Behavioral/Systems/Cognitive

Prefrontal Dynamics Underlying Rapid Instructed Task Learning Reverse with Practice

Michael W. Cole,[1,2] **Anto Bagic,**[3,4] **Robert Kass,**[5] **and Walter Schneider**[1,6]

[1]Center for the Neural Basis of Cognition, Center for Neuroscience, and Learning Research and Development Center, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, [2]Department of Psychology, Washington University, St. Louis, Missouri 63130, [3]Department of Neurology and Neurosurgery, University of Pittsburgh Medical Center (UPMC) and [4]UPMC Center for Advanced Brain Magnetic Source Imaging, University of Pittsburgh, Pittsburgh Pennsylvania 15213, [5]Department of Statistics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, and [6]Department of Psychology, University of Pittsburgh, Pittsburgh, Pennsylvania 15260

The ability to rapidly reconfigure our minds to perform novel tasks is important for adapting to an ever-changing world, yet little is understood about its basis in the brain. Furthermore, it is unclear how this kind of task preparation changes with practice. Previous research suggests that prefrontal cortex (PFC) is essential when preparing to perform either novel or practiced tasks. Building upon recent evidence that PFC is organized in an anterior-to-posterior hierarchy, we postulated that novel and practiced task preparation would differentiate hierarchically distinct regions within PFC across time. Specifically, we hypothesized and confirmed using functional magnetic resonance imaging and magnetoencephalography with humans that novel task preparation is a bottom-up process that involves lower-level rule representations in dorsolateral PFC (DLPFC) before a higher-level rule-integrating task representation in anterior PFC (aPFC). In contrast, we identified a complete reversal of this activity pattern during practiced task preparation. Specifically, we found that practiced task preparation is a top-down process that involves a higher-level rule-integrating task representation (recalled from long-term memory) in aPFC before lower-level rule representations in DLPFC. These findings reveal two distinct yet highly inter-related mechanisms for task preparation, one involving task set formation from instructions during rapid instructed task learning and the other involving task set retrieval from long-term memory to facilitate familiar task performance. These two mechanisms demonstrate the exceptional flexibility of human PFC as it rapidly reconfigures cognitive brain networks to implement a wide variety of possible tasks.

Introduction

Humans are uniquely capable of rapidly learning a nearly infinite variety of possible new tasks from instruction (Braver and Barch, 2006; Monsell, 1996). This form of learning allows for rapid adaptation to novel situations and transfer of previously learned skills, vastly increasing the cognitive flexibility of our species. Despite the prominent contribution of rapid instructed task learning (RITL) to human cognition, little is known about how it is implemented in the brain. This is in contrast to practiced task preparation, which has been investigated extensively.

One thing that is clear about RITL is that it relies upon prefrontal cortex (PFC) (Burgess, 1997; Dumontheil et al., 2010; Ruge and Wolfensteller, 2010). Luria (1973) found that some PFC lesions produced a striking inability to convert novel instructions into task per-

formance, despite intact working memory (WM) and linguistic abilities. This suggests that there is a specific preparatory process within PFC that is necessary for RITL. We sought to identify this process here by using functional magnetic resonance imaging (fMRI) and magnetoencephalography (MEG).

Studies of task switching also strongly support a role for PFC in practiced task preparation (Sakai, 2008). Recently, evidence has emerged for a specific preparatory mechanism in which activity in anterior PFC (aPFC) drives activity in dorsolateral PFC (DLPFC) and other posterior regions to specify task sets (Sakai and Passingham, 2003, 2006). This preparatory mechanism may reflect an anterior-to-posterior hierarchical relationship within PFC (Badre, 2008) in which more anterior regions implement "higher-level" processing than posterior regions.

Based on this hierarchical relationship, we hypothesized that aPFC-to-DLPFC information flow during practiced task preparation would reverse (becoming DLPFC-to-aPFC) during novel task preparation. This reversal would reflect the difference between top-down preparation for practiced tasks [involving specification of an rule-integrated task set from cued memory (Mayr and Kliegl, 2000)] and bottom-up preparation during RITL (involving specification of task sets from individual instruction rules).

We specifically predicted that the reversal within PFC would be reflected in time, with aPFC becoming active earlier than DLPFC during practiced task preparation and DLPFC becoming active earlier than aPFC during novel task preparation (Fig. 1).

Such a result would be compatible with DLPFC rule representations (Miller et al., 2002) becoming integrated to form a higher-level task representation in aPFC (De Pisapia et al., 2007) during RITL, which would be reactivated from memory during practiced task preparation.

We developed a new cognitive paradigm to test these predictions (Fig. 2). Most studies of task preparation include practiced rather than novel tasks, likely because of difficulty assessing statistical reliability when even a single repetition invalidates the task's novelty. The paradigm used here employs 64 unique tasks to observe statistically reliable effects during novel task preparation. Importantly, four of the tasks are practiced, providing controls for stimulus novelty and task switching [both known to involve PFC (Daffner et al., 2000; Dove et al., 2000)]. In addition to testing our hypotheses regarding PFC, this new paradigm allowed us to verify whether humans are truly capable of rapidly learning a wide variety of complex novel tasks without prior practice.

Materials and Methods

Participants. We included 15 right-handed participants (eight male, seven female), aged 19–29 (mean age 22), in the fMRI experiment. We included eight right-handed participants (three male, five female), aged 18–30 (mean age 23), in the MEG experiment. These participants were recruited from the University of Pittsburgh (Pittsburgh, PA) and the surrounding area. Participants were excluded if they had any medical, neurological, or psychiatric illness, any contraindications for MEG or MRI scans, were nonnative English speakers, or were left handed. All participants gave informed consent.

Task paradigm. The fMRI and MEG experiments consisted of performing the permuted rule operations (PRO) cognitive paradigm, which was developed as part of this study. The PRO paradigm combines a set of simple rule components in many different ways, creating dozens of complex task sets certain to be novel to participants (Fig. 2). The paradigm was presented using E-Prime software (Schneider et al., 2002).

Four sensory-semantic, four logic, and four motor response rules were used in the paradigm. Each semantic judgment task consisted of one rule from each of these categories, allowing the creation of $4 \times 4 \times 4 = 64$ distinct tasks by permuting the possible rules. Of these tasks, four (counterbalanced across participants) were practiced (30 blocks, 90 trials each) during a 2 h behavioral session 1–7 d before the neuroimaging session. These "practiced" tasks were chosen for each subject such that each rule was included in exactly one of the four tasks, ensuring that all rules were equally practiced. During the neuroimaging session, half of the blocks consisted of the practiced tasks and half of novel tasks. Novel and practiced blocks were randomly interleaved for the fMRI experiment, with

the constraint that exactly six blocks of each type occur within every run. With 10 runs total per participant, each novel task was presented in one block and each practiced task was presented in 15 blocks.

The semantic rules consisted of sensory semantic decisions (e.g., "is it sweet?"). The logic rules specified how to respond based on the semantic decision outcome(s) for each trial (Fig. 2, large box). The motor response rules specified which button to press based on the logic decision out-

