Cerebral Cortex Advance Access published October 3, 2013

Cerebral Cortex
doi:10.1093/cercor/bht262

He Who Is Well Prepared Has Half Won The Battle: An fMRI Study of Task Preparation

Anna Manelis[1,2] and Lynne M. Reder[1,2]

[1]Department of Psychology, Carnegie Mellon University, Pittsburgh, PA, USA and [2]The Center for the Neural Basis of Cognition, Pittsburgh, PA 15213, USA

Address correspondence to Anna Manelis, Department of Psychology, Carnegie Mellon University, Baker Hall 342c, Pittsburgh, PA 15213, USA. E-mail: anna.manelis@gmail.com

The neural mechanism underlying preparation for tasks that vary in difficulty has not been explored. This functional magnetic resonance imaging study manipulated task difficulty by varying the working memory (WM) load of the *n*-back task. Each *n*-back task block was preceded by a preparation period involving a screen that indicated the level of difficulty of the upcoming task. Consistent with previous work, activation in some brain regions depended on WM load in the task. These regions were used as regions of interest for the univariate and multivariate (classification) analyses of preparation periods. The findings were that the patterns of brain activation during task preparation contain information about the upcoming task difficulty. (1) A support vector machine classifier was able to decode the *n*-back task difficulty from the patterns of brain activation during task preparation. Those individuals whose activation patterns for anticipated 1- versus 2- versus 3-back conditions were classified with higher accuracy showed better behavioral performance on the task, suggesting that task performance depends on task preparation. (2) Left inferior frontal gyrus, intraparietal sulcus, and anterior cingulate cortex parametrically decreased activation as anticipated task difficulty increased. Taken together, these results suggest dynamic involvement of the WM network not only during WM task performance, but also during task preparation.

Keywords: fMRI, multivoxel pattern classification, *n*-back task, task preparation, working memory

Introduction

Working memory (WM) is a limited capacity system involved in the maintenance and online manipulation of information in the mind (Baddeley and Hitch 1974; Baddeley 2010). WM load increases with the amount of information that an individual needs to process at one time. Neuroimaging studies consistently show that activation in the network of brain regions that include prefrontal, anterior cingulate, posterior parietal cortices, and some striatal regions (see Owen et al. 2005; Rottschy et al. 2012 for reviews) is modulated by WM load. In this study, we refer to these regions as the WM network. Damage to neural substrates of the WM network leads to impairment in learning, reasoning, and decision-making (e.g., Mishkin and Manning 1978; Petrides and Milner 1982; Curtis and D'Esposito 2003; Müller and Knight 2006).

Given that the defining characteristics of the WM system refer to its engagement during on-line processing of information, the functionality of the WM network has been investigated exclusively during task performance. However, both in real-life situations and when participating in experiments, people not only perform tasks but also spend time preparing to execute them. Until now, no study examined how activation in the WM network during task preparation is modulated by anticipated task difficulty or whether there is a relationship between this preparatory activation and subsequent behavior.

Previous behavioral studies suggest that, during task preparation, both attentional and WM resources are activated, which, in turn, facilitate task performance. For example, the presentation of a valid cue informing subjects about the type of the upcoming trial in Stroop tasks facilitates task performance and significantly reduces subjects' response times (RTs), compared with trials when either an invalid or no cue is presented (e.g., Logan and Zbrodoff 1982; Aarts et al. 2008). The neuroimaging studies of task preparation have demonstrated that brain activations preceding encoding predict encoding success (Adcock et al. 2006; Mackiewicz et al. 2006; Otten et al. 2006; Guderian et al. 2009; Park and Rugg 2010), subject's intentions (e.g., Haynes et al. 2007; Gallivan et al. 2011; Gilbert 2011), a task that follows the cue in a task-switching paradigm (e.g., Brass and von Cramon 2002), or the type of stimulus a subject expects to process (e.g., Sakai and Passingham 2003).

Based on these previous findings and also on the fact that brain activation in the WM network is modulated by task difficulty, we hypothesized that preparation for a WM task relies on functioning of the WM network. Moreover, brain activation in the WM network during task preparation may depend on the expected level of task difficulty. Given that only a small number of previous studies examined this question, the direction of such modulation is difficult to predict. One recent magnetoencephalographic (MEG) study reported greater event-related synchronization (ERS) in frontal, parietal, and temporal regions during preparation for a WM task, compared with preparation for a task that did not involve WM (Altamura et al. 2010). ERS is thought to correspond to decreases in brain activation (e.g., Pfurtscheller 2001; Neuper et al. 2006), suggesting that the expectation of greater cognitive demands may reduce activation in the brain regions that usually increase activation during WM task performance. On the other hand, some studies suggest that brain activation during task preparation resembles what is observed during task performance. For example, while the superior parietal lobule (a region involved in spatial processing) activates strongly during preparation for spatial compared with verbal tasks, Broca's and Wernicke's areas activate more strongly before verbal, relative to the spatial, tasks (Sakai and Passingham 2003). These findings suggest that the brain regions that increase in activation when WM load increases during task performance will also increase in activation (or preactivate) during task preparation.

We examined the WM network activation during task preparation as a function of expected WM load in the fMRI study using the *n*-back task (Braver et al. 1997; Cohen et al. 1997). Performance on the *n*-back task involves multiple cognitive processes necessary to maintain a set of items in memory, rapidly update the identity and sequential order of items when a new item is presented, and to match a current item with one presented *n* items prior (e.g., Chatham et al. 2011). The difficulty of the *n*-back task increases as the value of *n* increases. To make the examination of preparatory activations possible, we presented blocks of *n*-back in a completely random order. Each block of *n*-back was preceded by an 8-s instruction period (called a preparatory period) during which an instruction screen indicated the upcoming level of difficulty ("1-back," "2-back," or "3-back"). Along with the conventional univariate methods of fMRI data analysis (i.e., general linear model (GLM)), we used a multivoxel pattern classification analysis (MVPA) to examine whether anticipated WM load during task performance can be decoded from the patterns of brain activation during task preparation. While the GLM analysis assesses average brain activation, the MVPA can uncover the signal components independent of the average response to the voxels of interest (e.g., Mur et al. 2009; Kragel et al. 2012). We also examined whether brain activation during task preparation could predict subjects' behavioral performance (i.e., RT and accuracy) at task.

Materials and Methods

Subjects

Nineteen native English speakers (mean age = 24, right-handed, 11 female) participated in this study for $65 compensation. Three subjects were excluded from the study due to excessive motion. All subjects were treated in accordance with the CMU IRB guidelines.

Design and Procedure

The *n*-back task requires that a subject indicates whether the current stimulus is the same as one that appeared a specific number of presentations prior (Braver et al. 1997; Cohen et al. 1997). For example, in the 2-back condition, the current stimulus is a target when it has the same identity as one that appeared 2 presentations prior. All other stimuli that do not match the stimulus shown 2 presentations before are considered nontargets. Figure 1*A* illustrates the 2-back condition of the *n*-back task. The greater the value of *n* (the number of positions back in time), the more challenging the *n*-back task, because there are more stimuli to maintain and update. In our study, subjects performed 1-, 2-, and 3-back tasks. Subjects practiced each of these tasks with single digits (1–9) as stimuli (3 blocks of 12 trials for each of the 3 levels of *n*-back) prior to entering the scanner.

