cognition and social cognition (see Discussion). However, the Distracter *x* Accuracy interaction in each of these areas was positive, indicating that decreases in activity were most beneficial for accurate performance in the No-distracter condition. For example, as shown in Figure 7B, low accuracy in the No-distracter condition in particular was associated with increased SII activity. These results underscore the idea that different processes may be rate-limiting steps in performance with and without distraction.

In contrast, anterior medial prefrontal cortex (aMPFC), a part of the "default mode" network, showed a trend toward stronger de-activation with more accurate memory specifically on Distracter trials. Though this effect did not reach significance for the Accuracy *x* Distracter interaction in the ROI defined by overall subsequent accuracy effects (Table 1), an adjacent region just outside the accuracy-predictive search mask did show an Accuracy x Distracter effect at $p < .001$ (xyz = [6, 58, 8], 2 voxels, max t = 9.50; shown in Figure 7C). Thus, de-activation of aMPFC may also be important for creating distracter-resistant memory.

Discussion

In the present study, we assessed the role of PFC and other regions in creating memories resistant to dual-task distracters by using a WM paradigm with two novel elements. First,

we matched performance on Distracter and No-distracter versions of the task by adaptively titrating the memory set size, ensuring that we could efficiently estimate brain activity-performance relationships for both tasks. Second, we used a multi-part recognition probe with five levels of accuracy per trial, enhancing our ability to assess activity-performance relationships within-subjects. This approach is relatively unique, because in contrast to the hundreds of published WM studies, only a handful have reported relationships between PFC activity and behavioral performance (*25, 45-49*).

We replicated earlier results of Sakai et al. by showing that activation during encoding/maintenance in both the lateral and medial PFC predicted successful subsequent performance under distraction. However, among rigorously defined functional ROIs, only activation in the pre-SMA predicted performance *specifically* under distraction. Activity in VLPFC predicted subsequent performance in *both* Distracter and No-distracter conditions of the task, and was not selective to the distracter task. An exploratory search outside of the strict boundaries of our functional localizers also revealed an area in the right DLPFC (BA 46) close to Sakai et al.'s (*29*) results in which activation showed a more positive relationship with subsequent performance under Distracter than No-distracter conditions. However, activation of nearby and more widespread areas within bilateral IFG and DLPFC (bilateral BA 45 and left BA 47) predicted accurate performance only in the No-distracter task, in opposition to the distracter-resistance hypothesis of lateral PFC function.

The present results suggest that pre-SMA, which has been activated in numerous WM tasks but has received less attention in studies of WM than lateral PFC, plays a more important role than previously thought in maintaining representations that are resistant to dual-task interference, which is particularly important for complex cognition. Pre-SMA was a) activated during WM encoding/maintenance; b) correlated positively with subsequent WM accuracy across all trials on average; and c) correlated more positively for Distracter than No-distracter conditions. Furthermore, the activation-performance correlation appeared to be exclusively driven by the Distracter condition. This finding is interesting because there is substantial evidence that pre-SMA plays a very general role in goal-directed cognition and action-selection across many tasks (*41, 50-55*). Thus, it may be the strength of engagement of basic representation-selection and rehearsal

processes that create distracter-resistant WM, rather than a circuit dedicated specifically to managing distraction.

Our results also have implications for the role of lateral PFC in WM, and suggest the somewhat surprising conclusion that in contrast to the dominant view (*10, 19*), lateral PFC may play a limited role in managing dual-task situations in healthy individuals. The fact that multiple areas in VLPFC and DLPFC, along with temporal regions, predicted performance more strongly in the No-distractor condition implies that memory-set representations in prefrontal-temporal circuits are fragile and are disrupted by distracters. Performance correlations identify areas that implement 'performance-limiting' processes, or processes that fail intermittently and thus influence accuracy. VLPFC and DLPFC are almost certainly involved in task representation and maintenance, but the critical determinant of accuracy under dual-task distraction may be another facet of the representation—e.g., semantic or multi-modal associations that allow the memory to persist. The creation of a more fragile memory representation in the No-distractor condition is also consistent with the idea that activation of task-irrelevant representations (e.g., presumably somatosensory-related activity in S2) might be more disruptive in the No-distractor condition, which was the case here. Whether VLPFC supports an efficient but fragile type of encoding remains to be investigated in future experiments.

This conception of lateral PFC function is not inconsistent with previous work on distraction in WM, provided that the distinction between perceptual distracters and dual-task distraction is acknowledged. For instance, Chao and Knight (*21, 56*) found that patients with lateral PFC damage showed impaired WM performance with perceptual distracters (auditory tones) and increased electrophysiological responses to such tones, suggesting reduced distracter inhibition. In monkeys, Miller and Desimone (*23*) reported that activity in prefrontal neurons persisted across a distracter-filled delay period, but activity in inferior temporal neurons was disrupted by distracter presentation. In humans, Jha et al. (2004) reported greater LPFC activity during WM maintenance with concurrent presentation of perceptual distracters that were similar to the memory items. These findings and others support the idea that LPFC is important for stimulus selection and reduction of proactive interference (*37, 40, 57-60*), i.e., keeping irrelevant representations out of mind, but not management of concurrent task sets required for resistance to dual-

task distraction. As very few studies have examined the neural bases of dual-task distraction (for discussion see *61*), and management dual-task distraction is particularly predictive of complex cognitive abilities (*8, 62*), the present results serve as a launching point for explorations of the brain systems most critical for enabling complex cognition.

Another set of findings that argues against a lateral-PFC-dominated view of WM maintenance is the robust relationship between WM performance and deactivation of "default mode" and sensory processing regions. "Default mode" regions are associated with mentalization about the self (*63*), autobiographical memory retrieval (*64*), affective arousal and emotion (*65*), and task-irrelevant thoughts (*42, 66*)(*67, 68*). These findings parallel similar relationships between "default mode" decreases and performance on basic attention tasks (*69, 70*) and long-term memory encoding (*71*).

Thus, though lateral PFC increases may be a robust feature of WM processes, and an intact PFC may be critical for complex memory operations, the lateral PFC may be "necessary but not sufficient" for good performance in the context of dual-task distraction. In contrast, activation in Pre-SMA (likely related to response-generation) and deactivation of regions involved in spontaneous, internally generated cognition may be critical.

Limitations and methodological considerations

This study was designed to assess brain-performance relationships for distracting vs. non-distracting WM conditions, and so required that performance be matched in the two conditions, and that the task be difficult enough to be in a sensitive range for brain-performance correlations for each subject. To do this, we opted to titrate the memory set size for each subject for each task, resulting in a higher memory-set size for the ND than the Distracter condition. Because of this, we did not attempt to compare the main effects of the Distracter vs. ND conditions on brain activity. Previous studies (e.g., (*25-28*)) have supplied this information, reporting increased activity in PFC during processing of task-relevant distracters.

Another design choice was to block the Distracter and No-distracter conditions, thus preventing confusion and avoiding task switching on the part of subjects. Because subjects knew whether the distracters would appear or not, they could easily adopt different strategies in Distracter vs. No-distracter conditions. Thus, we cannot tell

whether it is the type of representation or the type of strategy employed that causes different regions to predict performance in each case. However, we regard this as an ecologically valid feature of the task: Knowing that one will be distracted affords the opportunity to engage in robust encoding and maintenance processes, strategic or otherwise. Such knowledge is a feature of complex WM span tasks that predict general fluid intelligence. Future studies could specifically investigate the impact of knowledge about upcoming distraction.

Third, it might be argued that the high WM load in our task produced "distraction" in the Distracter and No-distracter conditions alike. However, the double dissociations in the brain-performance correlations in Distracter and No-distracter conditions suggest that these two conditions are not equally distracting, or at least not in the same ways. As we argued above, interference due to memory load should be thought of as conceptually distinct from the requirement to manage a demanding secondary task.

Fourth, we did not attempt to separate encoding and maintenance processes, because doing so would have come at a substantial cost in power for our main fMRI comparisons. Future studies might fruitfully separate these processes.

Finally, the task we used was designed to study the mechanisms of WM maintenance, but is not isomorphic with the tasks and task parameters shown to predict fluid intelligence. Studying such tasks with fMRI (*61*) and correlating fMRI activity with fluid intelligence directly (*49*) is an important future direction.

Methods

Participants

Seventeen healthy right-handed participants (*M* age = 22.3 years, 6 female) were recruited in compliance with the human subjects regulations of Columbia University and provided informed consent in accordance with the Declaration of Helsinki. They were paid US$20/hour for voluntary participation, plus bonuses as described below. Handedness was assessed with the Edinburgh Handedness Inventory, and eligibility was assessed with a general health questionnaire and fMRI safety screening form.