Figure 1. Task preparation hypotheses predict reversal of information flow in PFC. **A**, Previous observations of practiced task preparation (Mayr and Kliegl, 2000; Sakai and Passingham, 2006) suggest that aPFC may initially load a goal/"higher-level task set" representation from LTM that subsequently drives activation of the individual rules in DLPFC. We sought to verify these predictions here. This pattern of results would be compatible with the cognitive processes demonstrated by Rubinstein et al. (2001) in which practiced task preparation involves higher-level goal shifting (during task encoding) followed by lower-level rule activation (during the first trial). **B**, Based on their hierarchical relationship, we predicted that aPFC and DLPFC would reverse their processing order and direction of influence during RITL. This would reflect the shift from primarily top-down processing for practiced task preparation to bottom-up processing for novel task preparation. During novel task preparation we expected individual task rules to be activated in DLPFC before being integrated into a goal/"higher-level task set" representation within aPFC for coordination of subsequent task performance. We refer to this process of activating and coordinating a novel set of task rules as "task set formation."
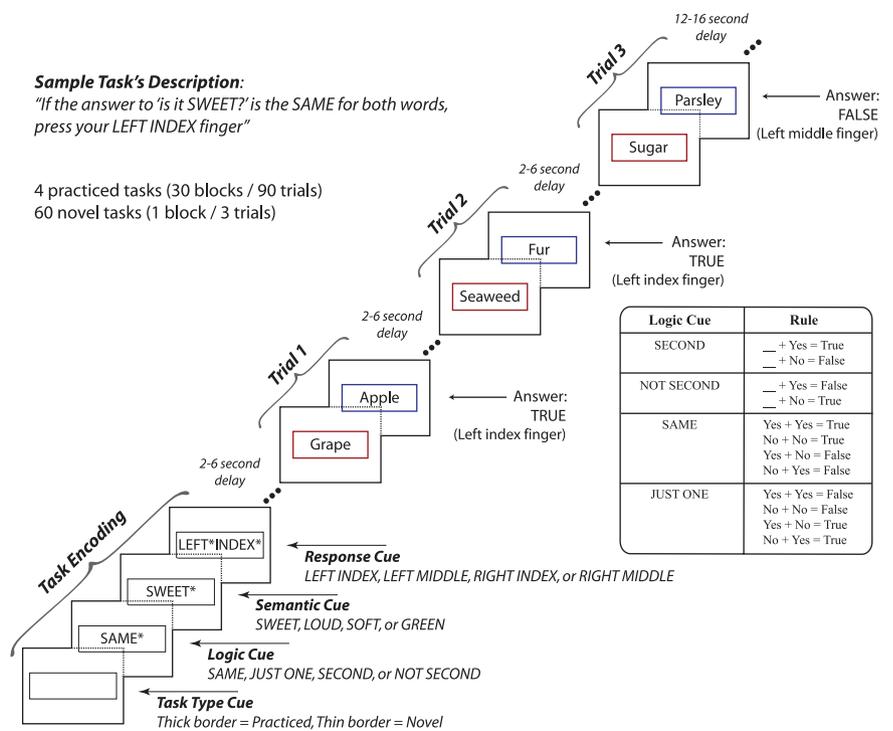
Figure 2. A novel cognitive paradigm for investigating rapid instructed task learning. The permuted rule operations paradigm involves permuting 12 task rules to create 64 unique tasks. Four of these tasks (counterbalanced across participants) were practiced in a prior behavioral session, while the remaining 60 were completely novel. Importantly, the individual rules were equally practiced while the task sets were not. Each task consisted of a logic rule, a sensory semantic rule, and a response rule, with four possible rules of each type. A block consisted of four encoding screens followed by three trials. The novel and practiced conditions were designed to be directly compared to each other, as they both control for stimulus novelty and task switching effects (both known to involve PFC) in addition to other general preparatory processes.

come. The task instructions made explicit reference to the motor response for a "true" outcome, while participants knew (from the practice session) to use the other finger on the same hand for a "false" outcome.

Figure 2 illustrates an example task block, which includes encoding and three trials. Each block began with a task type cue, indicating whether the upcoming task is novel (thin border) or practiced (thick border), followed by three instruction screens. This cue was included to aid participants implementing different strategies across the different task types. The order of the instructions following the task type cue was consistent for each participant, but counterbalanced across participants. Asterisks filled in extra spaces in each instruction screen to control for differences in visual stimulation across task rules. Each stimulus was presented for 800 ms with a 200 ms interstimulus interval. Inter-event intervals were randomly varied between 2 and 6 s (in 2 s intervals) for the fMRI experiment and 2 and 3 s (in 500 ms intervals) for the MEG experiment.

Stimuli were normalized by a separate group of participants (21 male, 33 female). Word stimuli were included in the neuroimaging experiments with exactly two category questions: one in which the normalizing group members answered "yes" >75% of the time and one in which they answered "no" >75% of the time. Exactly 45 stimuli were included per semantic category. Each stimulus was presented exactly 8 times (50% in a "yes" context, 50% in a "no" context) in the neuroimaging experiments, across both the behavioral practice (50%) and scanning (50%) sessions.

The cognitive paradigm used in the MEG experiment was virtually identical to the paradigm used in the fMRI experiment. There were, however, several differences: the intertrial delays were shortened (2–3 s jittered) to take advantage of the high temporal resolution of MEG; the amount of practice before scanning was doubled (same amount of practice time with half as many short intertrial delays and therefore twice as many trials); blocks were presented in "novel" and practiced epochs to encourage distinct strategies between the conditions; and no task repetitions were allowed.

MRI data collection. Image acquisition was performed on a 3T Siemens Trio MRI scanner. Thirty-eight transaxial slices were acquired every 2000 ms (field of view, 210 mm; echo time, 30 ms; flip angle: 90°; voxel dimensions, 3.2 mm^3) with a total of 216 echo-planar imaging volumes collected per run. Siemens' implementation of generalized autocalibrating partially parallel acquisition was used to double the image acquisition speed (Griswold et al., 2002). Three-dimensional anatomical MP-RAGE (magnetization prepared rapid acquisition gradient echo) images and T2 structural in-plane images were collected for each subject before fMRI data collection.

MEG data collection. Data acquisition was performed with a Vectorview (Elekta Neuromag) 306-sensor MEG system. To monitor eye movements and eye blinks, two bipolar electrode pairs were used for vertical and horizontal electrooculogram recordings. Two electrocardiogram electrodes were placed on each participant's chest to record heart activity. Four head position indicator (HPI) coils were attached onto each participant's scalp. Next, the three-dimensional locations of three cardinal landmarks (nasion and left and right periauriculars), HPI coils, and additional (20+) points on the scalp were digitized using a Fastrak system (Polhemus) to allow subsequent coregistration of the MEG data with each participant's structural MRI.

The HPI coils were used to determine the position of each participant's head relative to the MEG sensor array before every scanning run (10 runs, 5 min each). The MEG data were acquired and saved continuously at a sampling rate of 1000 Hz with a recording lowpass filter of 330 Hz.

Behavioral analysis. All behavioral analyses were carried out in R (R Foundation for Statistical Computing, Vienna, Austria). Accuracy statistical tests were modeled in mixed-effect non-repeated-measures ANOVAs (subject as a random effect), while reaction time statistical tests were modeled in mixed-effect repeated-measures ANOVAs (subject as a random effect). All nonswitch blocks (1.2% of the blocks during fMRI) were removed before statistical analysis (for behavioral and neuroimaging analyses).

fMRI analysis. Preprocessing and analysis were performed using AFNI (Cox, 1996). Preprocessing consisted of standard slice timing correction, motion correction, and spatial smoothing (6 mm full width at half max-

imum). Freesurfer (Desikan et al., 2006) anatomical segmentations of the participants' skull-stripped brain volumes were used as masks to remove nonbrain voxels.

The preprocessed data were analyzed with a rapid event-related general linear model with no assumption of hemodynamic response shape. An 11 repetition time (22 s) finite impulse response window was modeled for each event type to ensure that the modeled time series returned to baseline. The two encoding (novel and practiced) and six trial (three novel, three practiced) event types were modeled separately, with rest periods acting as baseline. Incorrect trials were removed from analyses by modeling them with separate regressors (reducing contamination of correct trial responses).

The resulting statistical maps were transformed into Talairach space (Talairach and Tournoux, 1988) and analyzed using group ANOVAs. The first ANOVA estimated mean across-subject fMRI responses to novel and practiced encoding events. Condition (novel and practiced), time (11 time points), and subject (as a random effect) factors were estimated, and the condition × time interaction was analyzed for each voxel. The resulting map illustrates the statistical reliability of response shape differences between conditions.

