

Dissociation of Verbal Working Memory System Components Using a Delayed Serial Recall Task

Jason M. Chein and Julie A. Fiez

Department of Psychology, Center for the Neural Basis of Cognition, University of Pittsburgh, Pennsylvania, USA

Functional magnetic resonance imaging (fMRI) was used to investigate the neural substrates of component processes in verbal working memory. Based on behavioral research using manipulations of verbal stimulus type to dissociate storage, rehearsal, and executive components of verbal working memory, we designed a delayed serial recall task requiring subjects to encode, maintain, and overtly recall sets of verbal items for which phonological similarity, articulatory length, and lexical status were manipulated. By using a task with temporally extended trials, we were able to exploit the temporal resolution afforded by fMRI to partially isolate neural contributions to encoding, maintenance, and retrieval stages of task performance. Several regions commonly associated with maintenance, including supplementary motor, premotor, and inferior frontal areas, were found to be active across all three trial stages. Additionally, we found that left inferior frontal and supplementary motor regions showed patterns of stimulus and temporal sensitivity implicating them in distinct aspects of articulatory rehearsal, while no regions showed a pattern of sensitivity consistent with a role in phonological storage. Regional modulation by task difficulty was further investigated as a measure of executive processing. We interpret our findings as they relate to notions about the cognitive architecture underlying verbal working memory performance.

Introduction

In the present study we use functional magnetic resonance imaging (fMRI) to characterize the properties of brain regions comprising the verbal working memory system. Verbal working memory can be defined as a limited capacity processing system that supports the online maintenance and manipulation of linguistic information over brief intervals. Neuroimaging studies have consistently implicated a distributed network of brain regions that contribute to verbal working memory, including dorsolateral prefrontal (BA 9/46), inferior parietal (BA 40), left inferior frontal (BA 44/45), supplementary motor (medial BA 6), premotor (lateral BA 6), and cerebellar regions (Fiez *et al.*, 1996; D'Esposito *et al.*, 1998; Smith and Jonides, 1998; Cabeza and Nyberg, 2000).

To interpret neuroimaging results, many researchers have turned to an influential model of working memory proposed by Baddeley and colleagues (Baddeley, 1986). According to this model, working memory is made up of multiple interacting components, including material-specific maintenance subsystems (*verbal* and *visuo-spatial*), and a *central executive* control system. Neuropsychological evidence (Shallice and Warrington, 1970; Vallar and Shallice, 1990; Van der Linden *et al.*, 1992; Gainotti *et al.*, 1994; Gathercole, 1994; Smith and Jonides, 1995) and neuroimaging studies (D'Esposito *et al.*, 1998, 1999; Smith *et al.*, 1996; Collette *et al.*, 1999; Smith and Jonides, 1999) are generally consistent with this model. The verbal maintenance subsystem, referred to as the 'phonological loop,' is the most

thoroughly characterized component in Baddeley's model, and it is this system that is of primary interest in the present investigation. Based on a complex pattern of behavioral results, discussed below, Baddeley proposed that this verbal maintenance subsystem may be further divided into a storage component and a rehearsal component. The storage component, or *phonological store*, is thought to buffer information coded at the level of phonology (speech perception). Representations in this store are subject to rapid decay, but can be 'refreshed' by a rehearsal component, the *articulatory control process*, which cycles through the contents of the store using an articulatory code resembling inner-speech.

There is some evidence from neuroimaging research indicating that these different components of verbal maintenance can be mapped to different brain regions. Specifically, it has been suggested that the left inferior parietal cortex supports storage, while a network of areas associated with speech production (left inferior frontal, premotor, supplementary motor, and cerebellar regions) mediates articulatory rehearsal (Paulesu *et al.*, 1993; Awh *et al.*, 1996; Jonides *et al.*, 1998; Smith *et al.*, 1998). However the extent to which these conclusions account for both neuroimaging and neuropsychological results has been debated (Becker *et al.*, 1999; Fiez, 2001).