Materials and procedures

Stimuli. Stimuli were selected from a list of two-syllable words ranging from four to six letters in length (mean 5.5), generated using the online version of the MRC Psycholinguistic Database (http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm). Fifty nouns and fifty verbs were then each selected randomly from that list for use in the experiment. The words had intermediate average levels of concreteness (mean = 335), imageability (mean = 357), and Kucera-Francis frequency (mean = 85). Words were randomized across task conditions for each participant, with random-with-replacement selection, under the constraint that words were not repeated during the same trial.

Experimental structure. The experiment consisted of 80 working memory trials. As shown in Figure 1, on each trial, participants intentionally encoded a series of words ("Memory Set"). The set size was chosen for each individual based on a pre-scan calibration (described below). During the Encoding phase, each word was presented on a computer screen for 500 ms (with an approximate visual angle of 1 degree). Participants then maintained the word set over a 6000 ms Maintenance period, during which a fixation cross was presented on the computer screen. How FMRI activity varied during these combined encoding and maintenance periods was of primary interest.

Following Maintenance, trials were divided into two types, which occurred with equal frequency. On Distraction trials (D), participants performed a secondary judgment task. They viewed a series of four words, each presented for 1200 ms, and indicated whether each distracter was a noun or a verb by pressing the index or middle finger of the right hand, respectively, on an MRI-compatible button box. No words were members of the Memory Set. Participants were instructed to maintain the original memory set during the secondary task. A 2000 ms fixation period followed the set of distracters. On No-distraction trials (ND), participants viewed a fixation cross for the same amount of total time as the distraction task (6800 msec) and continued to maintain the Memory Set.

D and ND trials were grouped in blocks of 10 trials, so that an entire scanning run consisted of either D or ND trials. Participants were told in advance that the trials would be blocked, and therefore had information that would allow them to prepare differently during Encoding/Maintenance for D vs. ND trials.

Following the continued maintenance with or without the distracter task, participants viewed a series of 4 unique Probe words, each presented for 1500 msec. Participants indicated whether each word was part of the Memory Set with a "Yes" (index finger) or "No" (middle finger) response on the button box. This multistage probe period allowed for five levels of accuracy (0-4 probes correct) on each trial, providing a more sensitive measure of WM performance than has previously been used. Following the probe period, a 20 sec inter-trial interval allowed fMRI responses to return to approximate baseline values, which permitted us to quantify activity on individual trials.

Performance incentives. Subjects were incentivized with instructions that they could receive a $0.10 monetary reward on every trial for successful performance, with the potential to earn an additional $8.00 over the course of the experiment. On ND trials "successful" performance for reinforcement purposes was defined solely in terms of memory performance, whereas on D trials also required 3 or 4 correct responses to the four Noun/Verb or Yes/No judgments. Also, on D trials, subjects received equivalent monetary incentives ($0.05 each) for good performance on both the Probe and the distracting Noun/Verb judgments to ensure that they would attend to both tasks. Participants did not receive any feedback about monetary earnings until the end of each block (scanning run). Bonuses ranged from $3.70 to $5.30, with an average of $4.61.

Memory Set Titration. D and ND trials were matched for difficulty within and across participants during a titration phase that occurred before fMRI scanning. This titration phase also served to familiarize the subjects with the 100 words used in the task to eliminate novelty effects. Subjects performed the task using Memory Sets of varying sizes until they were achieving successful performance on 80% of the trials. Across subjects and conditions, Memory Set size varied from 7 to 11 words. For each of the 17 subjects, the titrated Memory Set size was larger in the ND condition (mean = 10.77 words, s.d. = 0.56) than in the D condition (mean = 9.24 words, s.d. = 0.75), indicating that the distracting Noun/Verb judgment task was having a deleterious effect on working memory ability (mean difference = 1.53 words, $t(16) = 12.25$, $p < .001$). These set sizes were used in the subsequent fMRI task. Thus, the comparison in brain activity for [ND – D] trials differed in the number of words encoded, and might be expected to produce an

overall [ND > D] difference in standard working memory networks (though if encoding/maintenance processes were more strongly engaged during D blocks, it would tend to offset the set size effect by creating a [D > ND] difference). However, the [ND – D] comparison was not of primary interest; we designed the study to focus on relationships between brain activity and accuracy within each condition.. The titration ensured a homogenous distribution of accuracy values across subjects and a distribution across a range of accuracy values within subjects, for both the D and ND conditions, which is desirable for sensitive detection of brain-accuracy relationships.

Data acquisition

Whole-brain fMRI data were acquired on a 1.5T GE Signa Twin Speed Excite HD scanner (GE Medical Systems). Functional and anatomical images were acquired with a T2*-sensitive EPI BOLD sequence with a TR of 2000 ms, TE of 41 ms, flip angle of 60°, ascending interleaved acquisition, field of view of 22 cm, 24 slices and 3.44 x 3.44 x 4.5 mm voxels, with whole-brain coverage. Stimulus presentation and data acquisition were controlled using Matlab (http://www.mathworks.com/) and the Psychophysics Toolbox (http://psychtoolbox.org/). An LCD projector displayed stimuli on a back-projection screen mounted in the scanner suite. Responses were made with the right hand via a 5-finger button response unit with a molded hand brace (Resonance Technologies, Inc.).

Image processing and data analysis

Image denoising. Functional images were initially examined for spike artifacts using custom software (http://www.columbia.edu/cu/psychology/tor/). Global outlier time points were identified by computing both the mean and the standard deviation of values in each image for each slice. Mahalanobis distances for the matrix of mean values (one per slice) x functional volumes were computed, and images with a value above 3 standard deviations were considered outliers. Indicators for each individual outlier were entered as regressors in the general linear model (GLM). Next, principal components analysis was used to identify the first 10 components, and component scores were regressed on both outlier-related and task-related regressors (the design matrix; see

below). Components that had strong outlier-related correlations and weak task-related (e.g., r-squared values < .02) correlations were manually identified and removed from the data.

Preprocessing. After denoising, functional images were slice-acquisition timing and motion corrected using SPM5 software (Wellcome Department of Cognitive Neurology, UCL), using default parameters. Structural T1-weighted images were coregistered to the mean functional image for each subject using an iterative procedure of automated registration using mutual information coregistration in SPM5 and manual adjustment of the automated algorithm's starting point by a trained analyst until the automated procedure provided satisfactory alignment. Structural images were normalized (spatially warped) to a standard template brain (the MNI avg152T1.nii) SPM5's combined segmentation/normalization procedure with default options, and the warping parameters were applied to functional images for each subject. Normalized functional images were interpolated to 2 x 2 x 2 mm voxels and spatially smoothed with an 8-mm Gaussian filter.

First-level GLM model. First-level GLM analyses for each subject were estimated in SPM5. Quantification of single-trial response magnitudes was done by constructing a GLM design matrix with separate regressors for each trial, as in the "beta series" approach of Rissman et al. (*72*). The model is a three-level mixed effects model, with trial nested within condition the first level, condition (D vs. ND x accuracy) within subjects at the second level, and subject as a random effect at the third level. For each ~40 sec trial, we included regressors for 1) the Encoding/Maintenance period, an 8.5 - 10.5 s epoch, depending on the number of words encoded; 2) The Distraction/Maintenance period, a 6.8 s epoch; and the Probe period, a 6 s epoch. As each trial was modeled individually, separate regression estimates were obtained for each D and ND trial. To assess activity during the Encoding/Maintenance period *prior* to the onset of distracters, and without assuming a response shape, we modeled epochs as box-car functions (indicators) without convolving with a hemodynamic response function. All box-cars were delayed for 3 sec to compensate for the expected hemodynamic lag. An intercept regressor for each run modeled mean run activity.

The second level analysis was conducted on first-level regression parameter estimates (activation estimates) for the Encoding/Maintenance period for each trial. To minimize the effects of MR artifacts on statistical results, trials with images coded as outliers or whose global brain signal values were greater than 3 s.d. from the mean were excluded (mean = 1.44 outliers per 100 images, s.d. = 0.51). For each voxel within each subject, a generalized least squares model was fit with an AR(1) error structure. Model predictors were Distracter Type (D vs. ND), Accuracy (an integer from 0 to 4), and the Distracter Type x Accuracy interaction (all predictors were centered). The intercept term in this model corresponded to the overall [Encoding/Maintenance vs. Baseline] contrast, where baseline was the 20 sec inter-trial interval. These contrast images were subjected to group analysis (the third level of the model). A one-sample t-test across subjects (treating subject as a random effect) was performed for each of the four effects separately using robust regression, which increases statistical power and decreases false positive rates in the presence of outliers (Wager, Keller, Lacey, & Jonides, 2005).