The second group ANOVA estimated fMRI responses to trial events. The ANOVA included condition (novel or practiced), trial number (1, 2, or 3), and subject (as a random effect) factors. The effect of condition for the first trial and the main effect of condition (across the three trials) are reported.

fMRI multiple-comparisons correction. Multiple comparisons were corrected for the statistical maps using familywise error (FWE) cluster size thresholding (Forman et al., 1995), or false discovery rate (FDR) (Genovese et al., 2002). FWE cluster thresholds were estimated using Monte Carlo simulations (AFNI AlphaSim). The whole-brain FWE simulation was restricted to an anatomical mask of all gray matter (including both cortical and subcortical structures) and used a cluster-defining threshold of $p < 0.05$, resulting in a 0.05 FWE-corrected cluster threshold of 89 voxels.

Since we specifically hypothesized effects in PFC, we determined an FWE cluster threshold using an anatomical mask restricted to lateral PFC. This FWE simulation used a cluster-defining threshold of $p < 0.05$, resulting in a 0.05 FWE-corrected cluster threshold of 28 voxels.

FDR was used to correct multiple comparisons for the encoding period statistical map (for regions besides aPFC and DLPFC) to identify smaller regions that may have been missed with the FWE approach (no voxels survived the FDR threshold for the trial period statistical map). This exploratory contrast should be interpreted with caution given evidence that while FDR corrects for voxelwise false positives, it may not properly correct for clusterwise false positives (Chumbley and Friston, 2009).

fMRI region of interest analyses. The whole-brain statistical maps were used to define regions of interest (ROIs) for subsequent analyses. We eliminated "circular" selection biases (Kriegeskorte et al., 2009) when analyzing encoding-period data by using ROIs identified with trial period data and vice versa. See supplemental Results, available at www.jneurosci.org as supplemental material, for evidence supporting independent estimation of encoding and trial events. The MEG directed connectivity results used the fMRI ROIs, eliminating circularity in that analysis as well.

MEG dSPM data analysis. MEG was used to test the hypothesis that information flow reverses between novel and practiced task preparation. As a first step, we localized MEG activity to the cortical surface to test for a replication of the fMRI results using MEG.

The MEG data were preprocessed using signal space separation and MaxMove (to spatially align head position between runs) as implemented by MaxFilter software (Taulu et al., 2004, 2005). Using MATLAB 2007a (The MathWorks), electrooculogram and electroencephalogram signals were regressed out of the MEG signals to reduce eye blink (and cardiac) artifacts (Wallstrom et al., 2004) with a 60 s moving window. Any trials with MEG artifacts exceeding 200 picoteslas were removed from further analysis.

The data were further analyzed using minimum norm estimate (MNE) software (Hämäläinen and Ilmoniemi, 1994). Noise and baseline values were estimated using data during the interblock delays. Depth

weighting was used to help counter the localization bias toward the superficial surface (Lin et al., 2006b), and current orientation was loosely constrained to surface normal (default constraint value of 0.4) (Lin et al., 2006a). Dynamic statistical parametric mapping (dSPM) was used to localize statistically significant brain regions (Dale et al., 2000).

For group analysis, each subject's dSPM data were projected onto the cortical surface and morphed (Fischl et al., 1999) to a common Talairach (Talairach and Tournoux, 1988) template surface. All subjects' data were averaged on this common template surface. These group averaged dSPM values were subtracted between conditions and projected back to the surface for condition contrast analyses.

MEG directed connectivity analyses. Given that the fMRI results were replicated using MEG, we used directed connectivity to test the hypothesis that information flow reverses between aPFC and DLPFC across novel and practiced task preparation. We used two directed connectivity methods to test this hypothesis: Granger causality (GC) and the phase slope index (PSI).

GC analyses were implemented using the Causal Connectivity Toolbox (Seth, 2005, 2010), while the PSI analyses were implemented using MATLAB functions developed by Nolte et al. (2008). Before running the connectivity analyses, the aPFC and DLPFC ROIs from the fMRI experiment were morphed to each participant's brain surface, and MNE time series were extracted based on MNE software's inverse estimates (fully constrained to surface normal). We then preprocessed the MNE data by removing outlier time series (those with fluctuations >6 SDs from the median) and removing the linear trend and the temporal mean from each time series.

Each block of each condition was modeled separately to account for interblock and intersubject differences in timing. A large GC model order was used (100 ms) to account for a range of possible dynamics, as well as to avoid fitting the bandpass filter cutoffs (Seth, 2010). GC was estimated after application of a bandpass filter in the gamma band (40–150 Hz) and conversion to gamma band power (root mean square of nonoverlapping 50 ms time bins) because of evidence of long-distance neural interactions in this frequency range (Fries et al., 2007; Gregoriou et al., 2009), as well as correspondence of MEG gamma power with fMRI data (Niessing et al., 2005; Zumer et al., 2010). GC magnitudes were estimated as the log F ratio (Geweke, 1982).

PSI estimation was used in addition to the GC analysis based on evidence that PSI better accounts for volume conduction artifacts and high levels of noise (Nolte et al., 2008). Unlike GC, PSI uses the "imaginary" component of coherence, which reflects nonzero lag connectivity such that it is less sensitive to mixtures of independent sources (Nolte et al., 2004, 2008). The dominant direction of influence between two signals is estimated from the direction of increasing phase differences with increasingly high frequencies (i.e., an increase in the phase slope). Nonnormalized PSI estimates were obtained from each block separately (segment size was 500 ms).

GC and PSI estimates were obtained separately for each of the three encoding cues (1000 ms each, excluding the "task type" cue) and first trials (2000 ms each). *t* tests were used to test for dominant influences and differences in influences between conditions. All GC and PSI *t* tests were two tailed and paired by block. Quantile–quantile plots were used to verify that the GC/PSI values were approximately normally distributed before running the statistical tests.

Simulations verified the statistical specificity and sensitivity of the GC and PSI analyses. The simulations consisted of 1000 *t* tests for dominant directions of influence, with each *t* test based on GC and PSI estimates of 400 independently simulated trials (the MEG data included 413 blocks after rejection of outlier blocks). Simulated trials consisted of 1000 time points each, randomly sampled from a normal distribution, that were copied to a second time series with lag introduced to simulate directed connectivity. The lag was randomly selected for each trial from a range of 10–100 ms. Model parameters were identical to those used in the MEG analyses. The first simulation included equal bidirectional connectivity, with double the noise in one simulated channel (50 vs 25% Gaussian noise) to ensure that differential signal-to-noise ratio does not inflate GC estimates (with the particular approach used here), as has been suggested previously (Andersson, 2005; Schoffelen and Gross, 2009). The results

indicated a reasonably small rate of false positives (4.0% for gamma power GC, and 3.9% for PSI, with alpha at 5%). The second simulation included influences in only one direction (with 50% Gaussian noise) and indicated a reasonable rate of correct detections (100% for both methods).

MEG signal correlations were analyzed to support conclusions based on the region-to-region interactions inferred by the GC/PSI results. These correlations were based on mean encoding activity (gamma power across all three encoding cue periods) in aPFC and DLPFC for each block. Importantly, we sought to reduce the influence of zero-lag volume conduction effects that might artificially inflate inter-regional correlations. In addition to our use of source localization and separate-hemisphere ROIs to reduce volume conduction artifacts, we removed direct ROI-to-ROI zero-lag correlations using alternating 50 ms segments to estimate the mean encoding activities for each region. Odd numbered segments were used for aPFC while even numbered segments were used for DLPFC (alternating odd/even between the regions every block). Thus, the mean encoding activity correlations reflect interactions with lags greater than ± 25 ms.

Results
Behavioral results
The experimental paradigm was designed to maximize performance on novel task trials to test whether humans are truly capable of rapidly (i.e., in 5–10 s) learning novel tasks from verbal instruction. Accuracy on the very first performance of the 64 novel tasks (60 per subject, counterbalanced to include 64 total) was 91%. This high accuracy (chance was 25%) demonstrates that individuals are capable of accurately forming and implementing complex novel task sets within seconds.