Our objective is to further differentiate the neural substrates of verbal rehearsal and storage on the basis of temporal and stimulus-dependent differences in regional activation. In order to focus on verbal maintenance, we employ a delayed serial recall task. On each trial of this task, subjects are shown a list of verbal items, and explicitly rehearse these items over an extended delay, until cued for overt recollection. Because the delayed serial recall task makes use of temporally extended trials, it allows for the identification of brain regions that are active during maintenance, after the encoding of stimulus items into memory has been completed, and prior to the need for retrieval of these items from memory (Fiez *et al.*, 1996; Jonides *et al.*, 1998; D'Esposito *et al.*, 1999; Rypma and D'Esposito, 1999; Postle *et al.*, 2000). With fMRI, the extended delay interval can be further exploited to reveal dynamics in regional activation that coincide with changes in cognitive processing occurring *within* the maintenance stage of a trial.

In addition to honing in on maintenance with an extended delay interval, we take the further step of using manipulations of stimulus condition to examine the dissociability of the cognitive processes and neural substrates that support verbal maintenance. This approach is analogous to that employed in behavioral research, in which manipulations of stimulus condition have been used to derive evidence supporting the notion of separate storage and rehearsal components in the phonological loop. Prior behavioral studies of recollection in working memory have shown that compared to short, phonologically distinct

words, recall accuracy is attenuated for phonologically similar words (phonological similarity effect) (Conrad and Hull, 1964; Baddeley, 1966; Coltheart, 1993), for words of longer articulatory duration (word-length effect) (Baddeley *et al.*, 1975; Roe, 1993), and for pseudowords (pronounceable nonword items) (Hulme *et al.*, 1991; Multhaup *et al.*, 1996). It is argued that the phonological similarity effect proceeds from confusions between partially overlapping representations in the storage component, whereas the word-length effect stems from the operation of the articulatory rehearsal component, and reflects increased decay resulting from the additional rehearsal time needed for longer words (Baddeley *et al.*, 1984; Longoni *et al.*, 1993). Two alternative mechanisms may account for the lexicality effect. One view is that it results from interactions between the short-term and long-term memory systems, conferring an advantage to items having existing representations in long-term memory (real words, but not pseudowords). Alternatively, lexicality effects may proceed from additional demands associated with the sublexical phonological and phonetic processing of pseudoword items (Levelt *et al.*, 1999), which may impact on the operation of the verbal rehearsal system. Given the distinct mechanisms presumed to underlie the effects of phonological similarity, articulatory length, and lexicality, it is likely that they are associated with the operation of distinct neural regions.

According to Baddeley's model, the performance of a verbal working memory task should engage not only the verbal maintenance subsystem, but also the central executive. The central executive is thought to coordinate the behavior of the maintenance subsystems, and to further mediate the allocation of attention, the inhibition of task irrelevant processes, and the coding of contextual and temporal order information associated with representations held in memory (Baddeley, 1986; Baddeley and Logie, 1999; Smith and Jonides, 1999). A common approach used in previous investigations of central executive function has been to manipulate task variables that influence these executive processes. Examples include studies in which memory load is manipulated (Braver *et al.*, 1997; Cohen *et al.*, 1997; Rypma *et al.*, 1999), secondary tasks are used to burden the working memory system (D'Esposito *et al.*, 1995; Bunge *et al.*, 2000), or the duration of a memory delay interval is varied (Barch *et al.*, 1997). These previously utilized manipulations can be applied equivalently with verbal and visuo-spatial materials, and can therefore be thought of as domain-independent. In contrast, we examine the brain's response when task difficulty is varied through differences in verbal stimuli. Manipulations of verbal stimulus type are domain-specific in that they produce variations in task difficulty by differentially taxing processes thought to be specific to the verbal maintenance subsystem. The use of this domain-specific manipulation allows us to gain a novel perspective on the interaction between executive and maintenance-based processes.