Search volume and Localization of results. Normalized structural T1 images were averaged across participants to create an anatomical underlay for visualizing significant regions of activation, and for visually assessing normalization quality. In our experience, this is advantageous because the quality of nonlinear warping can vary across brain regions, resulting in greater differences between the standard brain and subjects' actual T1 images. The average T1 image was segmented using SPM5's unified segmentation/normalization routine, and the union of gray and white matter voxels was used as a mask in all analyses.

Region of interest (ROI) analyses and multiple comparisons correction. We defined ROIs based on the conjunction of three criteria: 1) gray/white matter, 2) functional response during the Encoding/Maintenance period, and 3) *a priori* interest based on previous studies. Our *a priori* hypotheses focused on two sets of regions. One set consisted of regions that were strongly activated in the overall [Encoding/Maintenance vs. Baseline] contrast ($q < .05$ false discovery rate (FDR) corrected, $p < .0095$, see Figure 2 and Supplementary Table 1) and were additionally located in lateral or medial prefrontal cortex, which have been strongly implicated in cognitive control and resistance to distracters. A second set consisted of "default mode"

regions (ventromedial PFC, posterior cingulate, STS) and non-visual sensory processing regions (S1/S2, A1/S2) that are increasingly thought to play a role in task-irrelevant thought (*42, 68, 73*), self-directed attention (*63*), and poorer cognitive performance (*69, 70*), and were also strongly deactivated in the [Encoding/Maintenance vs. Baseline] contrast ($q < .05$ FDR-corrected). Within ROIs, we report results at a threshold of $p < .001$, which is both the modal threshold used in fMRI studies and sufficient to control the family-wise error rate at $p < .05$ across voxels within each ROI. However, for completeness and archival purposes, Supplementary Figure 1 (see Supporting Online Materials) shows results for all regions that met criteria (1) and (2) above.

References

1. J. W. de Fockert, G. Rees, C. D. Frith, N. Lavie, *Science* **291**, 1803 (Mar 2, 2001).
2. A. Baddeley, *Working Memory*. (Oxford University Press, 1986., New York, 1986).
3. G. Hitch, A. Baddeley, *The Quarterly Journal of Experimental Psychology* **28**, 603 (1976).
4. G. Keppel, B. Underwood, *Journal of Verbal Learning and Verbal Behavior* **1**, 153 (1962).
5. L. R. Peterson, M. J. Peterson, *J Exp Psychol* **58**, 193 (Sep, 1959).
6. C. Bundesen, *Psychological review* **97**, 523 (1990).
7. A. Treisman, *Computer vision, graphics, and image processing* **31**, 156 (1985).
8. A. Miyake *et al.*, *Cognit Psychol* **41**, 49 (Aug, 2000).
9. R. Engle, M. Kane, S. Tuholski, *Models of working memory: Mechanisms of active maintenance and executive control*, 102 (1999).
10. M. Kane, R. Engle, *Psychonomic Bulletin & Review* **9**, 637 (2002).
11. A. Fry, S. Hale, *Psychological science* **7**, 237 (1996).
12. A. R. Conway, M. J. Kane, R. W. Engle, *Trends Cogn Sci* **7**, 547 (Dec, 2003).
13. A. Conway, N. Cowan, M. Bunting, D. Therriault, S. Minkoff, *Intelligence* **30**, 163 (2002).
14. M. J. Kane, R. W. Engle, *J Exp Psychol Gen* **132**, 47 (Mar, 2003).
15. M. K. Bleckley, F. T. Durso, J. M. Crutchfield, R. W. Engle, M. M. Khanna, *Psychon Bull Rev* **10**, 884 (Dec, 2003).
16. N. Friedman, A. Miyake, *Journal of Experimental Psychology-General* **133**, 101 (2004).
17. A. Baddeley, S. Della Sala, *Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences* **351**, 1397 (1996).
18. R. Knight, W. Richard Staines, D. Swick, L. Chao, *Acta Psychologica* **101**, 159 (1999).
19. E. K. Miller, J. D. Cohen, *Annu Rev Neurosci* **24**, 167 (2001).
20. R. Malmo, *Journal of Neurophysiology* **5**, 295 (1942).
21. L. L. Chao, R. T. Knight, *J Cogn Neurosci* **10**, 167 (Mar, 1998).
22. L. L. Chao, R. T. Knight, *Neuroreport* **6**, 1605 (Aug 21, 1995).
23. E. K. Miller, C. A. Erickson, R. Desimone, *J Neurosci* **16**, 5154 (1996).
24. Y. Tsushima, Y. Sasaki, T. Watanabe, *Science* **314**, 1786 (Dec 15, 2006).
25. A. P. Jha, S. A. Fabian, G. K. Aguirre, *Cogn Affect Behav Neurosci* **4**, 517 (Dec, 2004).
26. F. McNab, T. Klingberg, *Nat Neurosci* **11**, 103 (Jan, 2008).
27. B. R. Postle, *J Cogn Neurosci* **17**, 1679 (Nov, 2005).
28. K. K. Sreenivasan, A. P. Jha, *J Cogn Neurosci* **19**, 32 (Jan, 2007).
29. K. Sakai, J. B. Rowe, R. E. Passingham, *Nat Neurosci* **5**, 479 (May, 2002).
30. M. D'Esposito, B. R. Postle, *Neuropsychologia* **37**, 1303 (Oct, 1999).
31. A. Ptito, J. Crane, G. Leonard, R. Amsel, Z. Caramanos, *Neuroreport* **6**, 1781 (Sep 11, 1995).
32. S. Funahashi, M. Inoue, K. Kubota, *Behav Brain Res* **84**, 203 (Mar, 1997).

33. S. Funahashi, C. J. Bruce, P. S. Goldman-Rakic, *J Neurophysiol* **61**, 331 (Feb, 1989).
34. T. Sawaguchi, I. Yamane, *J Neurophysiol* **82**, 2070 (Nov, 1999).
35. F. Dolcos, B. Miller, P. Kragel, A. Jha, G. McCarthy, *Brain Res* **1152**, 171 (Jun 4, 2007).
36. D. Artchakov *et al.*, *Cereb Cortex* **19**, 2680 (Nov, 2009).
37. J. Fuster, *The prefrontal cortex*. (Academic Pr, 2008).
38. S. L. Thompson-Schill *et al.*, *Cogn Affect Behav Neurosci* **2**, 109 (Jun, 2002).
39. D. Badre, R. A. Poldrack, E. J. Pare-Blagoev, R. Z. Insler, A. D. Wagner, *Neuron* **47**, 907 (Sep 15, 2005).
40. D. E. Nee, J. Jonides, *Neuroimage* **45**, 963 (Apr 15, 2009).
41. J. X. van Snellenberg, T. D. Wager, in *Luria's Legacy in the 21st Century,* A. Christensen, E. Goldberg, D. Bougakov, Eds. (Oxford University Press, Oxford, 2009), pp. 30-61.
42. J. R. Andrews-Hanna, J. S. Reidler, C. Huang, R. L. Buckner, *J Neurophysiol* **104**, 322 (Jul, 2010).
43. D. A. Gusnard, M. E. Raichle, *Nat Rev Neurosci* **2**, 685 (Oct, 2001).
44. S. B. Eickhoff *et al.*, *Neuroimage* **25**, 1325 (May 1, 2005).
45. L. Pessoa, E. Gutierrez, P. Bandettini, L. Ungerleider, *Neuron* **35**, 975 (Aug 29, 2002).
46. B. Rypma, M. D'Esposito, *Brain Res Cogn Brain Res* **16**, 162 (Apr, 2003).
47. B. Rypma, J. S. Berger, M. D'Esposito, *J Cogn Neurosci* **14**, 721 (Jul 1, 2002).
48. E. E. Smith *et al.*, *Proc Natl Acad Sci U S A* **98**, 2095 (Feb 13, 2001).
49. J. R. Gray, C. F. Chabris, T. S. Braver, *Nat Neurosci* **6**, 316 (Mar, 2003).
50. R. E. Passingham, Y. C. Chen, D. Thaler, in *Neural Programming,* M. Ito, Ed. (S. Karger AG, Basel, 1989), pp. 13-24.
51. K. Shima, J. Tanji, *Science* **282**, 1335 (1998).
52. D. T. Stuss, M. P. Alexander, *Philosophical Transactions of the Royal Society B* **362**, 901 (2007).
53. A. Hampshire, J. Duncan, A. M. Owen, *J Neurosci* **27**, 6219 (Jun 6, 2007).
54. M. K. Johnson *et al.*, *Cogn Affect Behav Neurosci* **5**, 339 (Sep, 2005).
55. C. L. Raye, K. J. Mitchell, J. A. Reeder, E. J. Greene, M. K. Johnson, *J Cogn Neurosci* **20**, 852 (May, 2008).
56. L. L. Chao, L. Nielsen-Bohlman, R. T. Knight, *Electroencephalogr Clin Neurophysiol* **96**, 157 (Mar, 1995).
57. E. Feredoes, G. Tononi, B. R. Postle, *Proc Natl Acad Sci U S A* **103**, 19530 (Dec 19, 2006).
58. J. K. Nelson, P. A. Reuter-Lorenz, C. Y. Sylvester, J. Jonides, E. E. Smith, *Proc Natl Acad Sci U S A* **100**, 11171 (Sep 16, 2003).
59. D. Badre, A. D. Wagner, *Cereb Cortex* **15**, 2003 (Dec, 2005).
60. S. L. Thompson-Schill, M. D'Esposito, G. K. Aguirre, M. J. Farah, *Proc Natl Acad Sci U S A* **94**, 14792 (Dec 23, 1997).
61. J. M. Chein, A. B. Moore, A. R. Conway, *Neuroimage* **54**, 550 (Jan 1, 2011).
62. M. J. Kane *et al.*, *J Exp Psychol Gen* **133**, 189 (Jun, 2004).
63. G. Northoff *et al.*, *Neuroimage* **31**, 440 (May 15, 2006).