We used mixed-effect repeated-measures ANOVAs (subjects as a random effect) to test for accuracy differences across novel and practiced trials. Analyses were collapsed across the fMRI and MEG experiments (results were similar when analyzed separately). Overall accuracy in the PRO paradigm was 92.1%. Accuracy for the novel tasks across all trials was $91.0 \pm 0.8\%$ (mean $\pm$ between-subject SE), and accuracy for the practiced tasks was $93.1 \pm 0.8\%$. There was a significant increase in accuracy for practiced relative to novel trials ($F_{(1,22)} = 9.22$, $p = 0.006$). Neither the main effect of trial number ($F_{(2,42)} = 1.45$, $p = 0.25$, NS) nor the trial number by condition interaction ($F_{(2,42)} = 0.43$, $p = 0.65$, NS) were significant.

The overall mean reaction time (RT) for all trials was 1295 ms. The mean RT for novel trials was 1306 ± 61 ms, while the mean RT for practiced trials was 1283 ± 60 ms. This difference was statistically significant: $F_{(1,22)} = 8.82$, $p = 0.007$. The main effect of trial number was marginally significant ($F_{(2,42)} = 2.96$, $p = 0.063$), while the trial number by condition interaction was not significant ($F_{(2,42)} = 0.016$, $p = 0.98$, NS).

The expected task switching effect—collapsed across novel and practiced conditions and assessed by comparing trial 1 (1313 ± 61 ms) with trial 2 (1281 ± 58 ms)—was statistically significant ($F_{(1,22)} = 4.40$, $p = 0.048$). This effect was virtually identical for novel and practiced trials (novel first vs second trial mean, 32.10 ms; practiced first vs second trial mean, 32.79 ms; both separately, $F_{(1,22)} = 4.85$, $p = 0.038$). This suggests that the amount of task interference was the same for novel and practiced tasks. Note that there was no significant difference between trials 2 (1281 ms) and 3 (1286 ms; $F_{(1,22)} = 0.28$, $p = 0.60$, NS).

fMRI results: double dissociations within PFC across novel and practiced task preparation
We hypothesized that the processing order would reverse between aPFC and DLPFC across novel and practiced task prepa-
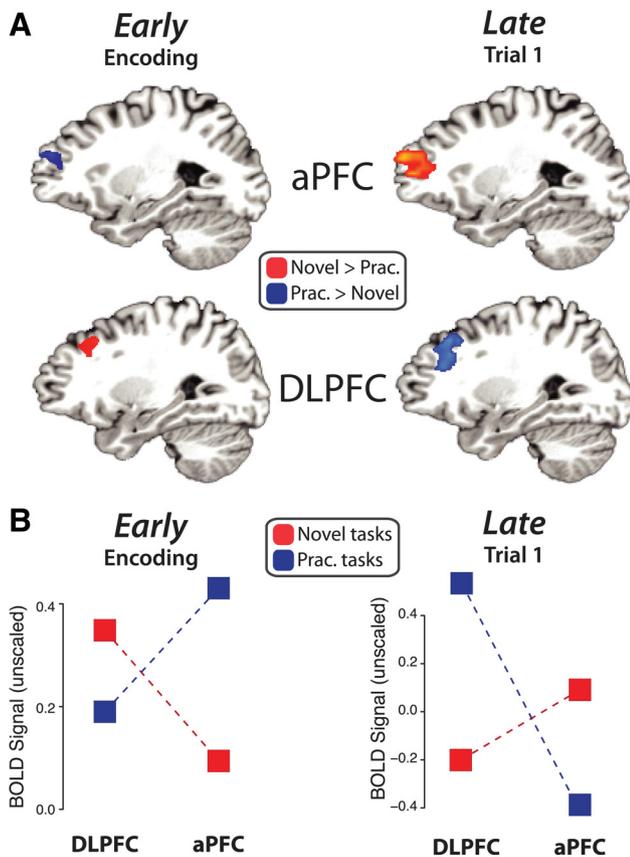
Figure 3. A double dissociation in PFC demonstrates a reversal of processing order. **A**, The activity patterns reversed between novel and practiced task preparation across aPFC and DLPFC. We compared the fMRI activity across novel and practiced tasks during encoding (left) and first trials (right). The encoding activity was identified using a condition × time interaction and verified to have multiple time points significantly different in the indicated directions. The main effect of condition (collapsed across time) was used for the trial activity. **B**, Left, DLPFC was more active early (encoding) for novel tasks, while aPFC was more active early for practiced tasks. Right, In contrast, DLPFC was more active late (trial 1) for practiced tasks, while aPFC was more active late for novel tasks. These patterns are consistent with our hypotheses (see Fig. 1). The double dissociation between aPFC and DLPFC was highly statistically significant ($p = 0.0033$). We ensured the two time periods involved the same regions by plotting the encoding activity using the trial period ROIs and plotting the trial activity using the encoding period ROIs. This pattern of results was replicated using MEG (see supplemental Fig. 1, available at www. jneurosci.org as supplemental material).

ration, reflecting a reversal of preparatory processing stages with practice. We expected the first preparatory stage to occur during encoding (early) and the second preparatory stage to occur during the first trial (late), based on the behavioral results and previous findings of task switching costs during first trials despite extensive preparation time (Rubinstein et al., 2001; Monsell, 2003). Consistent with our predictions, we found a significant practice-induced reversal of these regions' activities across time (region × task type × time period interaction; $F_{(1,14)} = 12.51$, $p = 0.0033$). DLPFC was more active for novel (than practiced) tasks during encoding, but more active for practiced (than novel) tasks during first trials (Fig. 3). In contrast, aPFC was more active for practiced (than novel) tasks during encoding, but more active for novel (than practiced) tasks during first trials.

To avoid circularity (Kriegeskorte et al., 2009), these analyses used encoding activity estimated with the trial period ROIs and the trial activity estimated with the encoding period ROIs (see Materials and Methods for details). Results were similar to the

results reported here when either set of ROIs was used for estimating both encoding and trial period activity.

This dissociation consisted of two double dissociations: an encoding period double dissociation (region × task type interaction; $F_{(1,14)} = 6.45$, $p = 0.02$) and a first trial period double dissociation (region × task type interaction; $F_{(1,14)} = 11.58$, $p = 0.004$). All novel versus practiced mean activity differences were statistically significant ($p < 0.05$), except for mean DLPFC activity during the encoding period. However, DLPFC showed significantly greater novel than practiced activity late in the encoding period (novel > practiced at 14 s poststimulus; $p = 0.00095$).

fMRI results: distinct sets of regions match the reversal patterns of aPFC and DLPFC

Several sensory, motor, and semantic regions were coactive with DLPFC during task preparation (encoding and first trial). These coactive regions were identified in the same way as DLPFC and aPFC (Fig. 4) and included premotor cortex (PMC), M1, A1, S1, posterior parietal cortex (PPC), inferior parietal lobe (IPL), cingulate/supplementary motor area (CMA/SMA), posterior temporal lobe (pTL), and several others (Tables 1, 2). These additional regions likely supported DLPFC during task preparation, since task rule representations are thought to be stored in more posterior regions and maintained in an active state via interactions with DLPFC (Miller and Cohen, 2001; Yeung et al., 2006).

In contrast to the above set of regions coactive with DLPFC, right anterior temporal lobe (aTL) was coactive with aPFC during practiced task encoding (Fig. 4A). Consistent with these regions' involvement in long-term memory (LTM) retrieval, posterior cingulate cortex (PCC), a region strongly implicated in LTM processing (Shannon and Buckner, 2004), was also active during practiced task encoding.

MEG results replicate the major components of the fMRI results

dSPM estimates for the encoding and first trial periods were contrasted across novel and practiced conditions and thresholded at $p < 0.05$ (supplemental Fig. 1, available at www.jneurosci.org as supplemental material). Left aPFC was more active for practiced than novel tasks during the first encoding period. This pattern was present throughout encoding but switched for the first trial, with left aPFC being more active for novel than practiced tasks. In contrast, right DLPFC was more active for novel than practiced tasks during encoding, while it was more active for practiced than novel tasks for the first trial. This set of MEG results replicates the main fMRI findings in PFC.