In summary, in the present study, we employ the delayed serial recall task, together with stimulus manipulations of phonological similarity, articulatory length, and lexicality, with the following broad goals in mind: (i) to identify the neural substrates of verbal maintenance and to determine if they are selectively engaged in maintenance-based processing; (ii) to further characterize the particular functions served by brain regions engaged in specific rehearsal and storage processes through an investigation of stimulus-specific and temporal effects, and (iii) to examine the response of central executive regions to changes in task difficulty stemming from the operation of the verbal maintenance subsystem.

Materials and Methods

Subjects

Twelve right-handed, native English speaking, adults (six females; mean age 22.8 years) participated in scanning sessions. These subjects were recruited from the undergraduate and graduate student populations at the University of Pittsburgh (Pennsylvania, USA). All subjects gave informed written consent prior to participation in the study. Subjects participated in a short practice session, before being scanned, in order to develop familiarity with the task procedure.

Cognitive Task Procedure

A schematic diagram of a single task trial is presented in Figure 1. Each trial lasted for 52 s. At the start of a trial, subjects were presented with a sequential list of five word, or word-like, items. The items were presented visually in white, all capital Geneva font against a black background, positioned a few millimeters below a visually centered fixation point. Each item was displayed for 1 s, with a 600 ms interval between items. Following presentation of the list, a rehearsal indicator (—) appeared on the screen to denote the start of a 20 s maintenance interval. Subjects were explicitly instructed to covertly rehearse the presented items for the duration of the maintenance interval. After the 20 s passed, the rehearsal indicator was replaced by a recall cue (#####), to which subjects responded by overtly recalling the remembered list items. Four seconds were given to complete recall, and at the conclusion of this time the recall cue was removed. Subjects were instructed to say 'skip' as a place-holder for items that could not be accurately recalled, and all responses were tape recorded (a funnel placed proximally to the subject's mouth carried the voice signal through a pneumatic tube to a standard cassette-tape recording system). For the 20 s following recollection, subjects passively viewed a fixation point.

The specific nature of the presented list varied from trial to trial. Four distinct list types were used: (i) one-syllable phonologically distinct words (e.g. pit, stem, match); (ii) one-syllable phonologically similar words (e.g. fight, site, height); (iii) three-syllable phonologically distinct words (e.g. telescope, computer, suggestion); and (iv) one-syllable phonologically distinct pronounceable pseudowords (e.g. blick, rame, scote). Eight sets of five items each were developed for each stimulus condition. Word items having a written frequency of >30 (Francis and Kucera, 1982) formed the pool of potential word stimuli. The lists of phonologically similar items were composed by selecting one-syllable rhyming words from this pool. Lists of one- and three-syllable phonologically distinct words were then selected from the remaining set of words. To match items for individual character frequency, pseudoword items were formed by swapping initial consonant clusters from the short, phonologically distinct words, and the sets were re-randomized so that the pseudoword lists did not closely resemble the word lists. Across the one-syllable stimulus types, sets were further matched for average number of characters per word, to avoid systematic differences in orthographic length.

Data were collected in blocks, consisting of eight consecutive trials, that lasted 7 min and 12 s (which included a 16 s initial fixation period). Blocks consisted of two trials from each stimulus condition, and trial types were counterbalanced within and across blocks, as well as across subjects. Subjects were given ~2 min to rest between blocks, and completed eight blocks in all, giving a total of 16 trials for each stimulus condition. Since there were eight sets of stimuli for each stimulus condition, each set was used once in the first half of the session, and once in the second half of the session with a new item ordering.