64. K. B. McDermott, K. K. Szpunar, S. E. Christ, *Neuropsychologia* **47**, 2290 (Sep, 2009).
65. T. D. Wager *et al.*, in *Handbook of Emotions,* M. Lewis, J. M. Haviland-Jones, L. F. Barrett, Eds. (Guilford Press, New York, 2008), pp. 249-271.
66. K. Wang *et al.*, *PLoS One* **4**, e4867 (2009).
67. K. Christoff, A. M. Gordon, J. Smallwood, R. Smith, J. W. Schooler, *Proc Natl Acad Sci U S A* **106**, 8719 (May 26, 2009).
68. M. F. Mason *et al.*, *Science* **315**, 393 (Jan 19, 2007).
69. D. H. Weissman, K. C. Roberts, K. M. Visscher, M. G. Woldorff, *Nat Neurosci* **9**, 971 (Jul, 2006).
70. T. Eichele *et al.*, *Proc Natl Acad Sci U S A* **105**, 6173 (Apr 22, 2008).
71. S. M. Daselaar *et al.*, *Front Hum Neurosci* **3**, 13 (2009).
72. J. Rissman, A. Gazzaley, M. D'Esposito, *Neuroimage* **23**, 752 (Oct, 2004).
73. K. A. McKiernan, B. R. D'Angelo, J. N. Kaufman, J. R. Binder, *Neuroimage* **29**, 1185 (Feb 15, 2006).

Table 1. Regions activated during encoding/maintenance that predict subsequent accuracy

		Coordinates			Cluster size		Effect size	[Enc/M - Base]		Accuracy		Distractor (D/ND)		Acc. x Distractor	
Region group	**Region**	**x**	**y**	**z**	**Vox**	**mm³**	**Max -log(P)**	**t**	**p**	**t**	**p**	**t**	**p**	**t**	**p**
Positive: Increasing activation associated with greater accuracy															
Prefrontal regions of interest (ROIs)															
Inferior lateral PFC	IFG/BA44 (L)	-40	10	26	116	928	12.34	**6.80**	<.0001	**2.90**	0.0052	**1.83**	0.0433	0.00	0.4990
	IFG/BA44 (R)	50	14	26	5	40	8.55	**3.28**	0.0024	**5.23**	0.0001	-0.68	0.2522	0.79	0.2204
Premotor	PMC/BA6 (L)	-40	-4	54	2	16	10.19	**3.68**	0.0010	**2.03**	0.0296	**2.61**	0.0094	-0.57	0.2881
	PMC/BA6 (L)	-28	-16	72	5	40	7.48	**3.61**	0.0012	**3.41**	0.0018	**1.95**	0.0347	-0.80	0.2166
Medial prefrontal	Pre-SMA	-2	10	50	5	40	6.99	**10.48**	<.0001	**4.44**	0.0002	1.49	0.0778	0.85	0.2049
	Pre-SMA (R)	6	8	52	20	160	11.04	**7.63**	<.0001	**3.06**	0.0038	0.24	0.4082	**1.86**	0.0407
Insula/operculum	dAINS (R)	34	20	0	2	16	10.15	**6.31**	<.0001	**4.07**	0.0005	-0.41	0.3425	1.15	0.1328
Other regions within [Encoding/Maintenance - Baseline] mask															
Parietal cortex	IPL/PGa (L)	-34	-68	46	7	56	9.57	**4.69**	0.0001	**6.55**	<.0001	0.47	0.3223	0.25	0.4034
Temporal cortex	STG (L)	-48	10	-14	4	32	8.29	**3.89**	0.0007	**3.60**	0.0012	-0.31	0.3792	0.64	0.2642
	ITG (L)	-42	-54	-8	16	128	12.12	**5.41**	0.0001	**3.52**	0.0014	-1.56	0.0696	1.14	0.1358
Occipital cortex	LOCC/V4 (L)	-44	-84	-8	23	184	8.04	**6.40**	<.0001	**3.47**	0.0016	-0.99	0.1681	0.89	0.1921
Negative: deactivation or reduced activity associated with greater accuracy															
Default mode/sensory processing ROIs: stronger deactivation associated with greater accuracy															
Ventromedial PFC	pgACC/BA10/BA32	-4	54	12	378	3024	9.26	**-4.26**	0.0003	**-4.09**	0.0005	-0.68	0.2525	-1.03	0.1594
	rMPFC/BA9/BA32 (R)	10	50	16	17	136	7.17	**-3.44**	0.0017	**-3.32**	0.0022	-1.68	0.0564	-0.16	0.4361
Rostral DMPFC	DMPFC/BA8 (R)	8	36	48	21	168	7.5	**-3.62**	0.0012	**-3.59**	0.0013	-0.25	0.4041	-1.21	0.1212
Somatic sensory	SII (R)	50	-26	22	108	864	6.3	**-5.23**	0.0001	**-3.64**	0.0011	-1.67	0.0571	**4.35**	0.0003
	dpINS (R)	36	-18	8	233	1864	10.99	**-4.86**	0.0001	**-3.76**	0.0009	-1.05	0.1552	**2.43**	0.0137
Inferior parietal	IPL/PGp (R)	52	-66	30	92	736	12.49	**-4.39**	0.0003	**-3.52**	0.0015	0.24	0.4079	-0.84	0.2055
	IPL/PGp (L)	-48	-76	30	76	608	9.31	**-3.56**	0.0013	**-4.62**	0.0002	-1.15	0.1340	-0.42	0.3395
	IPL/PF (L)	-66	-46	30	3	24	9.07	**-3.60**	0.0012	**-3.84**	0.0007	-1.08	0.1478	-0.61	0.2741
Other regions within [Encoding/Maintenance - Baseline] mask															
Medial temporal lobe	HCMP (R)	30	-36	-2	48	384	7.16	**5.73**	0.0000	**-2.99**	0.0044	**-1.75**	0.0493	0.11	0.4562
	HCMP (L)	-34	-28	-6	4	32	8.83	**3.73**	0.0009	**-3.82**	0.0008	-1.11	0.1412	-0.24	0.4063
	HCMP (L)	-20	-38	10	2	16	8.72	**3.68**	0.0010	**-3.23**	0.0026	-0.02	0.4919	-0.11	0.4558
Temporal cortex Audit	TOP/TE (R)	52	-6	4	8	64	6.27	**-3.39**	0.0019	**-3.00**	0.0042	-0.56	0.2932	1.16	0.1323
	ITG (R)	42	-42	-14	63	504	6.9	**4.92**	0.0001	**-4.16**	0.0004	0.57	0.2889	-1.27	0.1105
Superior temporal	STS (R)	56	-22	-4	132	1056	7.35	**6.98**	0.0000	**-3.64**	0.0011	-1.32	0.1031	**2.33**	0.0166
Occipital cortex	LOCC/BA17 (L)	-26	-76	6	5	40	8.34	**3.54**	0.0014	**-3.96**	0.0006	-0.50	0.3135	0.50	0.3104
	LOCC/MTG (L)	-42	-52	8	4	32	7.62	**4.79**	0.0001	**-4.47**	0.0002	0.48	0.3194	0.54	0.2980
	LOCC/V3/V4 (R)	30	-80	-10	3	24	7.34	**3.81**	0.0008	**-3.96**	0.0006	-1.62	0.0626	0.50	0.3107
	LOCC (L)	-36	-76	4	4	32	7.4	**3.74**	0.0009	**-3.35**	0.0021	**-2.38**	0.0151	1.62	0.0624
Frontal cortex	IFG/BA45/BA46 (R)	40	28	14	4	32	8.92	**4.03**	0.0005	**-4.68**	0.0001	-0.96	0.1751	1.37	0.0950
	SMC (R)	60	-12	46	18	144	6.88	**5.77**	0.0000	**-3.35**	0.0021	-0.19	0.4265	0.86	0.2023
	PMC (R)	30	-16	66	6	48	10.57	**3.53**	0.0014	**-5.01**	0.0001	-0.18	0.4302	0.08	0.4677
Parietal cortex	IPS (L)	-32	-52	34	1	8	7.7	**4.27**	0.0003	**-4.39**	0.0003	0.07	0.4724	0.85	0.2027
Medial frontal cortex	MCC (R)	8	-4	26	5	40	6.66	**3.64**	0.0011	**-2.42**	0.0140	**-2.09**	0.0267	1.63	0.0612
Anterior cingulate	aMCC (R)	8	28	24	27	216	9.12	**3.55**	0.0014	**-5.26**	0.0001	-1.79	0.0460	1.54	0.0717
Basal ganglia	Cau (R)	12	26	6	16	128	10.47	**4.15**	0.0004	**-8.11**	0.0000	0.44	0.3329	1.05	0.1547
	Cau (L)	-12	22	0	189	1512	9.68	**4.20**	0.0004	**-3.39**	0.0019	**2.44**	0.0133	-0.45	0.3291