MEG directed connectivity results: DLPFC to aPFC influences reverse with practice

The higher temporal resolution afforded by MEG was used to more directly test whether aPFC and DLPFC reverse their direction of influence (rather than just their order of processing) across practiced and novel task preparation, as suggested by the fMRI and dSPM MEG results. First trials and the three encoding screens were chosen for analysis based on the likelihood of task preparatory processing during these time periods. We were also especially interested in early time periods because of evidence that long-distance gamma band cortical interactions quickly become equally bidirectional after initial dominant unidirectional information flow (Gregoriou et al., 2009). The same ROIs used in the fMRI analysis (encoding period) were used for these MEG analyses.
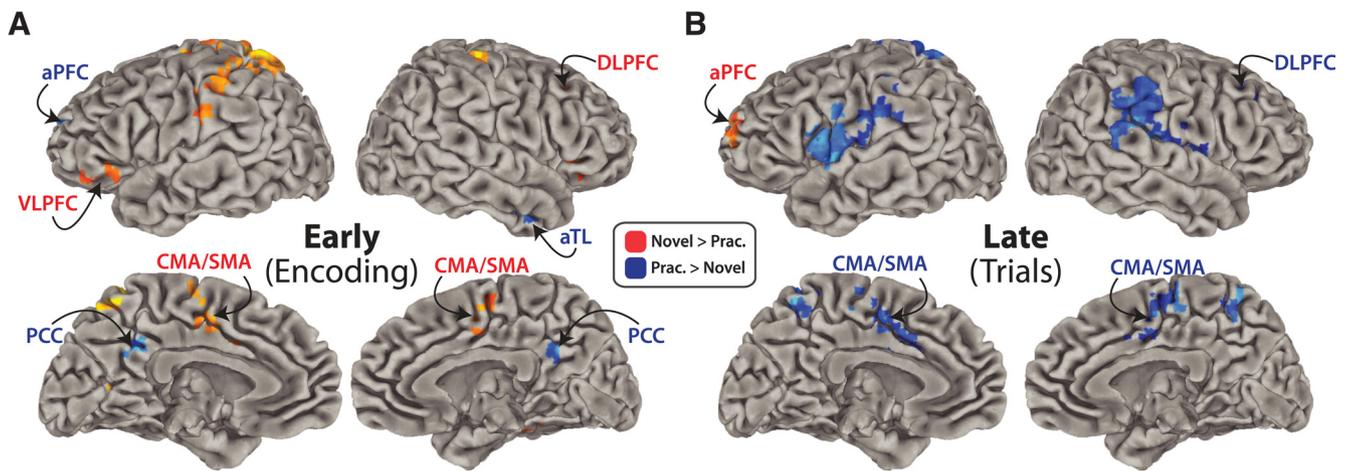
Figure 4. fMRI activity patterns across encoding and trial periods. ***A***, We used a condition × time interaction analysis to identify the brain regions with statistical differences in response shape across conditions. Red regions have statistically significant time points larger for novel tasks, while blue regions have time points larger for practiced tasks. Previous research has found that PCC, lateral temporal lobe (including aTL), and aPFC are involved in LTM retrieval. Their greater involvement in practiced task encoding suggests they are involved in LTM retrieval of practiced task sets. ***B***, Several regions outside PFC reversed between encoding and trial periods. Most of these regions (such as CMA/SMA) matched the DLPFC activity pattern, suggesting they are involved in rule activation. For instance, in the case of CMA/SMA, this activation was likely for the motor rule.

Table 1. Novel versus practiced encoding activity voxel clusters

Cluster labels	Direction of signal	Hemisphere	Voxels (3.2 mm³)	Talairach coordinate x	Talairach coordinate y	Talairach coordinate z	Areas
aPFC	P > N	Left	31	−22.3	48.1	18.5	10
DLPFC	N > P	Right	29	29.6	26.5	35.3	9
IPL	N > P	Left	51	−52	−27.2	25.7	40
PCC	P > N	Both	40	−2	−46.8	29.7	31
M1/S1/S2	N > P	Left	39	−18.3	−29.3	66.7	4, 3, 5
Fusiform/cerebellum	N > P	Right	24	25.7	−37.8	−19.7	20, 36
ACC	N > P	Left	24	−11.5	1.4	28.7	24
M1/S1	N > P	Right	24	41.7	−25.4	58	3, 4
VLPFC	N > P	Left	23	−46.8	29.3	−4.2	47, 45
BG (Medial globus pallidus and putamen)	N > P	Left	16	−17	−2.8		
aTL	P > N	Right	15	51.1	−2.5	−24.8	20, 21
IFJ	N > P	Right	15	27.8	0.5	32	9, 6
PMC	N > P	Right	15	22.9	−13.4	47.5	6
VLPFC	N > P	Right	14	48.9	25.5	−0.7	47, 45
PMC	N > P	Left	14	−32	−15.5	60.5	6
Cuneus	N > P	Left	13	−10.2	−62.6	5.6	18, 19

The contrast was defined by a condition × time (novel vs practiced × 11 time points) interaction ($p < 0.05$, FWE corrected for lateral PFC, FDR corrected for all other voxels). P > N, Practiced significantly greater than novel; N > P, novel significantly greater than practiced; ACC, anterior cingulate cortex; BG, basal ganglia.

Table 2. Novel versus practiced trial activity voxel clusters

Cluster labels	Direction of signal	Hemisphere	Voxels (3.2 mm³)	Talairach coordinate x	Talairach coordinate y	Talairach coordinate z	Areas
aPFC	N > P	Left	148	−29.4	46.6	11.5	10, 11
DLPFC	P > N	Right	113	25.5	28.7	29.7	9
PMC/M1/S1/A1/IPL/pTL	P > N	Left	318	−51.2	−19.9	20.3	44, 6, 4, 3, 13, 22, 41, 43, 40
PPC	P > N	Both	240	−8	−52.2	58	7
PMC/M1/S1/A1/IPL/pTL	P > N	Right	228	57.9	−24.4	21.4	40, 6, 22, 41, 42, 21
SMA/CMA	P > N	Both	204	−0.9	−7.6	47.9	6, 24

The contrast was defined by novel trials versus practiced trials ($p < 0.05$, FWE corrected). P > N, Practiced significantly greater than novel; N > P, novel significantly greater than practiced.

We found using GC and PSI that DLPFC dominantly influenced aPFC during novel task preparation (first encoding screen), as expected (Fig. 5*A*, *B*, Table 3, 4). This influence was reversed during practiced task preparation (second encoding screen, possibly reflecting additional time necessary for LTM retrieval of the practiced task set), again as predicted. These dominant influences significantly differed between conditions (novel greater than practiced) for DLPFC to aPFC in the GC gamma band analysis (first encoding screen; $t_{(412)} = 2.34$, $p = 0.0197$).

Analysis of first trials replicated the encoding period results (Fig. 5*C* and Table 4). Using PSI, we found a significant dominant influence from DLPFC to aPFC for novel first trials ($t = 2.4$, $p = 0.019$), with a significant reversal (novel vs practiced; $t = 3.3$, $p = 0.001$) to aPFC dominantly influencing DLPFC for practiced first trials ($t = 2.3$, $p = 0.023$). These effects were not present during second trials, and the DLPFC-to-aPFC influence was significantly higher for first than second trials ($t = 2.6$, $p = 0.01$). These first trial effects were not present in the gamma band GC results, possibly because of interactions occurring outside gamma band frequencies.