Behavioral Data Analysis

Subject performance was evaluated by determining the accuracy of the subject's recall responses for each trial. A response item was considered correct if it matched the presented item exactly, and was recalled in the correct serial position. The total number of correctly recalled items in each trial (maximum of five) was then tallied to give the overall trial span score. Overall trial span scores were pooled for trials performed under the same stimulus conditions, providing an average trial span for each stimulus condition. This scoring procedure was carried out for each subject independently. To confirm the presence of the desired behavioral effects, the average trial span data were first submitted to a mixed-model ANOVA

with subject as a random factor and condition as a within-subject factor. Pair-wise matched samples *t*-tests were then used to directly compare performance with phonologically similar, three-syllable, and pseudoword items to one-syllable phonologically distinct items. An additional *t*-test was conducted to demonstrate an overall difficulty effect, contrasting performance with one-syllable phonologically distinct items to the average performance under the three other conditions. Although behavioral piloting studies were used to equate performance for the phonologically similar, three-syllable, and pseudoword trial types, additional pair-wise matched samples *t*-tests were performed to test for differences between these stimulus conditions. Bonferroni correction methods were used to adjust for expected probabilities with multiple comparisons in all *t*-tests.

fMRI Data Collection

Scanning was conducted on a 1.5 T whole-body GE Signa magnet. To minimize changes in head position across the session, subjects used a visual parallax marker, which provided a constant reference point for subject position. To establish the reference, a small knot on a transparent fishing line, located proximal to the subject's eyes, was positioned so that it was in line with the displayed fixation point, when viewed through one eye. Between blocks, subjects were instructed to re-align to the reference position, if necessary. Task stimuli were back-projected onto a visual display positioned above the subject's chest, and were viewed through a mirror.

A 36-slice oblique-axial structural series was collected parallel to the AC-PC plane (Noll *et al.*, 1997), using a standard T_1 -weighted spin-echo pulse sequence ($T_E = 12$ ms, $T_R = 500$ ms, FOV = 24, slice thickness = 3.8 skip 0 mm). Functional series were then acquired in the same plane as the structural series, but with coverage limited to the 26 center slices. This prescription provided coverage from just below the superior vertex to approximately mid-way through the cerebellum. Functional images were collected with a T_2^* -weighted, gradient-echo, two-shot spiral pulse sequence ($T_E = 35$, $T_R = 2000$, FOV = 24, slice thickness = 3.8 skip 0 mm, flip angle = 70°, inplane resolution = 3.75 mm) (Noll *et al.*, 1995). This procedure provided a full volume functional acquisition once every 4 s. Figure 1 shows the relationship between functional MRI acquisitions (13 images acquired per trial) and the behavioral task design.

fMRI Data Analysis

Analysis of the fMRI signal was conducted off-line using the Neuro-Imaging Software package (NIS 3.3) developed at the University of Pittsburgh and Princeton University. Following reconstruction and inhomogeneity correction, images were corrected for subject motion using a six-parameter rigid-body automated registration algorithm (AIR 3.08) (Woods *et al.*, 1992). To form a group composite data set, the structural images collected from each subject were co-registered to a common reference anatomy using a 12-parameter affine transformation algorithm (Woods *et al.*, 1993). Functional images were then transformed into the same common space and were subjected to normalization (achieved by mean scaling of each image to match global mean image intensities across subjects) and Gaussian smoothing (8 mm FWHM, to account for local anatomical variability between subjects). The group composite data was used in all voxel-based statistical analyses, although individual subject time series were also examined to confirm results derived from the group data. Statistical maps were visualized and transformed into stereotaxic coordinate space (Talairach and Tournoux, 1988) using AFNI 2.26 (Cox, 1996).

Analysis 1: Maintenance Regions

Our first set of analyses focused on the localization of verbal maintenance regions and the correspondence of these regions to those engaged during encoding and retrieval. To identify regions associated specifically with the maintenance stage of a trial, a comparison between maintenance and baseline was conducted using a voxel-wise *t*-test. For this comparison, data were pooled across the four stimulus conditions. The maintenance activation state was identified by selective averaging of data collected during the last two acquisitions of the maintenance interval (scans 6 and 7). A baseline control state was similarly derived by selective averaging of signal intensities from the final two acquisitions in each trial, which corresponded to passive viewing of the fixation point. Given the typical shape of the hemodynamic response function, using only the last two

scans of the maintenance and baseline intervals increased the likelihood that our selective averages did not include residual activation from the preceding stages of the trial.