In the scanner, the subjects performed 30 blocks of the *n*-back task, with 12 trials per block using words as stimuli, with 10 blocks for each of the 3 levels of *n*-back. The order of blocks of 1-, 2-, and 3-back was randomly determined for each subject as was the assignment of words to positions within a block. Blocks were separated by 10–12 s of rest. Additionally, each block was preceded with an 8-s instruction screen indicating the upcoming *n*-back condition (Fig. 1*A*). The instruction screen displayed "1-back," "2-back,"or "3-back." A behavioral pilot study indicated that randomized presentation of *n*-back blocks made it more difficult for subjects to remember the current *n*-back condition (e.g., subjects may respond as if they are in the 1-back condition, when, in fact, they are in the 3-back condition). To ameliorate this problem, we assigned different background color and font combinations for each level of *n*-back as well as giving explicit instructions prior to each block. The background/font color schemes were randomly assigned to each *n*-back condition for each subject. The practice trials outside the scanner familiarized subjects with the instruction screens and the color/font assignments to the level of *n*-back.

The stimuli were 180 nouns selected from the MRC Psycholinguistic Database. The words were between 4 and 7 letters and consisted of no more than 2 syllables. Six different words were selected for each block of trials and not repeated in any other block. Each word was repeated exactly once within its block, either in the target or nontarget position as defined by the level of *n*-back for that block. To discourage perceptual matching of words, words were repeated in the opposite case (lower or upper) than the first presentation. Following Braver et al. (2001), one-third of all trials were targets. Each word was displayed until the subject responded with a maximum duration of 4 s. There was a jittered interval of 2–8 s between words. Subjects responded to targets with one hand and to nontargets with the other. Hand assignment was counterbalanced across subjects.

fMRI Acquisition

The fMRI experiment was conducted using a Siemens 3T Verio MR system. A high-resolution structural image (0.8 × 0.8 × 0.8 mm) was acquired using magnetization-prepared rapid acquisition with gradient echo (MPRAGE) [time repetition (TR) = 1800 ms, time echo (TE) = 2.22 ms,

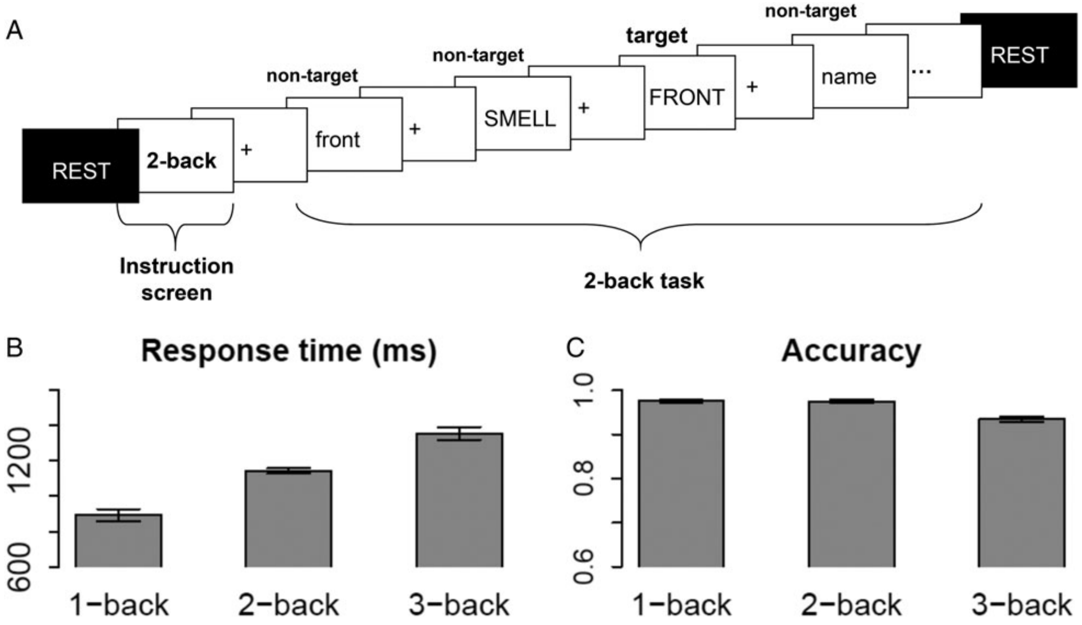

Figure 1. (*A*) Example of a partial run in the 2-back version of the *n*-back task. (*B* and *C*). Changes in RT and accuracy as a function of WM load across all subjects. Error bars refer to a ±1 standard error.

field of view (FOV) = 205, flip angle (FA) = 9°, number of slices = 256]. Functional data were collected using a gradient-echo, echo-planar sequence (TR = 2000 ms, TE = 30 ms, field of view (FOV) = 205, flip angle (FA) = 79°, 36 slices, 3.2 × 3.2 × 3.2 mm). Field maps were collected with the same resolution as the blood oxygen level-dependent (BOLD) images using a gradient-echo sequence (TR = 394 ms, FA = 60°, TE = 5.1 and 7.56 ms).

Our pilot study showed that when trials are presented at a fast pace, subjects tend to skip responses, resulting in a number of missed trials within a block. Because we planned to analyze blocks of trials (not separate events), we tried to reduce the number of missed trials within a block by presenting stimuli in a self-paced manner with the constraint that a trial length could be no more than 4 s. This resulted in a variable number of volumes in the subjects' fMRI data (ranging from 1009 to 1119 volumes).

fMRI Data Analysis

The images were preprocessed and analyzed with FSL 4.1.7 (www.fmrib.ox.ac.uk/fsl, last accessed on 09.12.2013). For each raw BOLD dataset, nonlinear noise reduction was performed using SUSAN (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/SUSAN, last accessed on 09.12.2013). Preprocessing included motion correction with MCFLIRT (Jenkinson et al. 2002), fieldmap-based echo-planar imaging unwarping using PRELUDE + FUGUE (Jenkinson 2003), nonbrain removal using BET (Smith 2002), spatial smoothing with a Gaussian kernel of full-width at half-maximum 6 mm; grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor; high-pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma = 50.0 s). Given that we used a block design, no slice-timing correction was applied.

Preprocessed data were subjected to the Probabilistic Independent Component Analysis (Beckmann and Smith 2004) that was implemented using FSL's MELODIC (Multivariate Exploratory Linear Decomposition into Independent Components, http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/MELODIC, last accessed on 09.12.2013) Version 3.10. In the input data, the nonbrain voxels were masked, the data were demeaned voxel-wise, and the voxel-wise variance was normalized. Preprocessed data were whitened and projected into a 20-dimensional subspace using principal component analysis. The whitened observations were decomposed into sets of vectors that describe signal variation across the temporal domain (time courses) and across the spatial domain (maps) by optimizing for nonGaussian spatial source distributions using a fixed-point iteration technique (Hyvärinen 1999). Estimated component maps were divided by the standard deviation of the residual noise and thresholded by fitting a mixture model to the histogram of intensity values (Beckmann and Smith 2004). The noise components were identified based on published recommendations (Tohka et al. 2008; Kelly et al. 2010) and removed using the fsl_regfilt script.

The denoised data were submitted to a first (subject)-level GLM analysis that was implemented using FEAT (FMRI Expert Analysis Tool, v5.98). This analysis examined the parametric changes in brain activation as a function of memory load by comparing BOLD signal among the blocks of 1- versus 2- versus 3-back task. The model included 6 regressors: 1-, 2-, and 3-back instruction periods, and 1-, 2-, and 3-back task performance blocks.

Coregistration was carried out using FLIRT (Jenkinson and Smith 2001; Jenkinson et al. 2002). BOLD images were registered to the high-resolution structural (MPRAGE) images, the high-resolution images were registered to the MNI152_T1_2 mm template, and the 2 resulting transformations were concatenated and applied to the original BOLD image (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FLIRT, last accessed on 09.12.2013) to transform it to the MNI space.