A GC of gamma power
Encoding
1 (novel) and 2 (prac.)

aPFC
$t=2.73$
$p=.007$
$t=2.38$
$p=.018$
DLPFC

B PSI of wide band
Encoding
1 (novel) and 2 (prac.)

aPFC
$t=2.2$
$p=.027$
$t=2.8$
$p=.006$
DLPFC

C PSI of wide band
Trial 1
(both stimuli)

aPFC
$t=2.4$
$p=.019$
$t=2.3$
$p=.023$
DLPFC

Dominant influences
→ Practiced   → Novel

D **Practiced Tasks**
Mean encoding activity,
gamma power (40-150 hz)

DLPFC (z-normalized) vs aPFC (z-normalized)
$r = .32, p = 2e{-}10$

E **Novel Tasks**
Mean encoding activity,
gamma power (40-150 hz)

DLPFC (z-normalized) vs aPFC (z-normalized)
$r = .35, p = 1e{-}12$

Figure 5. MEG directed connectivity confirms the reversal of DLPFC and aPFC information flow. GC and PSI were used to estimate the directed influences between the DLPFC and aPFC regions (as defined by the fMRI experiment) during task encoding. Note that these directed connectivity methods are correlational in nature and are not able to assess true causality (though modulation of these measures by experimental manipulation can strengthen that inference). **A**, GC of gamma power indicated that aPFC and DLPFC influence each other bidirectionally during encoding (largest $p = 6e{-}37$). There was a significant dominant influence from DLPFC to aPFC during the first novel task encoding screen and a significant dominant influence from aPFC to DLPFC during the second practiced task encoding screen. This supports the prediction of bottom-up information flow during novel task preparation and top-down information flow during practiced task preparation. See Tables 3 and 4 for details. Practiced task interactions may be delayed relative to novel task interactions because of the additional time necessary for retrieval of higher-level task information from LTM. **B**, The PSI statistic, an alternative to GC with fewer false positives for noisy data (Nolte et al., 2008), supports the GC finding. **C**, This reversal of information flow was also present during first trials and was significantly different between novel and practiced task preparation ($p = 0.001$). **D**, **E**, The gamma power estimates used for GC analysis (averaged by block) were significantly correlated between aPFC and DLPFC, indicating strong functional connectivity and supporting the conclusions based on GC and PSI.

Supporting the directed connectivity results, we found significant correlations between aPFC and DLPFC mean encoding activities (based on nonzero lag mean gamma power). The regions correlated at $r = 0.32$ ($p = 2e{-}10$) for practiced task preparation and $r = 0.35$ ($p = 1e{-}12$) for novel task preparation (Fig. 5D,E).

Discussion

Previous research suggests individual task rule representations are activated during task preparation along with a higher-level goal representation that integrates and coordinates those rules during task performance. We hypothesized that the order of these two processes would reverse between novel and practiced task preparation, with novel preparation involving activation of individual rules before their integration by a higher-level task representation, and practiced preparation involving cued recall of a

higher-level task representation that then activates individual rules from memory. Accordingly, we predicted a reversal of information flow between DLPFC (a lower-level control region) and aPFC (a higher-level control region) in the transition from novel to practiced task preparation. This prediction was confirmed by an fMRI double dissociation of aPFC and DLPFC across novel and practiced task preparation (Fig. 3) and further supported by a reversal in MEG directed connectivity between aPFC and DLPFC (Fig. 5).

The present study investigated two ways in which the human brain is rapidly reconfigured for task performance. Importantly, novel task preparation can be characterized in terms of RITL, a powerful form of learning largely unique to humans. We show here that humans can rapidly (i.e., within 5–10 s) learn complex novel tasks and perform accurately with no practice (91% on first trials). Other animals typically use slower forms of learning such as operant conditioning, which can take days or months (especially for abstract or complex tasks). Some nonhuman primates can perform RITL using imitation, but imitation is typically ineffective for abstract and complex tasks like those used here given the many examples necessary to specify the appropriate task rules. Compatible with the present study's neural localization of this human ability, aPFC and DLPFC are among a select number of regions that grew substantially since our common ancestor with chimpanzees (Semendeferi et al., 2001; Avants et al., 2006), suggesting that development of these (and related) regions first allowed RITL in early humans. These regions likely provide humans with the enhanced WM integration and subgoaling skills (Braver and Bongiolatti, 2002; De Pisapia et al., 2007) necessary for rapidly integrating instructions into coherent task sets during RITL.

Previous studies have suggested that aPFC specifies task sets in DLPFC and more posterior regions during practiced task preparation (Sakai and Passingham, 2003, 2006). However, these studies did not demonstrate the direction of information flow between these regions, leaving open the possibility that posterior regions actually drive activity in anterior regions during task preparation. Given that these studies all used practiced tasks and given that practiced task preparation is likely a top-down process involving retrieval of task sets from LTM (Mayr and Kliegl, 2003), we expected (and demonstrated) that aPFC indeed drives activity in DLPFC during practiced task preparation (aPFC to DLPFC). This is in sharp contrast to novel task preparation, which likely involves a bottom-up series of transformations from instruction stimuli to rule representations to an integrated task representation. Compatible with a shift from bottom-up to top-down processing, we show here that DLPFC originally drives aPFC activity

(DLPFC to aPFC) during RITL, but the direction of information flow reverses with practice (as task sets become familiar enough to be retrieved from memory).

This reversal of information flow is compatible with—indeed, it was predicted based on—theories of an anterior-to-posterior hierarchy in PFC (Koechlin et al., 2003; Badre, 2008). The present set of results suggests the need to consolidate seemingly opposing theories of the hierarchical organization of PFC. Rather than supporting just a representational PFC hierarchy [involving increasing relational complexity (Bunge et al., 2005; De Pisapia et al., 2007) and conceptual abstraction (Badre and D'Esposito, 2007; Christoff et al., 2009)] or a control PFC hierarchy [involving actions/procedures in a temporal or goal hierarchy (Koechlin et al., 1999; Fuster, 2001; Botvinick, 2008)], the data support each one in turn. The aPFC-to-DLPFC information flow is most compatible with a control PFC hierarchy in which higher-level regions influence lower-level regions to control/coordinate their activities. In contrast, the DLPFC-to-aPFC information flow is most compatible with a representational PFC hierarchy in which lower-level regions feed information upward to be integrated into higher-level representations. However, it may be that the PFC hierarchy functions by representing higher-order relationships at each hierarchical stage, with the activation of these representations controlling lower-level stages via feedback connections. According to this interpretation, the DLPFC-to-aPFC activation during novel task preparation forms a higher-level task set representation in aPFC (which can be later retrieved from LTM during practiced task preparation) that controls/coordinates the rule representations in DLPFC for subsequent task performance. Given that there are likely other plausible interpretations of the present results, however, further research is necessary to verify this interpretation.

In addition to a hierarchical account of PFC, the present results are also compatible with theories suggesting DLPFC controls activity in posterior regions in preparation for task performance (Miller and Cohen, 2001; Yeung et al., 2006). Most posterior regions were coactive with DLPFC (rather than aPFC), suggesting that DLPFC encodes rule representations in WM and actively maintains them via interactions with these regions. Some of the posterior regions included portions of auditory, visual, motor, and somatosensory cortices, which is compatible with recent evidence that even when considered abstractly (i.e., with words and in imagined situations), semantic rules activate the same brain areas that would be involved in processing that information if it were directly experienced (e.g., visual cortex for color rules, auditory cortex for sound rules, SMA for motor rules, etc.) (Goldberg et al., 2006). Other posterior regions that coactivate with DLPFC have been implicated in logic rule representation [in addition to DLPFC itself (Miller et al., 2002)], including pTL (Bunge et al., 2003; Donohue et al., 2005), PMC (Wallis and Miller, 2003), and PPC (Stoet and Snyder, 2004). Note that ven-

Table 3. DLPFC↔aPFC: GC directional connectivity estimates

	GC (gamma power)	
	aPFC→DLPFC	DLPFC→aPFC
Encoding stimulus 1 (1–1000 ms)		
Practiced	$d = 0.77, t = 18/p = 4.6e{-}55$	$d = 0.81, t = 18/p = 2e{-}55$
Novel	$d = 0.67, t = 16/p = 6e{-}45$	$d = 0.88\ (*D), t = 21/p = 2.8e{-}64$
Encoding stimulus 2 (1–1000 ms)		
Practiced	$d = 0.80\ (*D), t = 19/p = 1.5e{-}57$	$d = 0.59, t = 14/p = 6e{-}37$
Novel	$d = 0.77, t = 18/p = 1.9e{-}54$	$d = 0.82, t = 18/p = 1.9e{-}53$
Encoding stimulus 3 (1–1000 ms)		
Practiced	$d = 0.81, t = 19/p = 2e{-}56$	$d = 0.77, t = 17/p = 2e{-}50$
Novel	$d = 0.80, t = 18/p = 8e{-}55$	$d = 0.76, t = 17/p = 6.7e{-}51$
Trial 1 (1–2000 ms)		
Practiced	$d = 0.45, t = 8/p = 8e{-}15$	$d = 0.46, t = 7/p = 6e{-}12$
Novel	$d = 0.47, t = 7/p = 1e{-}10$	$d = 0.42, t = 6/p = 2e{-}9$

Note that unlike PSI, the GC results are of gamma band activity only and therefore may have filtered out some effects in other frequency bands (such as during trial 1). *d*, Effect size (Cohen's *d*) of Granger causality magnitudes; *t*, *t* value of magnitudes versus 0; *p*, *p* value of the *t* test versus 0; *D, Significant dominant influence (A→B vs B→A).