Figure Captions

Figure 1. Task design. During Encoding, memoranda were presented every 500 msec, with the number of words for each participant determined before the main task using an adaptive procedure targeting 75% accuracy. During Maintenance, participants rehearsed the memoranda. On Distracter trials, Maintenance was followed with a secondary task in which participants made noun/verb judgments on four words. On No-Distracter trials, participants continued maintaining the items for the same duration as the secondary task on Distracter trials (6800 msec). In the Probe period, participants saw four words and made yes/no judgments as to whether they were part of the memory set.

Figure 2. fMRI activity related to working memory encoding/maintenance. A) Activations and deactivations in the [Encoding/Maintenance – Baseline] contrast, shown in red/yellow colors and purple/blue, respectively. The primary threshold was $p < 10^{-6}$ (yellow/dark blue for positive/negative effects) and 3 contiguous voxels. Contiguous voxels at lower thresholds that are contiguous with those at the primary threshold are also shown, at $p < 10^{-5}$ (orange/light blue) and $p < .009$ ($q < .01$ False Discovery Rate corrected, shown in pink/purple). This stringent threshold was chosen to provide anatomically specific regions of interest, though activity in this contrast was widespread. The anatomical underlay is the average T1-weighted image from this sample, after normalization to Montreal Neurological Institute space. B) Results from meta-analyses of working memory (WM), long-term memory (LTM), task switching, and inhibition tasks for comparison with (A). WM is divided into executive (Exec) and storage-only tasks. The black lines show anatomical divisions between Brodmann's Areas as implemented in Caret software (http://brainvis.wustl.edu/wiki/index.php/Caret:About). Figure adapted from Van Snellenberg et al. (2009).

Figure 3. Regions both responsive to working memory demand and predictive of subsequent accuracy. A) Surface rendering of areas significant in the conjunction of [Encoding/Maintenance – Baseline] and subsequent accuracy. In yellow/orange areas, activity increases predicted higher accuracy, and in blue/purple areas, activity decreases predicted higher accuracy. Right: axial slices showing results for this conjunction. The

threshold is $p < .001$ (yellow/dark blue), $p < .005$ (orange/light blue), and $p < .01$ (pink/purple), with voxels at the latter thresholds contiguous with those at higher thresholds. B) Left: Detailed localization of sensory regions, compared with the anatomical regions from the SPM Anatomy Toolbox v.1.7. Deactivations predicting accurate memory were localized in somatosensory areas OP1 (SII) and dorsal posterior insula (extending into Ig1) and auditory areas TE1.1/1.0. Right: Increases in pre-SMA and decreases in ventromedial prefrontal cortex predictive of accurate performance. The full extent of activation is shown for display purposes, although only the more dorsal portion was significant in the [Encoding/Maintenance – Baseline] contrast.

Figure 4. Pre-supplementary motor area results. A) The significant region in used as a functional region of interest (ROI) for distraction analyses. Left, [Encoding/Maintenance – Baseline] results, with thresholds as in Figure 2. Right: Voxels within this area correlating with accuracy, with thresholds as in Figure 3, located in the inferior frontal gyrus. B) Average activity in the accuracy-predictive ROI across the entire trial, for each level of accuracy. During the critical Encoding/Maintenance period, activity on trials with the highest level of accuracy (3 or 4 correct) is higher than trials with lower levels of accuracy. High levels of activation were also found during the probe interval, which is typical for studies of WM and consistent with findings on retrieval-related frontal activation, but is not of primary interest here. C) Bar plots of activity in the critical Encoding/Maintenance period (*y*-axis) as a function of accuracy (*x*-axis) and Distracter status. The positive overall relationship between brain activity and accuracy is driven by the Distracter condition.

Figure 5. Ventrolateral prefrontal cortex results. A) The significant region in used as a functional region of interest, as Figure 4. B) Average activity in the accuracy-predictive ROI across the entire trial, for each level of accuracy. C) Bar plots of activity as a function of accuracy and Distracter status. The relationship between brain activity and accuracy held for both Distracter and No-Distracter trials.

Figure 6. Accuracy *x* Distraction interaction results across the functional cortex. These results are not constrained to functional regions of interest. In yellow/orange areas, the activity-performance relationship was more positive for the Distracter trials. In blue areas, the relationship was more positive for the No-Distracter trials. Thresholds and colors are as in Figure 3. Whole-brain results for the Accuracy x Distraction interaction are shown in Supplementary Figure 1.

Figure 7. Accuracy *x* Distraction interaction results in default mode and sensory regions of interest. A) Left, regions predictive of accuracy and significantly activated during Encoding/Maintenance. Right: Voxels showing Accuracy *x* Distraction interactions. Thresholds and colors are as in Figure 3 In yellow/orange areas, the activity-performance relationship was more positive for the Distracter trials. B) Detail for SII. In sensory areas, including SII and dorsal-posterior insula, deactivation was more strongly related to performance in the No-distracter condition. C) In anterior medial prefrontal cortex (aMPFC), deactivation was more strongly related to performance in the Distracter condition.

Figure 1.