A group analysis was performed using a permutation method (Nichols and Holmes 2002) implemented through the Randomise v2.1 tool (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Randomise, last accessed on 09.12.2013), with a whole brain as a mask, 5 mm smoothing, 5000 permutations, and correction for multiple comparisons at the voxel-wise FWE-controlled threshold $P < 0.05$. Then, those regions that parametrically increased (Increase network) or decreased (Decrease network) in activation across the 3 levels of *n*-back were used as the regions of interest (ROIs) in subsequent analyses.

The Analyses of Preparation Periods

Univariate Analysis (GLM)

The first of those subsequent analyses contrasted activations during the instruction periods preceding 1-, 2-, and 3-back task blocks in the Increase and Decrease networks identified in the previous analysis. The group analysis was also performed using the Randomise tool with 5 mm smoothing, 5000 permutations, threshold-free cluster enhancement (Smith and Nichols 2009) correction for multiple comparisons, and FWE-controlled threshold $P < 0.05$.

Multivoxel Pattern Classification

The second of the follow-up analyses used MVPA implemented using PYMVPA (Hanke et al. 2009). This analysis examined whether the patterns of brain activation in the Increase and Decrease networks during "task preparation" can predict the upcoming *n*-back condition (1- vs. 2- vs. 3-back). The voxels of interest for MVPA were selected based on the results of the GLM analysis, so no further feature selection was performed. MVPA was conducted using a linear support vector machine (SVM; Vapnik 1995), with the default parameter $C = 1$ and a leave-one-sample-out cross-validation strategy on all voxels in the Increase network and separately on all voxels in the Decrease network. The masks of Increase and Decrease networks were coregistered to each subject's BOLD images. Classification accuracies were computed for each subject in the subject's space. A one-versus-one multiclass classification was used (for detailed explanation see http://www.csie.ntu.edu.tw/~cjlin/libsvm/, last accessed on 09.12.2013). The resulting accuracies were then averaged across the subjects, but not across the networks, and compared against chance (33% for three conditions). While classifying 1- versus 2- versus 3-back conditions during task performance was not the focus on our study, we still conducted this analysis to confirm that the three *n*-back conditions differ not only in the magnitude of activation, but also in the patterns of activation.

The samples used as inputs to MVPA were volumes (or TRs) pertaining to each condition of interest. There were 10 preparation periods for each of the three *n*-back conditions each lasting for 4 TRs (or 8 s). Overall, there were a total of 40 samples associated with preparation for each of the three *n*-back conditions. The length of the *n*-back blocks varied across subjects and conditions (because the task was self-paced and because 3-back was slower than other conditions). The length of each block was not, however, <42 s (21 TRs) for any *n*-back condition for any subject. Having a different number of samples for different conditions can bias a classifier. To avoid this bias, only the first 21 TRs from each *n*-back block were used for the analysis, resulting in a total of 210 samples (21 TRs × 10 blocks) per *n*-back condition.

Relationship Between Preparatory Activation and Behavior

The relationship between classification accuracy during preparation periods and behavioral measures was examined in 2 ways. We correlated subjects' RT for each level of *n*-back and classification accuracy for 1- versus 2- versus 3-back preparation periods separately in the Increase and Decrease networks. A total of 6 correlation coefficients were computed (three *n*-back conditions in 2 networks). In addition, we also used 6 behavioral measures (response accuracy and RT for the 3 levels of *n*-back) as inputs to the K-mean cluster analysis ($k = 2$) to partition subjects into 2 groups based on their behavioral performance. These 2 groups of subjects were then compared for their classification accuracy in the Increase and Decrease networks. In addition, the same correlation analyses were performed between the differences in average activation for 1- versus 3-back preparation periods in Increase and Decrease networks and behavioral measures of performance.

Results

Behavioral

As in previous studies (Carlson et al. 1998; Nystrom et al. 2000), subjects were slower ($F_{2,30} = 46.4$, $P < 0.001$, partial $\eta^2 = 0.76$; Fig. 1*B*) and less accurate ($F_{2,30} = 22.5$, $P < 0.001$, partial $\eta^2 = 0.6$; Fig. 1*C*) as the value of *n* increased in the task.

Neuroimaging

The first analysis allowed us to identify a network of regions that parametrically increased activation when WM load increased (Increase network) and also a network of regions that decreased activation with an increase in WM load (Decrease network). Consistent with the previous work (for a review, see Owen et al. 2005; Rottschy et al. 2012), the Increase network consisted of prefrontal cortex, paracingulate/anterior cingulate cortex (ACC), parietal regions [including intraparietal sulcus (IPS)], and subcortical structures parametrically increased activation with increases in WM load (Table 1). Medial frontal cortex, posterior cingulate gyrus, and the inferior division of the lateral occipital cortex, among other regions, parametrically decreased activation as a function of WM load and comprised the Decrease network (Table 1).

Parametric Changes in the WM Load Regions During Preparation for 1-, 2-, and 3-Back Tasks

This analysis tested the hypothesis that activations in the WM Increase and Decrease networks (Fig. 2*A*) are modulated by expected levels of difficulty. A GLM analysis supported this hypothesis and showed that although left inferior frontal gyrus (LIFG), left IPS, and ACC increased activation linearly as a function of WM load during the task, they decreased activation during task preparation as a function of expected WM load (LIFG: z-max = 4.69, n-voxels = 118 [−40 6 24]; LIPS: z-max = 3.85, n-voxels = 99 [−42 −50 52]; ACC: z-max = 3.85, n-voxels = 12 [−4 12 48]) (Fig. 3*B*,*C*). The time courses associated with task preparation in LIFG, LIPS, and ACC were extracted using PEPATE (http://www.jonaskaplan.com/peate/index.php, last accessed on 09.12.2013) and are given in Supplementary Figures 1–3. A subsequent correlation analysis showed that the activation changes in these regions did not correlate with behavioral measures of performance (either RT or accuracy).

Multivoxel Pattern Classification Analysis

This analysis tested whether the expected n-back conditions can be decoded from the activation patterns in the Increase and Decrease networks. The Increase network on average consisted of 3413 (standard deviation (SD) = 284) voxels (in subjects' space) across subjects. The Decrease network consisted on average of 3726 (SD = 314) voxels across subjects. Not surprisingly, MVPA accurately decoded the 3 levels of n-back during task

Table 1

Parametric changes in brain activation as a function of WM load during task performance and classification accuracies for 1- versus 2- versus 3-back conditions during instruction periods and task performance