Table 4. DLPFC↔aPFC: PSI directional connectivity estimates

	PSI (dominant direction only)	
	aPFC→DLPFC	DLPFC→aPFC
Encoding stimulus 1 (1–1000 ms)		
Practiced		$d = 0.04, t = 0.62/p = 0.53$
Novel		$d = 0.06\ (*D), t = 2.2/p = 0.027$
Encoding stimulus 2 (1–1000 ms)		
Practiced	$d = 0.14\ (*D), t = 2.8/p = 0.006$	
Novel	$d = 0.06, t = 1.4/p = 0.17$	
Encoding stimulus 3 (1–1000 ms)		
Practiced	$d = 0.0008, t = 0.15/p = 0.88$	
Novel		$d = 0.065, t = 1.4/p = 0.15$
Trial 1 (1–2000 ms)		
Practiced	$d = 0.10\ (*C)\ (*D), t = 2.3/p = 0.023$	
Novel		$d = 0.09\ (*C)\ (*D), t = 2.4/p = 0.019$

d, Effect size (Cohen's *d*) of PSI magnitudes; *t*, *t* value of magnitudes versus 0; *p*, *p* value of the *t* test versus 0; *C, significant condition (novel versus practiced) effect; *D, Significant dominant influence (A→B versus B→A).

trolateral PFC (VLPFC) was also involved in novel task encoding, compatible with evidence that it is involved in LTM retrieval of individual rules (as opposed to entire task sets) (Donohue et al., 2005). This set of results, in conjunction with previous findings, suggests that individual rule representations in posterior regions are prepared for task performance via interactions with lateral PFC.

We have emphasized task preparation during first trials, although this process might be better characterized as initial task execution. Importantly, it has been suggested that initial task execution involves preparatory processes that require the presence of task stimuli to complete (Rubinstein et al., 2001; Monsell, 2003). Supporting this view, we found that first trials were slower than second trials (while second and third trials showed no differences), suggesting an additional preparatory process might be present during first trials only. Also supporting this conclusion, we found with MEG that the DLPFC-to-aPFC influence during novel task encoding remained during first trials (suggesting a continuation of preparatory processes) but not second trials. However, if these first trial effects were caused by some process other than task preparation, the present results would need to be reinterpreted in terms of task preparation versus task execution (rather than two preparatory processes). This would entail a shift from DLPFC during task preparation to aPFC during task implementation for novel tasks and from aPFC during task preparation to DLPFC during task implementation for practiced tasks, suggesting a shift in the location of ongoing task control with practice. Further research is necessary to assess the extent to which

first trial effects that differentiate between novel and practiced tasks are specific to task execution rather than task preparation.

Although the hemispheric lateralization effect in PFC was useful for accurate MEG source localization, it was not expected a priori. It is possible that right DLPFC was involved because of the need to represent sensory semantic categories [a process associated with right-lateralized function (Seger et al., 2000)] in the particular task paradigm used here. Left aPFC may have been involved because of the need to perform WM integration during novel task preparation, as WM integration has been better associated with left than right aPFC (De Pisapia et al., 2007). This suggests that the hemispheric effect in aPFC would replicate in a different experimental context, while a different set of task rules would possibly involve left DLPFC instead of (or in addition to) right DLPFC. This conclusion is also supported by findings indicating that the same aPFC region interacts with different posterior PFC regions depending on the particular task being prepared for (Sakai and Passingham, 2003).

Implicit in task preparation is the need to resolve interference from previously used task sets. Some of the preparatory brain activity certainly reflected this interference resolution process. Importantly, however, the results indicate that the level of interference did not differ between novel and practiced tasks, suggesting that the observed results do not reflect interference-related processes. Evidence for this comes from two sources. First, assuming there was additional interference for one condition (and assuming PFC resolves that interference), then that condition should always involve greater PFC activity than the other; yet, we observed a reversal in which both conditions involved substantial PFC activity. Second, the task switching costs [which partially reflect task interference (Wylie and Allport, 2000)] were virtually identical (within 1 ms) for novel and practiced tasks, strongly suggesting that any interference from previous tasks was equated across conditions.

It was perhaps surprising that there was relatively little performance benefit from practice (just 2% accuracy, 23 ms RT). One possible explanation is that excessive time to prepare (5–10 s) and respond (average 1295 ms) allowed a lazy preparation strategy that reduced differences between conditions. Alternatively, the human brain may be highly effective at transferring rules to new task contexts during RITL (despite potentially massive rule interference) to the point where performance is little affected by novel rule combinations. Singley and Anderson (1989; chapter 3) demonstrated this ability using complex text editing tasks, showing that transfer of rules practiced in one task improves performance of related novel tasks beyond any performance decrement from interference. This benefit of practiced rule transfer may explain subjects' minimal rule interference as well as their high performance on novel tasks here.

We have shown a reversal in processing order and information flow between aPFC and DLPFC that suggests a fundamental difference between novel and practiced task preparation processes. Novel task preparation likely involves activation of individual rules in posterior regions and DLPFC before those rules are integrated into a unified task set by aPFC for coordinated task performance. In contrast, practiced task preparation likely involves retrieval of a higher-level task set representation from LTM that is loaded into aPFC for reconfiguration of DLPFC and posterior regions representing individual rules. This shift in dynamics between novel and practiced task preparation illustrates that there are two ways in which humans are able to rapidly reconfigure their minds via instruction. Importantly, the novel task preparation process begins to explain how we are able to rapidly learn a virtually infinite variety of possible tasks (i.e., RITL), allowing our species to efficiently adapt to the many unique situations and new technologies of an ever-changing world.

References

Andersson J (2005) Testing for Granger causality in the presence of measurement errors. Econ Bull 3:1–13.

Avants BB, Schoenemann PT, Gee JC (2006) Lagrangian frame diffeomorphic image registration: morphometric comparison of human and chimpanzee cortex. Med Image Anal 10:397–412.

Badre D (2008) Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. Trends Cogn Sci 12:193–200.

Badre D, D'Esposito M (2007) Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. J Cogn Neurosci 19:2082–2099.

Botvinick MM (2008) Hierarchical models of behavior and prefrontal function. Trends Cogn Sci 12:201–208.

Braver TS, Barch DM (2006) Extracting core components of cognitive control. Trends Cogn Sci 10:529–532.

Braver TS, Bongiolatti SR (2002) The role of frontopolar cortex in subgoal processing during working memory. Neuroimage 15:523–536.

Bunge SA, Kahn I, Wallis JD, Miller EK, Wagner AD (2003) Neural circuits subserving the retrieval and maintenance of abstract rules. J Neurophysiol 90:3419–3428.

Bunge SA, Wendelken C, Badre D, Wagner AD (2005) Analogical reasoning and prefrontal cortex: evidence for separable retrieval and integration. Cereb Cortex 15:239–249.

Burgess P (1997) Theory and methodology in executive function research. In: Methodology of frontal and executive function (Rabbit P, ed), pp 81–116. Hove, UK: Psychology.

Christoff K, Keramatian K, Gordon AM, Smith R, Mädler B (2009) Prefrontal organization of cognitive control according to levels of abstraction. Brain Res 1286:94–105.

Chumbley JR, Friston KJ (2009) False discovery rate revisited: FDR and topological inference using Gaussian random fields. Neuroimage 44:62–70.

Cox RW (1996) AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res 29:162–173.