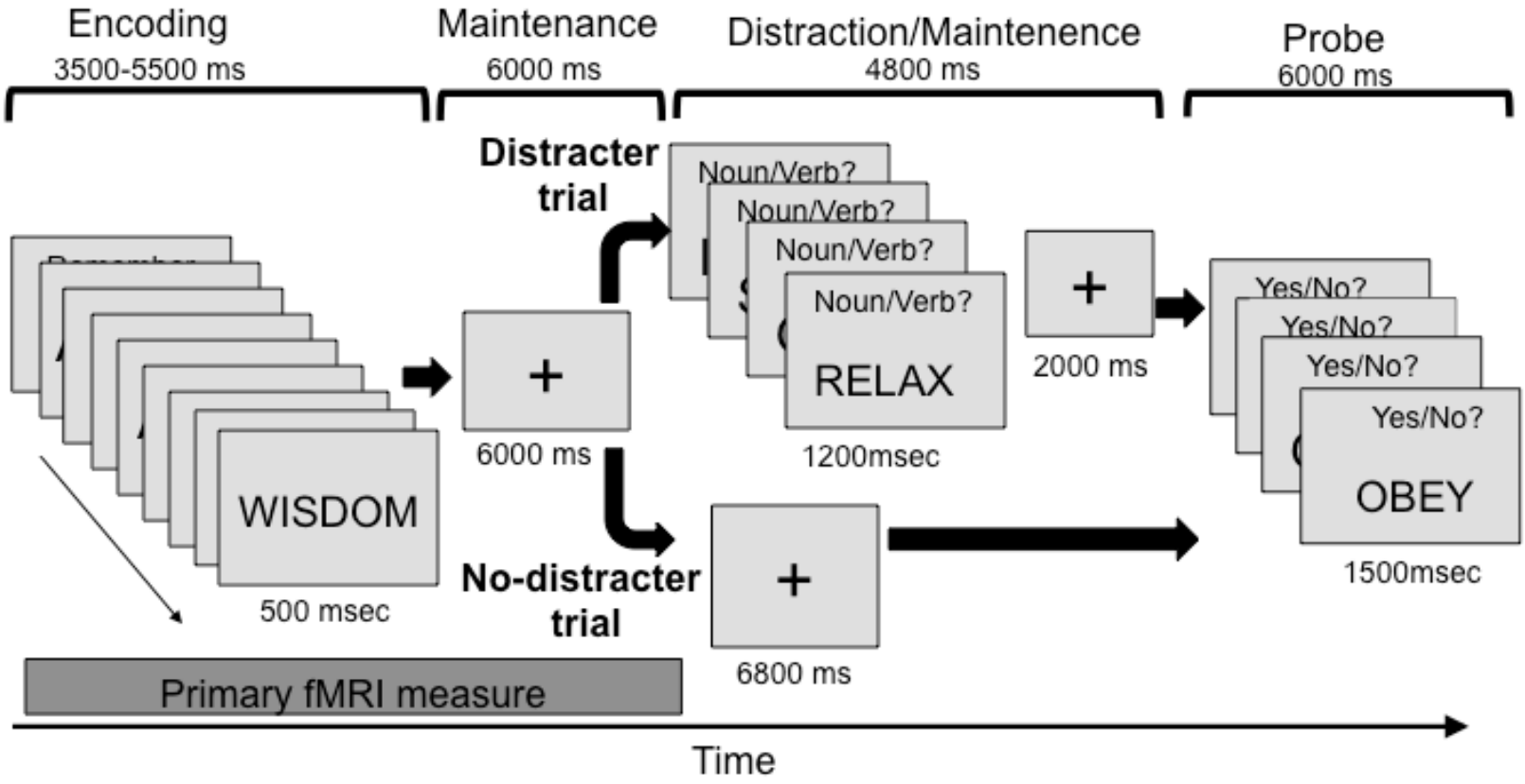

Encoding
3500-5500 ms
Maintenance
6000 ms
Distraction/Maintenence
4800 ms
Probe
6000 ms
Distracter trial
Noun/Verb?
Noun/Verb?
Noun/Verb?
Noun/Verb?
RELAX
1200msec
+
2000 ms
Yes/No?
Yes/No?
Yes/No?
Yes/No?
OBEY
1500msec
WISDOM
500 msec
+
6000 ms
No-distracter trial
+
6800 ms
Primary fMRI measure
Time

Figure 2. Activations and deactivations during encoding/maintenance, [Encoding/Maintenance – Baseline]

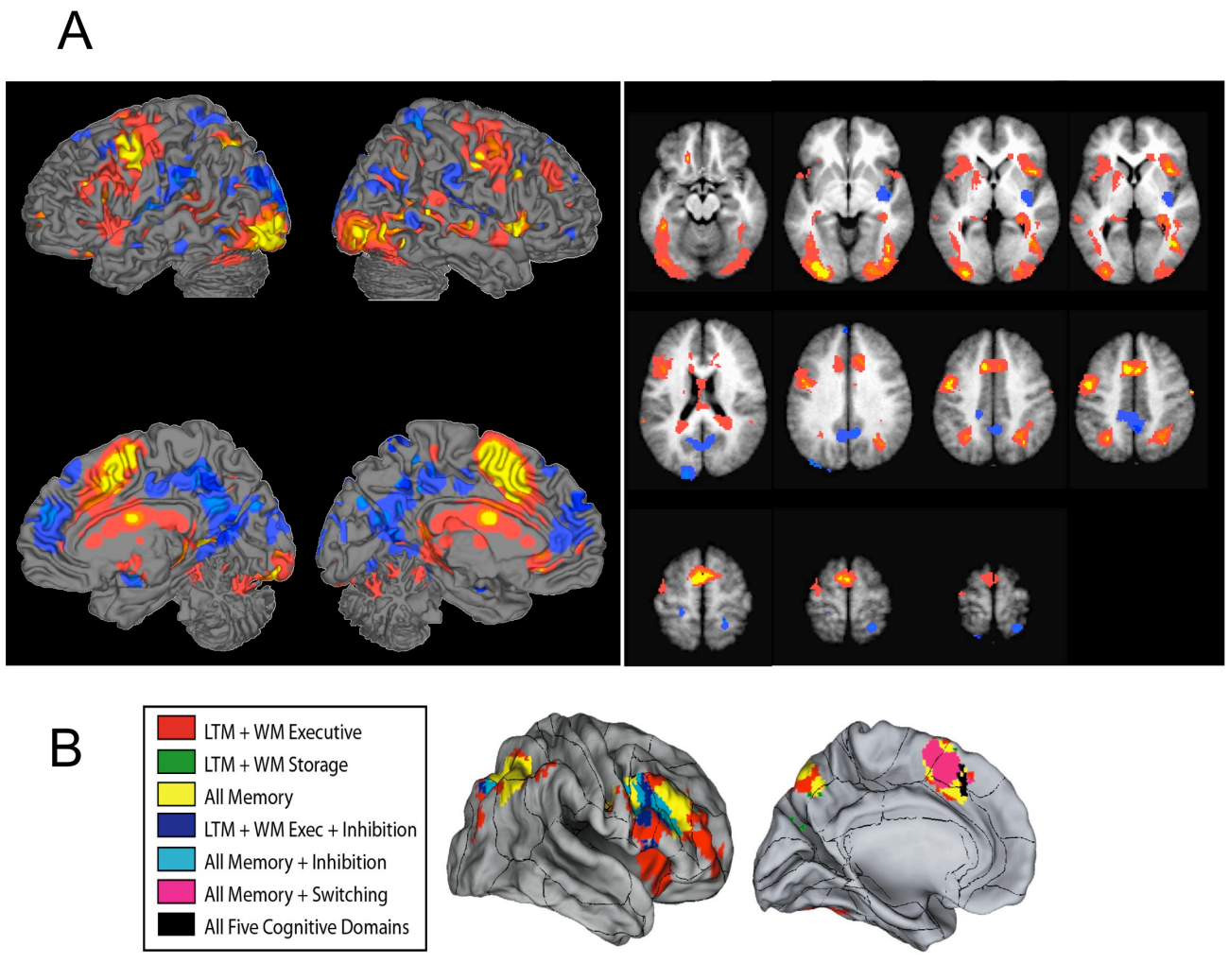


Figure 3. Increases predicting subsequent accuracy, within significant [Encoding/Maintenance – Baseline] regions