An analogous approach was used to identify encoding and retrieval activation sites. To examine regions showing activation during encoding, signal from the two scans following the onset of list presentation (scans 2 and 3) was selectively averaged and compared to the baseline control average. Likewise, to identify regions showing activation during recall, signal from the two acquisitions following the onset of the recall cue (scans 9 and 10) was averaged and compared to baseline. It should be noted that due to the relatively short duration of encoding and retrieval stages, the scans used for selective averaging in these stages may have contained residual effects from preceding task events.

The statistical maps resulting from the various stage-based contrasts were thresholded to a voxel-wise statistical probability of 0.0001, and were subjected to a clustering threshold of five contiguous voxels, to accommodate increased probabilities due to multiple tests (Forman *et al.*, 1995). Additional paired contrasts that directly compared encoding and recall to maintenance were conducted for each of the clustered regions identified in the contrasts of these conditions to baseline. A significance threshold of 0.05 was used for these ROI-based stage comparisons.

Analysis 2: Effects of Stimulus Condition on Storage and Rehearsal

Our next set of analyses focused on the effects of stimulus condition on regional activation. These analyses allowed us to address the separability of different brain regions involved in rehearsal and storage. Based on behavioral research using manipulations of stimulus condition to investigate the cognitive architecture of verbal maintenance, we postulated that the characteristics of regional processing could be inferred from the pattern of stimulus effects shown in maintenance regions. To examine stimulus condition effects, we conducted a regions-of-interest (ROI) analysis using the regions identified in the maintenance to baseline comparison. In this analysis, signal obtained from each subject during performance with each stimulus type was averaged across all voxels in a contiguous ROI, and for the last two images acquired during the maintenance interval (scans 6 and 7). The regional averages obtained from each subject were then normalized to a common mean. To test for a main effect of stimulus condition, we conducted a one-way mixed-model ANOVA, treating subject as a random factor and stimulus condition as a within-subjects factor. To reduce the probability that the results of this test would simply reflect differences in task difficulty, this analysis included data from only the three span-reducing conditions (phonologically similar words, three-syllable words, and pseudowords), which were approximately matched for difficulty. Based on a Bonferroni correction adjusting probabilities for the number of regions tested (10), regions showing a stimulus condition effect reaching a statistical probability below an alpha of 0.005 were considered significant. In regions demonstrating a significant main effect of stimulus condition, we conducted additional pair-wise *t*-tests (phonologically similar words versus three-syllable words, phonologically similar words versus pseudowords, and three-syllable words versus pseudowords) to further characterize the specific pattern of differences. For these pair-wise tests we used an alpha of 0.05.

Analysis 3: Effects of Difficulty on the Executive and Verbal Maintenance Systems

The next set of analyses addressed the sensitivity of brain regions to changes in task difficulty (as measured by recall performance) attributable to the operation of the verbal maintenance system. Again using the regions identified in the contrast of maintenance to baseline (maintenance regions), we conducted an ROI analysis that probed the data for significant overall effects of task difficulty. Cluster averaging, scan averaging (scans 6 and 7), and mean normalization were applied as had been done in analysis 2. However, for this analysis, we pooled data collected during trials with phonologically similar, three-syllable, and pseudoword items, and compared activation for these pooled, span-reducing ('difficult'), conditions to that for the short, phonologically distinct words ('easy'). This comparison was made via a matched samples *t*-test, pairing averages for the difficult and easy trials within subject. Regions showing activation differences with an alpha of 0.005 (reflecting

a Bonferroni correction of the significance threshold based on the 10 regions tested) were considered to show a significant difficulty effect.

Post Hoc Analyses

To address some specific limitations of our *a priori* analyses, we conducted three *post hoc* exploratory analyses. By virtue of having pooled data across stimulus types in our stage-based analyses, regions activated by only a subset of the stimulus conditions may have been overlooked. To account for this potential limitation, in one *post hoc* test we performed a voxel-wise one-way mixed-model ANOVA with stimulus condition (including all four conditions) as the within-subject factor, for scans collected during maintenance (scans 6 and 7). This analysis allowed us to identify voxels that were active during maintenance for certain stimulus types, but not for others.