	Region	BA	n-voxels	z-max	x	y	z	SVM classification accuracy 1- versus 2- versus 3-back	
								Instructions	Task
Parametric increases in activation (1- < 2- < 3-back) during task performance									
L	Anterior cingulate cortex (ACC)/paracingulate gyrus	32	818	9.99	−2	22	42	0.48 (0.08)	0.54 (0.06)
R	Inferior parietal sulcus (IPS)	40	503	8.29	40	−46	44	0.45 (0.05)	0.50 (0.07)
L	Orbitofrontal cortex (OFc)	13	447	9.66	−32	24	−6	0.43 (0.07)	0.47 (0.06)
L	Inferior parietal sulcus (IPS)	40	407	8.0	−34	−52	40	0.43 (0.07)	0.48 (0.06)
L	Superior frontal gyrus (SFG)	6	388	7.4	−28	8	62	0.48 (0.05)	0.53 (0.06)
R	Frontal pole (FP)	10	370	8.22	42	48	22	0.46 (0.08)	0.54 (0.07)
R	Precuneus (Prec)	7	314	7.99	2	−62	48	0.46 (0.08)	0.51 (0.05)
L	Frontal pole (FP)	10	313	7.58	−38	56	10	0.47 (0.06)	0.53 (0.06)
L	Inferior frontal gyrus (IFG)	9	300	8.21	−40	8	26	0.41 (0.06)	0.47 (0.06)
L	Middle frontal gyrus (MFG)	46	277	7.66	−48	36	22	0.45 (0.06)	0.49 (0.06)
R	Superior frontal gyrus (SFG)	6	262	6.87	28	12	60	0.43 (0.09)	0.50 (0.06)
R	Insular cortex (Ins)	13	177	8.88	34	24	−2	0.44 (0.07)	0.45 (0.04)
L	Basal ganglia (Bas)		160	7.2	−16	0	14	0.39 (0.07)	0.41 (0.05)
R	Lateral occipital cortex, superior (LOCs)	19	97	7.1	32	−66	46	0.43 (0.07)	0.48 (0.05)
L	Thalamus (Thal)		52	6.49	−8	−18	8	0.36 (0.12)[a]	0.40 (0.05)
R	Cerebellum (Cerebel)		27	5.83	36	−70	−28	0.39 (0.05)	0.47 (0.05)
L	Lateral occipital cortex, superior (LOCs)	7	26	6.42	−16	−74	52	0.40 (0.09)	0.48 (0.05)
R	Middle frontal gyrus (MFG)	8	10	5.92	44	34	38	0.41 (0.07)	0.47 (0.06)
Parametric decreases in activation (3- < 2- < 1-back) during task performance									
L	Planum polare	13	1599	10.4	−42	−16	−8	0.44 (0.04)	0.47 (0.06)
R	Parietal operculum	13	1305	8.67	48	−30	22	0.48 (0.07)	0.49 (0.08)
L	Anterior cingulate/medial frontal cortex	10	968	8.22	−2	54	−4	0.48 (0.09)	0.52 (0.06)
R	Juxtapositional lobule cortex	6	916	9.0	2	−8	48	0.45 (0.06)	0.48 (0.05)
L	Posterior cingulate gyrus	31	591	8.29	−10	−50	28	0.44 (0.08)	0.47 (0.07)
R	Right precentral gyrus	4	312	7.1	36	−18	48	0.41 (0.09)	0.47 (0.07)
R	Planum polare	38	228	7.14	42	0	−18	0.42 (0.05)	0.43 (0.05)
L	Posterior cingulate gyrus	31	200	8.47	−14	−30	38	0.40 (0.09)	0.43 (0.05)
L	Postcentral/precentral gyrus	4	64	8.2	−36	−18	40	0.40 (0.04)	0.42 (0.05)
R	Lateral occipital cortex, inferior	19	60	6.52	54	−70	8	0.40 (0.06)	0.44 (0.06)
R	Postcentral gyrus	2	31	6.41	56	−16	48	0.41 (0.09)	0.47 (0.07)
R	Temporal pole	38	17	6.55	40	24	−26	0.44 (0.07)	0.45 (0.04)
R	Subcallosal cortex	25	13	5.9	0	16	−12	0.37 (0.07)	0.40 (0.04)
R	Frontal pole	11	13	5.85	36	36	−16	0.41 (0.08)	0.44 (0.04)
R	Postcentral gyrus	2	12	6	30	−32	70	0.38 (0.09)	0.45 (0.06)
L	Occipital pole	19	11	5.69	−26	−94	24	0.38 (0.08)	0.45 (0.04)
R	Postcentral gyrus	40	10	5.71	26	−36	56	0.32 (0.09)[a]	0.40 (0.05)

Note: The images were thresholded at voxel-wise FWE-corrected $P < 0.05$. The abbreviations for region labels used in the text are in parentheses next to the labels.

Standard deviations for the mean classification accuracies are referred in parentheses.

L: left; R: right; BA: Brodmann areas.

[a] Classification accuracies at chance.

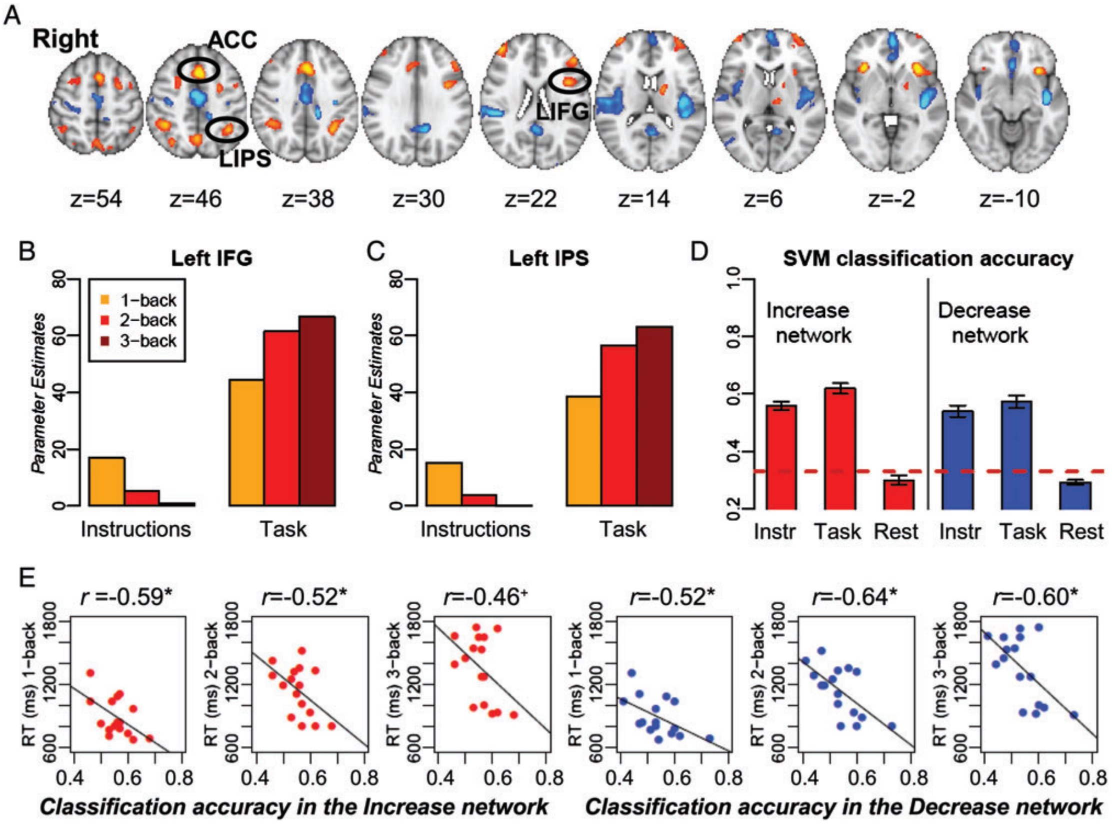

Figure 2. (*A*) The regions that parametrically increased (in red) and parametrically decreased (in blue) in activation with an increase in WM load. (*B*) LIFG and (*C*) LIPS parametrically increased in activation as a function of WM load during the task, but parametrically decreased activation during instruction periods. (*D*) SVM classification accuracy for 1- versus 2- versus 3-back conditions for instruction, *n*-back blocks, and random rest periods. Error bars refer to a ±1 standard error. (*E*) Correlation between the classification accuracy for 1- versus 2- versus 3-back instruction periods in the Increase (red) and Decrease (blue) networks and RT on the 1-, 2-, and 3-back task. The correlations denoted with "*" are significant at $P < 0.05$, and the correlation denoted with "+" is marginally significant at $P = 0.08$.

performance in the Increase and Decrease networks [Increase network: mean SVM accuracy = 0.63 (SD = 0.07); Decrease network: mean SVM accuracy = 0.58 (SD = 0.09)]. Given that this specific analysis was performed on the features (voxels) that have already been identified as informative by the GLM analysis of 1- versus 2- versus 3-back task, we will not discuss these MVPA findings further in the text as they are difficult to interpret separately from the univariate analysis.