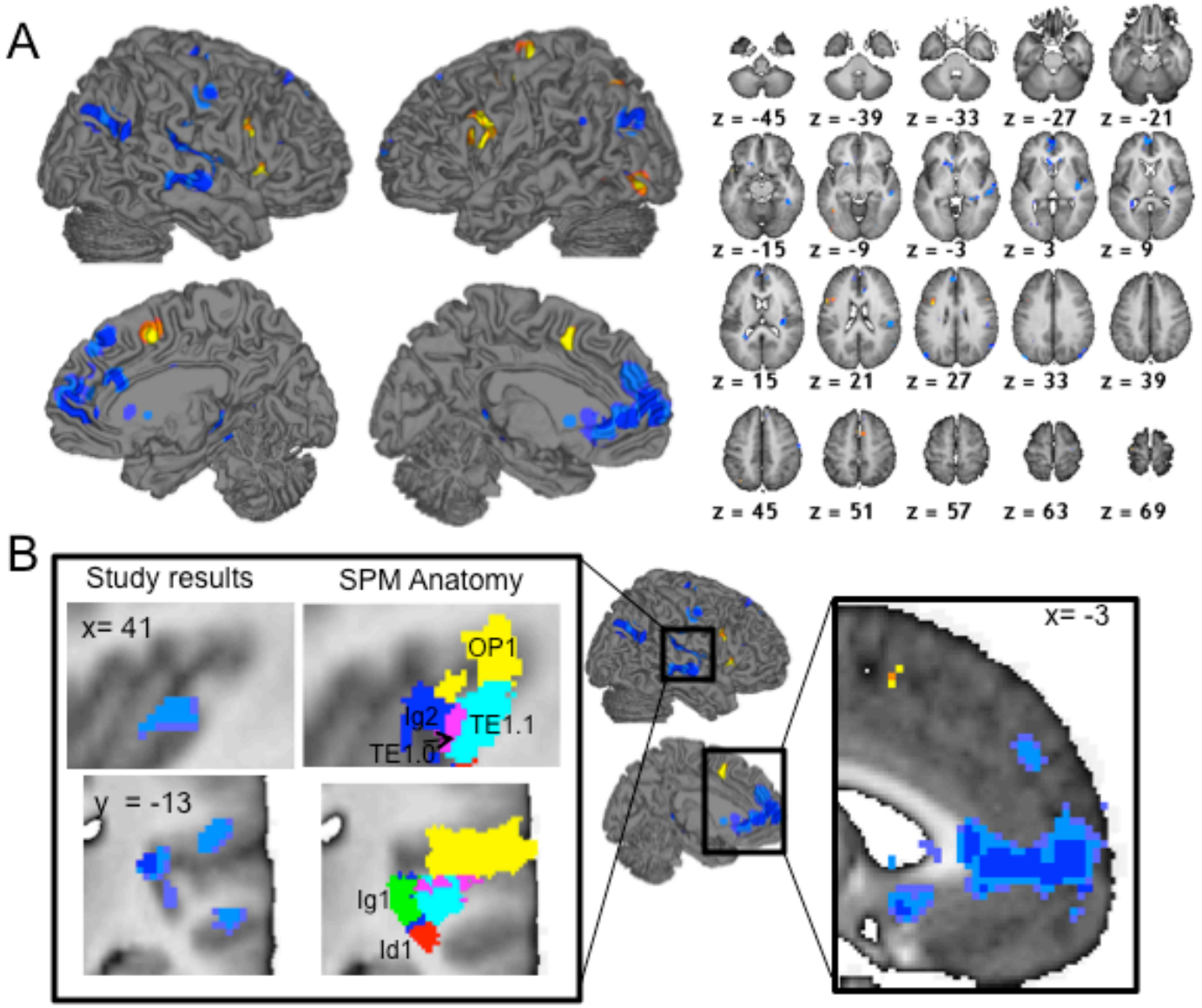


Figure 4. Increases in pre-supplementary motor area predicting subsequent accuracy.

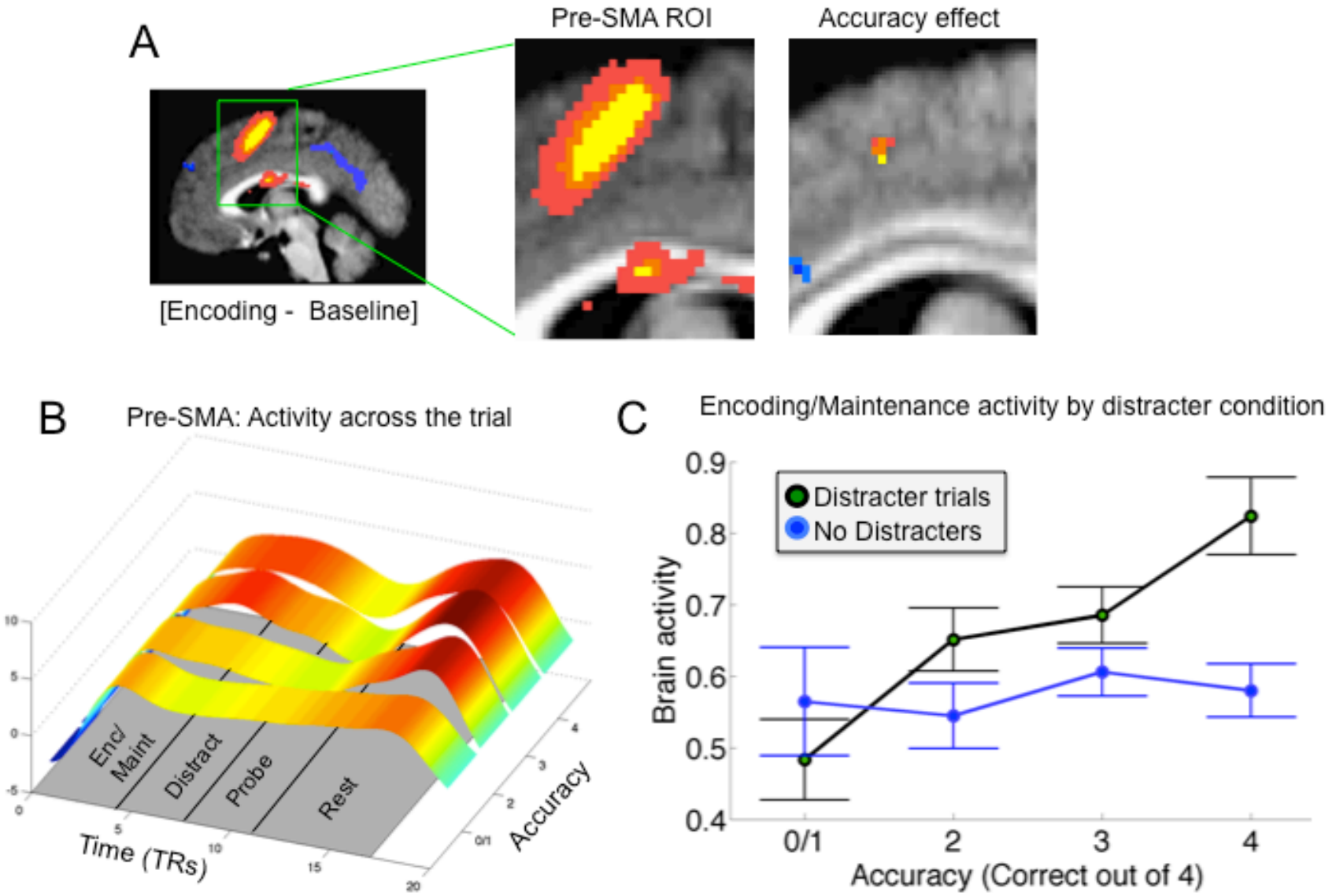


Figure 5. Increases in lateral prefrontal cortex predicting subsequent accuracy

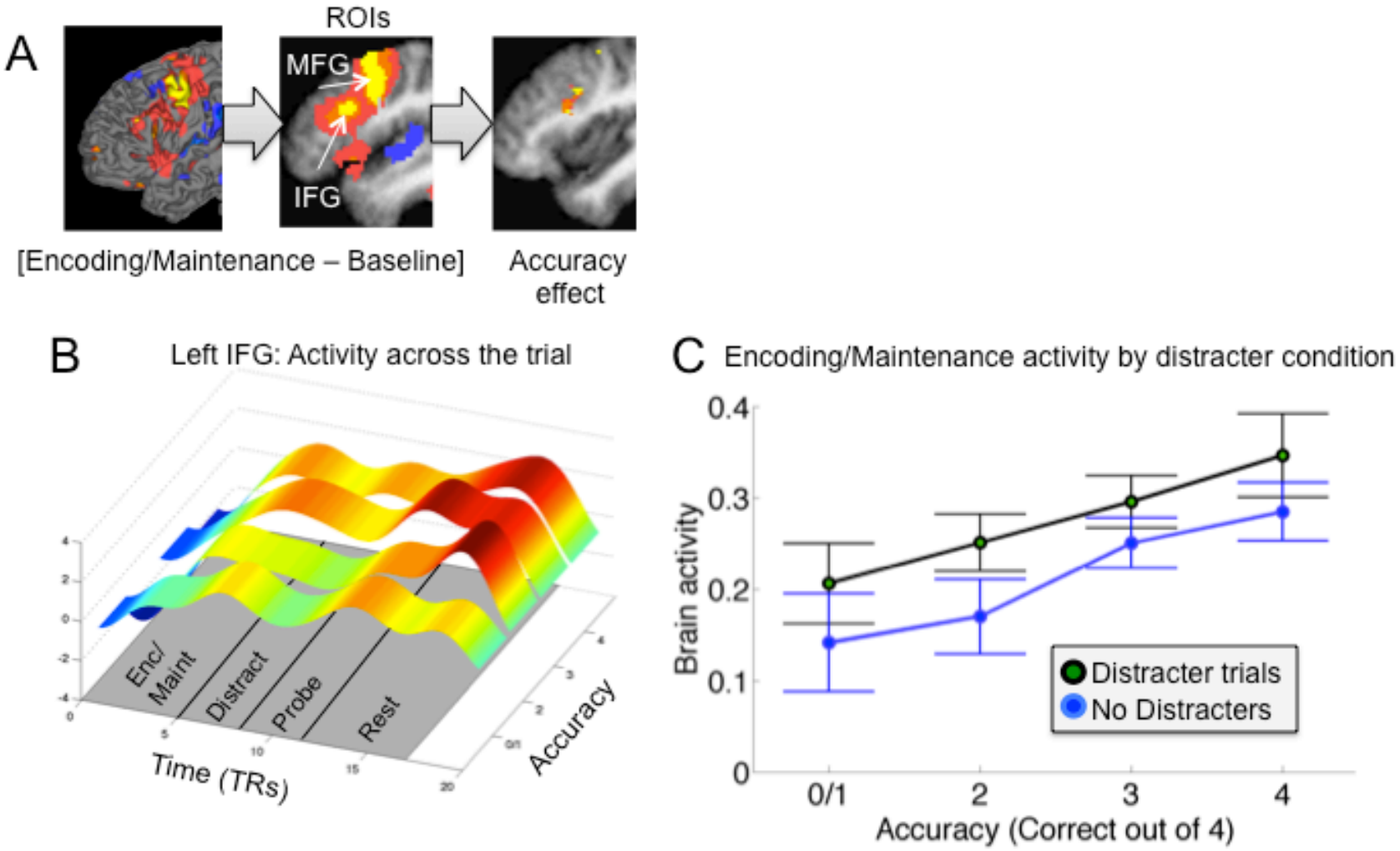


Figure 6.