Inspection of activity in regions identified in our primary analyses further suggested that some regions may exhibit modulation by stimulus condition only during the early portion of a trial (during encoding and early maintenance). Since our *a priori* analyses did not probe stimulus differences during this interval, in a second *post hoc* analysis we conducted another voxel-wise ANOVA test for a main effect of stimulus condition during encoding and early maintenance (scans 2–4).

The time-series data derived from regions identified in the maintenance to baseline comparison showed that activation in some regions decreased across the maintenance interval. Again, our *a priori* analyses offered no way to characterize or quantify the significance of this pattern. Consequently, in a third *post hoc* test we performed an additional ROI-based analysis using the maintenance regions. In this analysis, for each ROI being tested, the data from each subject was pooled across stimulus conditions and was normalized across scans to a common mean. Using an ANOVA test, we investigated the pooled and normalized data in these ROIs for a main effect of time (change in signal magnitude across scans 4–7), and for an interaction between time (scans 4–7) and region.

Results

Behavioral Data

Robust and consistent stimulus-based effects were produced with the delayed serial recall task. For three subjects background noise from the magnet made coding of responses unreliable, so behavioral data from these subjects were excluded. Across the remaining subjects, statistically significant [$P < 0.005$, $t(8) > 3.36$, one-tailed] reductions in memory span were associated with the phonologically similar (mean = 3.3, SD = 1.4), three-syllable (mean = 3.8, SD = 1.5), and pseudoword (mean = 3.0, SD = 1.2) items, as compared to the one-syllable phonologically distinct (mean = 4.6, SD = 1.8) items (Fig. 2). These results were highly reliable, with all subjects showing span reductions for each of the three ‘difficult’ stimulus types. A further contrast comparing performance across the span-reducing, ‘difficult,’ conditions to that for the ‘easy’ condition also produced a highly significant ($P < 0.001$) behavioral result (difficult mean = 3.3, SD = 0.72). Further contrasts between the phonologically similar, three-syllable, and pseudoword items were non-significant (based on Bonferroni adjusted significance thresholds for the three paired comparisons). However, the contrast between the three-syllable and pseudoword conditions did show an unpredicted trend toward significance ($P = 0.043$, not surpassing Bonferroni corrected threshold).

Imaging Data

Maintenance, Encoding and Retrieval Regions

Each of the paired voxel-wise comparisons conducted to identify stage-based processing produced several regions that surpassed the probability ($P < 0.0001$, $t(11) > 5.45$, one-tailed) and clustering (five contiguous voxels) criteria. Table 1 summarizes results

from these comparisons. The maintenance to baseline comparison revealed several regions exhibiting significant activation through the end of the maintenance interval, including multiple bilateral inferior frontal areas, the supplementary motor area (SMA), the left premotor cortex, the right dorsolateral prefrontal cortex (DLPFC) (BA 9/46), and bilateral caudate regions (see Fig. 3A). Absent from this set of regions were the inferior parietal cortex and the cerebellum, which have both been frequently reported in neuroimaging studies of verbal working memory. Time-series showing the pattern of activity for a subset of the regions identified in this analysis are shown in Figure 4.

The comparison of encoding to baseline revealed activation in almost all of the regions identified in the maintenance to baseline comparison, as well as in some additional areas. As in the maintenance comparison, bilateral inferior frontal, SMA, premotor, and basal ganglia activations were found for this encoding comparison. New regions also identified in this comparison were the right premotor (BA 6), anterior cingulate (BA 32/24), left inferior parietal (BA 40), bilateral extrastriate (BA 18), and bilateral cerebellar regions. A direct comparison of encoding to maintenance produced a set of regions that could be predicted from the contrasts of maintenance and encoding to baseline (see Table 1).