Importantly, the SVM classifier accurately decoded the *n*-back conditions not only during task performance, but also during the preparatory periods that preceded performance on 1-, 2-, and 3-back tasks [Increase network: mean SVM accuracy = 0.56 (SD = 0.06); Decrease network: mean SVM accuracy = 0.54 (SD = 0.08)] (Fig. 2*D*). When calculated across subjects, the mean classification accuracy was significantly above chance (33%) for all conditions (Increase network, *n*-back task: $t_{15} = 15.9$, $P < 0.001$; Decrease network, *n*-back task: $t_{15} = 11.0$, $P < 0.001$; Increase network, preparation: $t_{15} = 15.9$, $P < 0.001$; Decrease network, preparation: $t_{15} = 10.5$, $P < 0.001$). Table 2 summarizes the mean percent of correctly classified trials [i.e., true positives (sometimes called sensitivities)] for each of the three *n*-back conditions in each network for task preparation and task performance. The data in the table confirm that the above-chance classification performance was achieved because each of the three conditions was accurately classified [compared with a situation when the above-chance classification could be achieved because classification accuracy for 1 condition was very high (e.g., 95%), but classification accuracy for 2 other conditions was at chance (e.g., 33%)].

To make sure that the classification results are not due to some peculiar noise structure in the data, we conducted the classification analysis on 120 volumes (TR) randomly selected from the rest periods in the Increase and Decrease networks. These 120 TRs were randomly assigned to the three conditions (we call them random rest 1, random rest 2, and random rest 3) to resemble the structure of the classifier used for the analysis of the preparation periods. The logic was that if the above-chance classification accuracies for the preparation periods are explained by some specific noise structure in the data, we will find the above-chance classification accuracy for this random rest analysis. This was not, however, the case, as the classification accuracies of this random rest analysis were not above (and even slightly below) the chance level (Fig. 2*D*).

While we based our conclusions on the classification analysis of the whole Increase and Decrease networks, we performed an additional MVPA in each of the regions comprising a specific network for each individual subject. The goal of this analysis was to examine whether the classification accuracy in the whole network analysis was determined by the contribution from some specific regions, or by the contribution from the voxels in the distributed network. The results of this

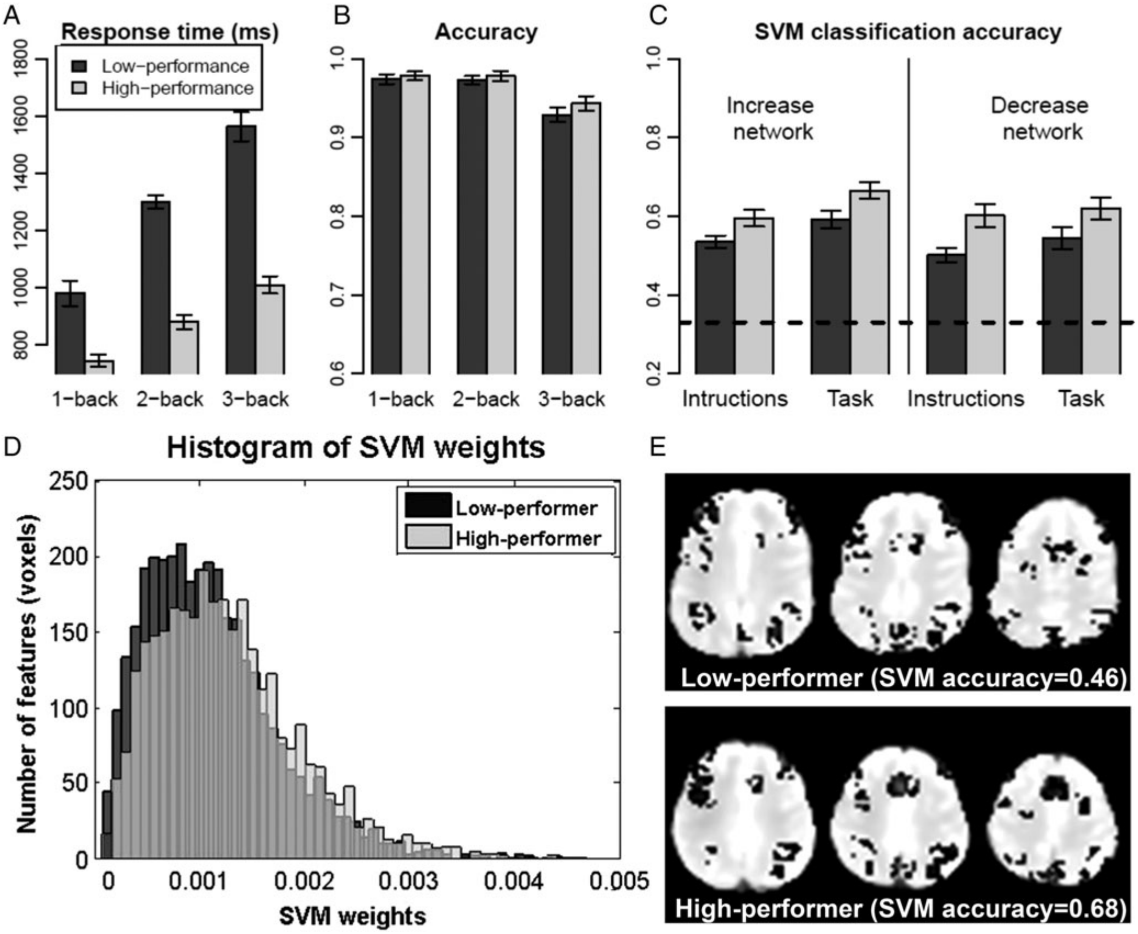

Figure 3. Comparison of higher and lower performing subjects. (*A*) RT for correct responses and (*B*) percent correct for the 3 levels of *n*-back. (*C*) SVM classification accuracy in the Increase and Decrease networks. (*D*) Histogram of the absolute SMV weights for one higher- and one lower-performance subjects in the Increase network. (*E*) The absolute SVM weights (thresholded at 0.001) in the Increase network are overlaid on the individual brains of the same higher- and lower-performance subjects shown in *D*. The figure illustrates that a subject with lower classification accuracy has fewer informative features compared with a subject with higher classification accuracy.

Table 2

Percent of correctly classified trials (true positives) for each condition of interest

	1-back	2-back	3-back
Increase network task preparation	58.3 (11.1)	54.7 (10.2)	54.7 (11.7)
Increase network task performance	72.8 (8.2)	53.4 (9.7)	59.5 (10.3)
Decrease network task preparation	54.7 (12.9)	53.0 (14.3)	54.5 (11.9)
Decrease network task performance	65.6 (12.3)	51.4 (9.5)	54.6 (9.5)

Note: Standard deviations for the mean classification accuracies are referred in parentheses.

second classification analysis were averaged across subjects for each region in the network, and the results are presented in Table 1 (2 right columns). The classification accuracy was above chance in all regions during *n*-back task performance. During task preparation, one region in the Increase network (left thalamus) and another region in the Decrease network (right postcentral gyrus) failed to classify the three *n*-back conditions above chance.