Daffner KR, Mesulam MM, Scinto LF, Acar D, Calvo V, Faust R, Chabrerie A, Kennedy B, Holcomb P (2000) The central role of the prefrontal cortex in directing attention to novel events. Brain 123:927–939.

Dale AM, Liu AK, Fischl BR, Buckner RL, Belliveau JW, Lewine JD, Halgren E (2000) Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. Neuron 26:55–67.

De Pisapia N, Slomski JA, Braver TS (2007) Functional specializations in lateral prefrontal cortex associated with the integration and segregation of information in working memory. Cereb Cortex 17:993–1006.

Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, Albert MS, Killiany RJ (2006) An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 31:968–980.

Donohue SE, Wendelken C, Crone EA, Bunge SA (2005) Retrieving rules for behavior from long-term memory. Neuroimage 26:1140–1149.

Dove A, Pollmann S, Schubert T, Wiggins CJ, von Cramon DY (2000) Prefrontal cortex activation in task switching: an event-related fMRI study. Brain Res Cogn Brain Res 9:103–109.

Dumontheil I, Thompson R, Duncan J (2010) Assembly and use of new task rules in fronto-parietal Cortex. J Cogn Neurosci. Advanced online publication. Retrieved May 10, 2010. doi: 10.1162/jocn.2010.21439.

Fischl B, Sereno MI, Tootell RB, Dale AM (1999) High-resolution intersubject averaging and a coordinate system for the cortical surface. Hum Brain Mapp 8:272–284.

Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995) Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. Magn Reson Med 33:636–647.

Fries P, Nikolić D, Singer W (2007) The gamma cycle. Trends Neurosci 30:309–316.

Fuster JM (2001) The prefrontal cortex–an update: time is of the essence. Neuron 30:319–333.

Genovese CR, Lazar NA, Nichols T (2002) Thresholding of statistical maps in functional neuroimaging using the false discovery rate. Neuroimage 15:870–878.

Geweke J (1982) Measurement of linear dependence and feedback between multiple time series. J Am Stat Assoc 77:304–313.

Goldberg RF, Perfetti CA, Schneider W (2006) Perceptual knowledge retrieval activates sensory brain regions. J Neurosci 26:4917–4921.

Gregoriou GG, Gotts SJ, Zhou H, Desimone R (2009) High-frequency, long-range coupling between prefrontal and visual cortex during attention. Science 324:1207–1210.

Griswold MA, Jakob PM, Heidemann RM, Nittka M, Jellus V, Wang J, Kiefer B, Haase A (2002) Generalized autocalibrating partially parallel acquisitions (GRAPPA). Magn Reson Med 47:1202–1210.

Hämäläinen MS, Ilmoniemi RJ (1994) Interpreting magnetic fields of the brain: minimum norm estimates. Med Biol Eng Comput 32:35–42.

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

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

Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI (2009) Circular analysis in systems neuroscience: the dangers of double dipping. Nat Neurosci 12:535–540.

Lin FH, Belliveau JW, Dale AM, Hämäläinen MS (2006a) Distributed current estimates using cortical orientation constraints. Hum Brain Mapp 27:1–13.

Lin FH, Witzel T, Ahlfors SP, Stufflebeam SM, Belliveau JW, Hämäläinen MS (2006b) Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates. Neuroimage 31:160–171.

Luria AR (1973) The frontal lobes and the regulation of behavior. In: Psychophysiology of the frontal lobes (Pribram KH, Luria AR, eds), pp 3–28. New York: Academic.

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

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

Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. Annu Rev Neurosci 24:167–202.

Miller EK, Freedman DJ, Wallis JD (2002) The prefrontal cortex: categories, concepts and cognition. Philos Trans R Soc Lond B Biol Sci 357:1123–1136.

Monsell S (1996) Control of mental processes. In: Unsolved mysteries of the mind: tutorial essays in cognition (Bruce V, ed), pp 93–148. Hove, UK: Erlbaum.

Monsell S (2003) Task switching. Trends Cogn Sci 7:134–140.

Niessing J, Ebisch B, Schmidt KE, Niessing M, Singer W, Galuske RA (2005) Hemodynamic signals correlate tightly with synchronized gamma oscillations. Science 309:948–951.

Nolte G, Bai O, Wheaton L, Mari Z, Vorbach S, Hallett M (2004) Identifying true brain interaction from EEG data using the imaginary part of coherency. Clin Neurophysiol 115:2292–2307.

Nolte G, Ziehe A, Nikulin VV, Schlögl A, Krämer N, Brismar T, Müller K-R (2008) Robustly estimating the flow direction of information in complex physical systems. Phys Rev Lett 100:1–4.

Rubinstein JS, Meyer DE, Evans JE (2001) Executive control of cognitive processes in task switching. J Exp Psychol Hum Percept Perform 27:763–797.

Ruge H, Wolfensteller U (2010) Rapid formation of pragmatic rule representations in the human brain during instruction-based learning. Cereb Cortex 20:1656–1667.

Sakai K (2008) Task set and prefrontal cortex. Annu Rev Neurosci 31:219–245.

Sakai K, Passingham RE (2003) Prefrontal interactions reflect future task operations. Nat Neurosci 6:75–81.

Sakai K, Passingham RE (2006) Prefrontal set activity predicts rule-specific neural processing during subsequent cognitive performance. J Neurosci 26:1211–1218.

Schneider W, Eschman A, Zuccolotto A (2002) E-Prime: user's guide. Pittsburgh: Psychology Software Tools.

Schoffelen JM, Gross J (2009) Source connectivity analysis with MEG and EEG. Hum Brain Mapp 30:1857–1865.

Seger CA, Desmond JE, Glover GH, Gabrieli JD (2000) Functional magnetic resonance imaging evidence for right-hemisphere involvement in processing unusual semantic relationships. Neuropsychology 14:361–369.

Semendeferi K, Armstrong E, Schleicher A, Zilles K, Van Hoesen GW (2001) Prefrontal cortex in humans and apes: a comparative study of area 10. Am J Phys Anthropol 114:224–241.

Seth AK (2005) Causal connectivity of evolved neural networks during behavior. Network 16:35–54.

Seth AK (2010) A MATLAB toolbox for Granger causal connectivity analysis. J Neurosci Methods 186:262–273.

Shannon BJ, Buckner RL (2004) Functional-anatomic correlates of memory retrieval that suggest nontraditional processing roles for multiple distinct regions within posterior parietal cortex. J Neurosci 24:10084–10092.

Singley MK, Anderson JR (1989) The transfer of cognitive skill. Cambridge, MA: Harvard UP.

Singley MK, Anderson JR (1989) The transfer of cognitive skill. Cambridge, MA: Harvard UP.

Stoet G, Snyder LH (2004) Single neurons in posterior parietal cortex of monkeys encode cognitive set. Neuron 42:1003–1012.

Talairach J, Tournoux P (1988) Co-planar stereotaxic atlas of the human brain. New York: Thieme.

Taulu S, Simola J, Kajola M (2004) MEG recordings of DC fields using the signal space separation method (SSS). Neurol Clin Neurophysiol 2004:35.

Taulu S, Simola J, Kajola M (2005) Applications of the signal space separation method. IEEE Trans Sign Proc 53:3359–3372.

Wallis JD, Miller EK (2003) From rule to response: neuronal processes in the premotor and prefrontal cortex. J Neurophysiol 90:1790–1806.

Wallstrom GL, Kass RE, Miller A, Cohn JF, Fox NA (2004) Automatic correction of ocular artifacts in the EEG: a comparison of regression-based and component-based methods. Int J Psychophysiol 53:105–119.

Wylie G, Allport A (2000) Task switching and the measurement of "switch costs". Psychol Res 63:212–233.

Yeung N, Nystrom LE, Aronson JA, Cohen JD (2006) Between-task competition and cognitive control in task switching. J Neurosci 26:1429–1438.

Zumer JM, Brookes MJ, Stevenson CM, Francis ST, Morris PG (2010) Relating BOLD fMRI and neural oscillations through convolution and optimal linear weighting. Neuroimage 49:1479–1489.