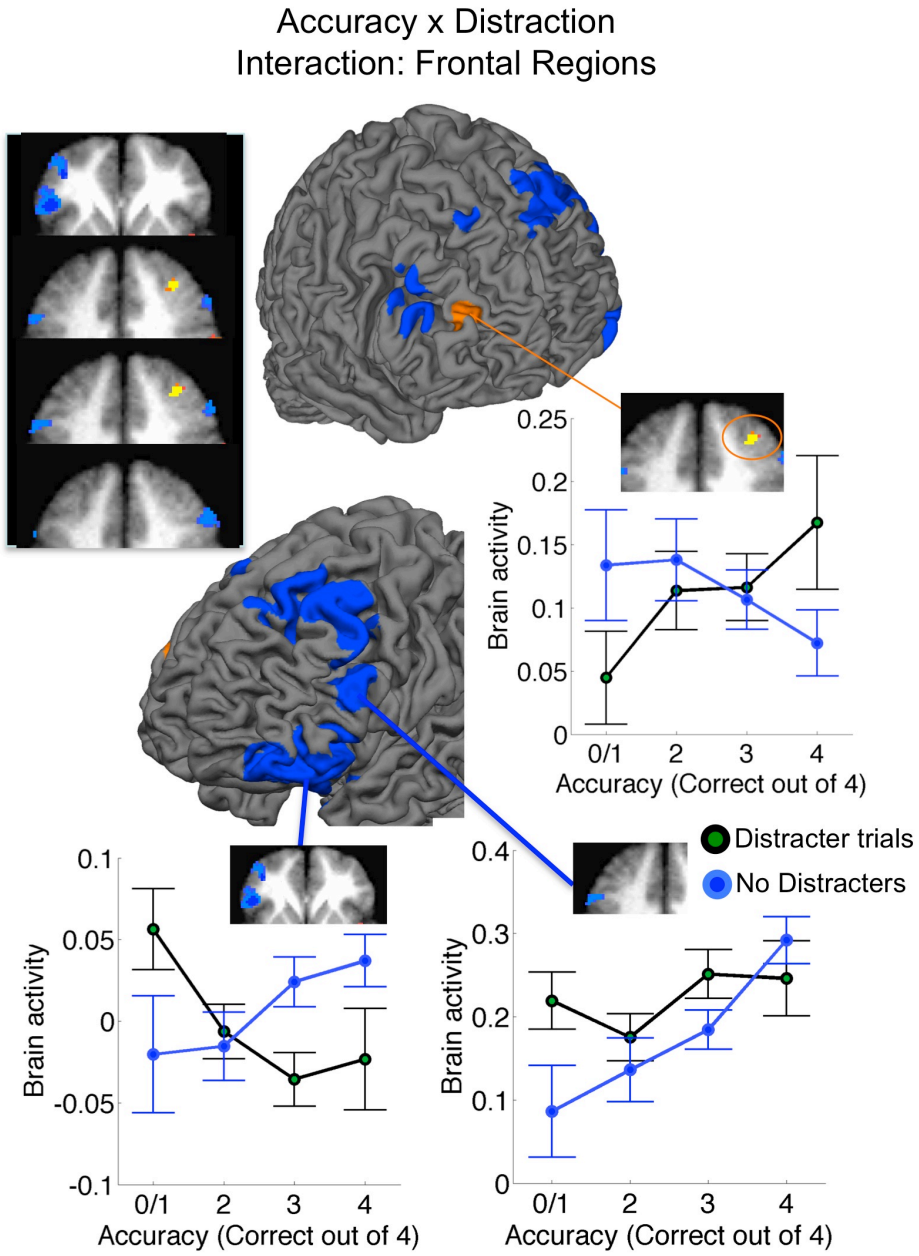
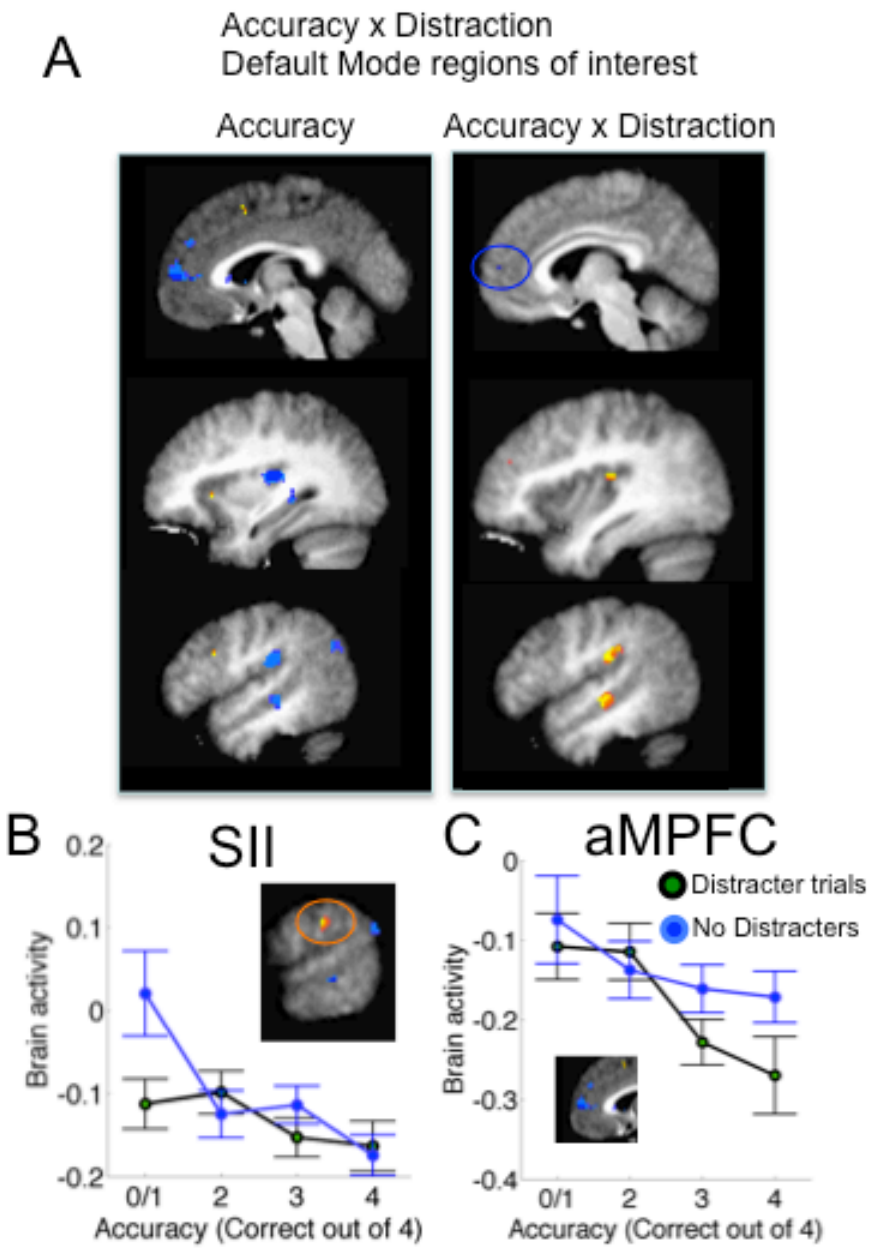

Accuracy x Distraction
Interaction: Frontal Regions
Brain activity
0.25
0.2
0.15
0.1
0.05
0
0/1
2
3
4
Accuracy (Correct out of 4)
Distracter trials
No Distracters
0.1
0.05
0
-0.05
-0.1
0.4
0.3
0.2
0.1

Figure 7.


A
Accuracy x Distraction
Default Mode regions of interest
Accuracy
Accuracy x Distraction
B
SII
Brain activity
0.2
0.1
0
-0.1
-0.2
0/1
2
3
4
Accuracy (Correct out of 4)
C
aMPFC
Distracter trials
No Distracters
Brain activity
0
-0.1
-0.2
-0.3
-0.4
0/1
2
3
4
Accuracy (Correct out of 4)