We further explored data by computing the mean accuracies across all regions in a specified network [the mean accuracy across 18 regions in the Increase network: preparation = 0.43 (SD = 0.03), *n*-back task = 0.48 (SD = 0.04); the mean accuracy across 17 regions in the Decrease network: preparation = 0.41 (SD = 0.04), *n*-back task = 0.45 (SD = 0.03)]. These classification accuracies were above the 33% chance level (Increase network, *n*-back task: $t_{17} = 16.2$, $P < 0.001$; Decrease network, *n*-back task: $t_{16} = 15.4$, $P < 0.001$; Increase network, preparation: $t_{17} = 9.0$, $P < 0.001$; Decrease network, preparation: $t_{16} = 8.6$, $P < 0.001$). However, they were significantly lower than the classification accuracies for the whole networks (Increase network, *n*-back task: $t_{17} = -15.2$, $P < 0.001$; Decrease network, *n*-back task: $t_{16} = -16.3$, $P < 0.001$; Increase network, preparation: $t_{17} = -16.3$, $P < 0.001$; Decrease network, preparation: $t_{16} = -13.3$, $P < 0.001$). This result suggests that discrimination between 1-, 2-, and 3-back conditions during both preparation and task performance relies on the distributed network of regions rather than on some specific region in that network.

Relationship Between Classification Accuracy and Behavioral Performance on the n-Back Task

The previous analyses showed that it is possible to decode task difficulty from the preparation periods that precede the task. What remains unclear is whether subjects' classification accuracy during anticipation of 1-, 2-, and 3-back tasks is related to subsequent task performance. A correlation analysis between classification accuracies in the Increase and Decrease networks during task preparation and RTs for 1-, 2-, and 3-back tasks showed that the subjects with higher classification accuracies (and, consequently, more distinct neural representations of the 3 task conditions) were faster at each level of the *n*-back task. Figure 2*E* reports correlation coefficients and *P*-values for all 6 analyses.

Another result that points to the same conclusion and supports the idea that task preparation determines the quality of task performance (e.g., Park and Rugg 2010) used a combination of all 6 behavioral measures (RT and response accuracy for 1-, 2-, and 3-back). These measures were entered into the K-mean cluster analysis ($k = 2$). This analysis partitioned 16 subjects into groups of 10 and 6 subjects based on their behavioral performance. Given that subjects in the larger group were slower ($F_{1,14} = 67.6$, $P < 0.001$, partial $\eta^2 = 0.83$; Fig. 3*A*), although not less accurate (Fig. 3*B*), we refer to them as the "lower-performance" group and the other subjects as the "higher-performance" group. "Higher-performers" were not only faster than "lower-performers", but their latencies were less affected by task difficulty ($F_{1,14} = 7.3$, $P < 0.05$, partial $\eta^2 = 0.34$). The important point is that "higher-performers", compared with "lower-performers," had higher SVM classification accuracies for expected 1- versus 2- versus 3-back conditions during preparation periods in the Increase ($t_{14} = 2.3$, $P < 0.05$) and Decrease ($t_{14} = 3.1$, $P < 0.01$) networks (Fig. 3*C*). These differences suggest that the voxels in the Increase or Decrease networks of the "higher-performers" contain more information about the upcoming *n*-back condition compared with "lower-performers." In the SVM classifier, a discriminative ability of a voxel is expressed as a weight (a parameter calculated by the SVM that can be interpreted as a distance of a specific data point from the hyperplane) with greater weights corresponding to more informative features. To illustrate how "higher-performers" and "lower-performers" differ in terms of the discriminative ability of the voxels in the ROIs, we plotted the absolute SVM weights taken from the Increase network classification of preparation periods for one "higher-performer" and one "lower-performer" (Figs 3*D*,*E*). A histogram shown in Figure 3*D* illustrates that a "higher-performer," compared with a "lower-performer," has more features (voxels) with higher absolute weights and fewer features with lower absolute weights in the Increase network. When the absolute SVM weights were mapped to the subjects' brains and thresholded to remove less informative features, the resulting representation of informative features was sparser for the "lower-performer" (Fig. 3*E*).

Discussion

The present study examined the novel question of how brain activation in the WM network is modulated by anticipated difficulty of the upcoming task during the instruction period that we refer to as task preparation. The WM network is operationally defined here as the network of regions whose activity is modulated by the changes in WM load during task performance. The regions comprising the WM network were identified by contrasting brain activations during performance on 1- versus 2- versus 3-back tasks. Consistent with multiple neuroimaging studies (see Owen et al. 2005; Rottschy et al. 2012 for reviews), parametric increases in WM load during task performance resulted in parametric increases in activation in prefrontal cortex (PFC), ACC, IPS, and some subcortical structures (Increase network) and parametric decreases in medial frontal cortex, posterior cingulate gyrus, and the inferior division of the lateral occipital cortex (Decrease network). The univariate (i.e., GLM) and multivariate (i.e., SVM) analyses revealed that the WM network dynamically adjusts to anticipated cognitive demands during task preparation.

Parametric Effects of Expected Difficulty on Brain Activation During Preparation

Our finding that the left IFG, IPS, and ACC increase in activation during task performance when WM load increases, but decrease in activation when the anticipated difficulty of the upcoming task increases, is inconsistent with the proposal that effective task preparation involves preactivation of the neural networks engaged during task performance (e.g., Wylie et al. 2006). Rather, it seems that, at least in some situations, the activation in task-related regions may be inhibited prior to task execution, a finding that is consonant with recent MEG results (Altamura et al. 2010). Their MEG study found decreased preparatory activation (expressed as a beta power increase) in several brain regions (including left IFG, DLPFC, and left parietal cortex) for a WM, compared with control, task that did not engage WM. The Altamura et al. study, however, did not vary WM load parametrically, so it is unclear whether this modulation is an all-or-none response or is graded as a function of expected cognitive load. Our results indicate that modulation of brain response by anticipated cognitive load occurs in a graded manner with greater decreases during anticipation of more difficult tasks.

Many previous neuroimaging studies have found that PFC and ACC play an important role in task preparation (e.g., MacDonald et al. 2000; Brass and von Cramon 2002, 2004; Luks et al. 2002; Sohn et al. 2007; Lavric et al. 2008). Specifically, it was proposed that the posterior left IFG, also called inferior frontal junction or IFJ, is involved in the processing and context-related updating of the task representations during task preparation (Brass and von Cramon 2004). The coordinates of the left IFJ reported in Brass and von Cramon [−37 8 35 Talairach] are very close to those reported in our study [−40 6 24 MNI] for the LIFG. The finding that this region responds parametrically to anticipated WM load provides converging evidence for the role of the left IFJ in task preparation by showing that this region not only supports processing and updating of the task representations, but also controls the amount of allocated cognitive resources based on expected cognitive load.

Preparation to perform a task is a complex process that includes (but is not limited to) encoding of task instructions, retrieval of the relevant rules for the current task, retrieval of the stimulus–response mapping, effort anticipation, planning to allocate attention to specific features of the stimuli, and what we call adjusting/cleaning memory storage. The last process involves suppression of information that was encoded for the previous task in order to reduce proactive interference. This last process may explain why left IFG, left IPS, and ACC decreased activation when anticipated WM load increased. Previous neuroimaging studies of WM suggest that all these regions are involved in maintenance of memory representations by increasing activation for tasks with greater WM load (Courtney et al. 1997; Rypma and D'Esposito 1999; D'Esposito et al. 2000; Pessoa et al. 2002; Narayanan et al. 2005). Given that WM is limited capacity (e.g., Baddeley and Hitch 1974), it is reasonable that people try to minimize interference from the sources not related to the task at hand as the task becomes more demanding (e.g., ask a friend to stop talking when one is engaged in a complex task such as driving in merging traffic or computing a complex sum). The load-related decreases during task preparation in the regions involved in maintenance of memory representations may be related to a brain mechanism that helps to "clear the mind" before starting a demanding WM task. For example, if a subject anticipates performing the 3-back task, she or he knows that it will involve maintaining and manipulating in memory 3 items compared with the one item required when performing the 1-back task. Consequently, the memories of stimuli, responses, and rules not related to the upcoming task (particularly those coming from previous blocks of trials) should be expunged to free up more processing resources to deal with a complex task.

It is unclear whether the difficulty of the preceding task influences how much one needs to "clear the mind" during preparing for an upcoming task. Unfortunately, we were unable to address this question, because the design of our current experiment did not provide sufficient power for such analyses: the order of *n*-back conditions was randomly determined for each subject. That means that the number of transitions between various conditions varied by subject. Future research should examine this question in an experiment that controls for the number and order of transitions among WM load levels.

"Clearing the mind" may also be a part of an effort calculation process that conceivably occurs during task preparation (Croxson et al. 2009). Croxson et al. examined the interaction between expected reward and expected effort during task preparation/anticipation. They found that activation in dorsal ACC decreased with an increase in anticipated effort given that the high reward was expected. Our study did not have a reward component, but the results are consistent with that of Croxson et al. The 3-back task requires greater effort than 1-back, and this difference appears to be anticipated during the preparation periods.

The results of our study demonstrate that neither encoding of task instructions, retrieval of task rules, nor the stimulus–response mapping can explain why the left IFG, left IPS, and ACC, the regions comprising the Increase network, parametrically decreased activation during task preparation when anticipated cognitive demands increased. First, the instruction screens contained just one word and it is unlikely that encoding of the phrase "1-back" was easier or more difficult than that of the phrase "3-back." Secondly, the stimulus–response mapping (e.g., press a button with your right hand if this is a target) was identical for all *n*-back conditions. Thirdly, a retrieval of more complex rules elicits increases in brain activation of the left ventro-lateral PFC [−42 6 30] and the left inferior parietal cortex [−36 −51 45] (Bunge et al. 2003). Given that the coordinates reported by Bunge et al. are very close to the coordinates of the left IFG and IPS in our study, it may follow that 1-back rule is more complex than 3-back. This, however, seems unlikely because subjects practiced all *n*-back conditions outside the scanner to achieve high proficiency on the task prior to performing the *n*-back task in the scanner and because, in general, it takes longer to explain to a subject the 2- or 3-back rule than the 1-back.

Multivoxel Pattern Classification Analysis

An important goal of cognitive neuroscience is to be able to predict people's future behavior from the brain activations that precede them (Haynes 2011). Several previous studies have successfully used MVPA to decode subjects' intentions (e.g., Gilbert 2007; Haynes et al. 2007; Gallivan et al. 2011). Using this technique, we were able to support our hypothesis about the involvement in the WM network during preparation for WM tasks of varying difficulty by showing that the upcoming *n*-back conditions (1- vs. 2- vs. 3-back) can be decoded from the activation patterns in the Increase and Decrease networks during preparatory periods. Because subjects' accuracy in the *n*-back task were uncorrelated with classification accuracy, we believe that subjects' ability to memorize, maintain, and manipulate the stimuli during the task did not depend on whether or not the Increase and Decrease networks contained information about the upcoming difficulty level during task preparation. However, based on the finding that greater 1- versus 2- versus 3-back classification accuracy during task preparation predicts faster behavioral responses and less of an increase in RT when WM load increases at task, we propose that the subjects who form more distinct neural representations of upcoming *n*-back conditions in the Increase and Decrease networks have more fluent (i.e., easier) access to the information during the task.

It could be argued that the differences in patterns across the WM load conditions during the preparatory periods revealed by MVPA reflect the overall changes in activity across load levels that have been revealed by the GLM analysis. If this were the case, the decreases in BOLD signal in the LIFG, LACC, and LIPS during task preparation should correlate with classification accuracies in these regions during task preparation. We tested this idea and found that the 3- minus 1-back BOLD signal changes did not correlate with the 1- versus 2- versus 3-back SVM classification accuracies in either ROI (the lowest P-value of >0.25).

In summary, the present work advances our understanding of the neural mechanisms of preparation for tasks of varying difficulty. The results of the present study suggest that neuroimaging research on WM should be extended to task preparation. These results also have implications for clinical research concerning cognitive impairments and learning disabilities. Specifically, a treatment may depend on whether subject's performance is impaired due to inability to prepare for the task or inability to process the information during the task.

Supplementary Material

Supplementary material can be found at: http://www.cercor.oxfordjournals.org/.

Funding

This work was supported by a NIMH (training grant T32MH019983).

Notes

Conflict of Interest: None declared.

References

Aarts E, Roelofs A, van Turennout M. 2008. Anticipatory activity in anterior cingulate cortex can be independent of conflict and error likelihood. J Neurosci. 28:4671–4678.

Adcock RA, Thangavel A, Whitfield-Gabrieli S, Knutson B, Gabrieli JDE. 2006. Reward-motivated learning: mesolimbic activation precedes memory formation. Neuron. 50:507–517.

Altamura M, Goldberg TE, Elvevåg B, Holroyd T, Carver FW, Weinberger DR, Coppola R. 2010. Prefrontal cortex modulation during anticipation of working memory demands as revealed by magnetoencephalography. Int J Biomed Imaging. doi:10.1155/2010/840416.

Baddeley A. 2010. Working memory. Curr Biol. 20:136–140.

Baddeley AD, Hitch GJ. 1974. Working memory. In: Bower GA, editors. Recent advances in learning and motivation. Vol. 8. New York: Academic Press. p. 47–90.

Beckmann CF, Smith SM. 2004. Probabilistic independent component analysis for functional magnetic resonance imaging. IEEE Trans Med Imaging. 23:137–152.

Brass M, von Cramon DY. 2004. Decomposing components of task preparation with functional magnetic resonance imaging. J Cogn Neurosci. 16:609–620.

Brass M, von Cramon DY. 2002. The role of the frontal cortex in task preparation. Cereb Cortex. 12:908–914.

Braver TS, Barch DM, Kelley WM, Buckner RL, Cohen NJ, Miezin FM, Snyder AZ, Ollinger JM, Akbudak E, Conturo TE et al. 2001. Direct comparison of prefrontal cortex regions engaged by working and long-term memory tasks. Neuroimage. 14:48–59.

Braver TS, Cohen JD, Nystrom LE, Jonides J, Smith EE, Noll DC. 1997. A parametric study of frontal cortex involvement in human working memory. Neuroimage. 5:49–62.

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.

Carlson S, Martinkauppi S, Rämä P, Salli E, Korvenoja A, Aronen HJ. 1998. Distribution of cortical activation during visuospatial n-back tasks as revealed by functional magnetic resonance imaging. Cereb Cortex. 8:743–752.

Chatham CH, Herd SA, Brant AM, Hazy TE, Miyake A, O'Reilly R, Friedman NP. 2011. From an executive network to executive control: a computational model of the n-back task. J Cogn Neurosci. 23:3598–3619.

Cohen JD, Perlstein WM, Braver TS, Nystrom LE, Noll DC, Jonides J, Smith EE. 1997. Temporal dynamics of brain activity during a working memory task. Nature. 386:604–608.

Courtney SM, Ungerleider LG, Keil K, Haxby JV. 1997. Transient and sustained activity in a distributed neural system for human working memory. Nature. 386:608–611.

Croxson PL, Walton ME, O'Reilly JX, Behrens TEJ, Rushworth MFS. 2009. Effort-based cost-benefit valuation and the human brain. J Neurosci. 29:4531–4541.

Curtis CE, D'Esposito M. 2003. Persistent activity in the prefrontal cortex during working memory. Trends Cogn Sci. 7:415–423.

D'Esposito M, Postle BR, Rypma B. 2000. Prefrontal cortical contributions to working memory: evidence from event-related fMRI studies. Exp Brain Res. 133:3–11.

Gallivan JP, McLean DA, Valyear KF, Pettypiece CE, Culham JC. 2011. Decoding action intentions from preparatory brain activity in human parieto-frontal networks. J Neurosci. 31:9599–9610.

Gilbert SJ. 2011. Decoding the content of delayed intentions. J Neurosci. 31:2888–2894.

Guderian S, Schott BH, Richardson-Klavehn A, Duzel E. 2009. Medial temporal theta state before an event predicts episodic encoding success in humans. Proc Natl Acad Sci USA. 106:5365–5370.

Hanke M, Halchenko YO, Sederberg PB, Hanson SJ, Haxby JV, Pollmann S. 2009. PyMVPA: a python toolbox for multivariate pattern analysis of fMRI data. Neuroinformatics. 7:37–53.

Haynes JD. 2011. Decoding and predicting intentions. Ann N Y Acad Sci. 1224:9–21.

Haynes JD, Sakai K, Rees G, Gilbert S, Frith C, Passingham RE. 2007. Reading hidden intentions in the human brain. Curr Biol. 17:323–328.

Hyvärinen A. 1999. Fast and robust fixed-point algorithms for independent component analysis. IEEE Trans Neural Netw. 10:626–634.

Jenkinson M. 2003. A fast, automated, n-dimensional phase unwrapping algorithm. Magn Reson Med. 49:193–197.

Jenkinson M, Bannister P, Brady M, Smith S. 2002. Improved optimisation for the robust and accurate linear registration and motion correction of brain images. Neuroimage. 17:825–841.

Jenkinson M, Smith S. 2001. A global optimisation method for robust affine registration of brain images. Med Image Anal. 5:143–156.

Kelly RE Jr, Alexopoulos GS, Wang Z, Gunning FM, Murphy CF, Morimoto SS, Kanellopoulos D, Jia Z, Lim KO, Hoptman MJ. 2010. Visual inspection of independent components: defining a procedure for artifact removal from fMRI data. J Neurosci Methods. 189:233–245.

Kragel PA, Carter RM, Huettel SA. 2012. What makes a pattern? Matching decoding methods to data in multivariate pattern analysis. Front Neurosci. 6:162.

Lavric A, Mizon GA, Monsell S. 2008. Neurophysiological signature of effective anticipatory task-set control: a task-switching investigation. Eur J Neurosci. 28:1016–1029.

Logan GD, Zbrodoff NJ. 1982. Constraints on strategy construction in a speeded discrimination task. J Exp Psychol Hum Percept Perform. 8:502–520.

Luks TL, Simpson GV, Feiwell RJ, Miller WJ. 2002. Evidence for anterior cingulate cortex involvement in monitoring preparatory attentional set. Neuroimage. 17:792–802.

MacDonald AW III, Cohen JD, Stenger VA, Carter CS. 2000. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. Science. 288:1835–1838.

Mackiewicz KL, Sarinopoulous I, Cleven KL, Nitschke JB. 2006. The effect of anticipation and the specificity of sex differences for amygdale and hippocampus function in emotional memory. Proc Natl Acad Sci USA. 103:14200–14205.

Mishkin M, Manning FJ. 1978. Non-spatial memory after selective prefrontal lesions in monkeys. Brain Res. 143:313–323.

Müller NG, Knight RT. 2006. The functional neuroanatomy of working memory: contributions of human brain lesion studies. Neuroscience. 139:51–58.

Mur M, Bandettini PA, Kriegeskorte N. 2009. Revealing representational content with pattern-information fMRI—an introductory guide. Soc Cogn Affect Neurosci. 4:101–109.

Narayanan NS, Prabhakaran V, Bunge SA, Christoff K, Fine EM, Gabrieli JDE. 2005. The role of the prefrontal cortex in the maintenance of verbal working memory: an event-related fMRI analysis. Neuropsychology. 19:223–232.

Neuper C, Wörtz M, Pfurtscheller G. 2006. ERD/ERS patterns reflecting sensorimotor activation and deactivation. Prog Brain Res. 159:211–222.

Nichols TE, Holmes AP. 2002. Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum Brain Mapp. 15:1–25.

Nystrom LE, Braver TS, Sabb FW, Delgado MR, Noll DC, Cohen JD. 2000. Working memory for letters, shapes, and locations: fMRI evidence against stimulus-based regional organization in human prefrontal cortex. Neuroimage. 11:424–446.

Otten LJ, Quayle AH, Akram S, Ditewig TA, Rugg MD. 2006. Brain activity before an event predicts later recollection. Nat Neurosci. 9:489–491.

Owen AM, McMillan KM, Laird AR, Bullmore E. 2005. N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. Hum Brain Mapp. 25:46–59.

Park H, Rugg MD. 2010. Prestimulus hippocampal activity predicts later recollection. Hippocampus. 20:24–28.

Pessoa L, Gutierrez E, Bandettini P, Ungerleider L. 2002. Neural correlates of visual working memory: fMRI amplitude predicts task performance. Neuron. 35:975–987.

Petrides M, Milner B. 1982. Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man. Neuropsychologia. 20:249–262.

Pfurtscheller G. 2001. Functional brain imaging based on ERD/ERS. Vision Res. 41:1257–1260.

Rottschy C, Langner R, Dogan I, Reetz K, Laird AR, Schulz JB, Fox PT, Eickhoff SB. 2012. Modelling neural correlates of working memory: a coordinate-based meta-analysis. Neuroimage. 60:830–846.

Rypma B, D'Esposito M. 1999. The roles of prefrontal brain regions in components of working memory: effects of memory load and individual differences. Proc Natl Acad Sci USA. 96: 6558–6563.

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

Smith S. 2002. Fast robust automated brain extraction. Hum Brain Mapp. 17:143–155.

Smith SM, Nichols TE. 2009. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. NeuroImage. 44:83–98.

Sohn MH, Albert MV, Jung K, Carter CS, Anderson JR. 2007. Anticipation of conflict monitoring in the anterior cingulate cortex and the prefrontal cortex. Proc Natl Acad Sci USA. 104:10330–10334.

Tohka J, Foerde K, Aron AR, Tom SM, Toga AW, Poldrack RA. 2008. Automatic independent component labeling for artifact removal in fMRI. Neuroimage. 39:1227–1245.

Vapnik V. 1995. The nature of statistical learning theory. New York: Springer.

Wylie GR, Javitt DC, Foxe JJ. 2006. Jumping the gun: is effective preparation contingent upon anticipatory activation in task-relevant neural circuitry? Cereb Cortex. 16:394–404.