Many of the regions implicated in maintenance were also found to be active during overt recollection, such as the premotor cortex, the SMA, and bilateral inferior frontal gyri (BA 45). The retrieval stage comparison further revealed activations in the anterior cingulate cortex and bilateral cerebellar regions, areas that were also identified in encoding. New regions specifically active during recall were the bilateral primary motor, bilateral somatosensory, bilateral superior temporal (BA 41/42/22), precuneus, medial cerebellar, and thalamic regions. Again, the direct contrast of recall with maintenance produced a set of regions that could be predicted from the contrasts of maintenance and recall to baseline (see Table 1).

Effects of Stimulus Condition

The pattern of stimulus condition effects produced within the maintenance regions was assessed using only the span-reducing stimulus conditions (three-syllable words, phonologically similar words, and pseudowords). The results of this analysis, which are summarized in Table 2, were intended to identify distinct components of the verbal maintenance subsystem. Two distinct subregions within the left inferior frontal gyrus showed large effects of stimulus condition [$F(2,33) > 6.24$, $P < 0.001$]. One of these subregions was located in the dorsal portion of the inferior frontal gyrus (BA 44), while the other was located in a more ventral portion of the inferior frontal gyrus (BA 45). These ROIs correspond closely with two dissociable inferior frontal foci identified through meta-analysis of prior neuroimaging studies of working memory (Chein *et al.*, 2001). In both regions, stimulus condition effects were shown to be driven by greater activation for pseudoword items than for long or phonologically similar items; a lexicality effect. However, in the more dorsal region, levels of activity associated with each condition were further graded, and ordered with subject performance levels. In the SMA, the overall stimulus condition effect also passed our Bonferroni adjusted alpha criterion [$F(2,33) = 5.69$, $P = 0.002$]. Paired comparisons for each of the difficult conditions showed significantly larger activations in SMA for the pseudoword and phonologically similar word conditions when compared to the long word condition. No other regions identified in the maintenance to baseline comparison showed a significant effect of stimulus condition (none showed a statistical probability below

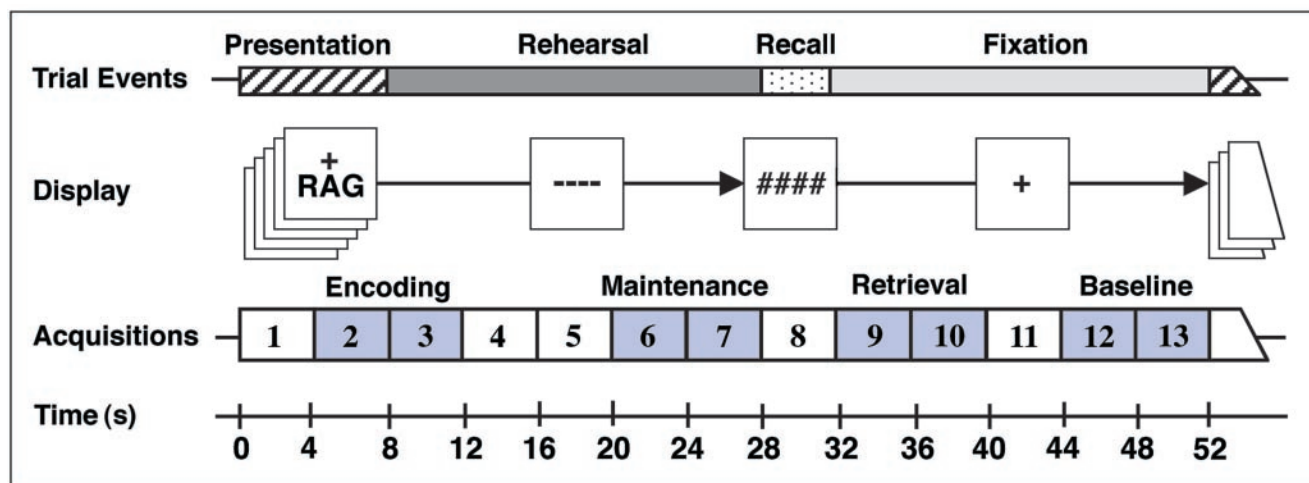


Figure 1

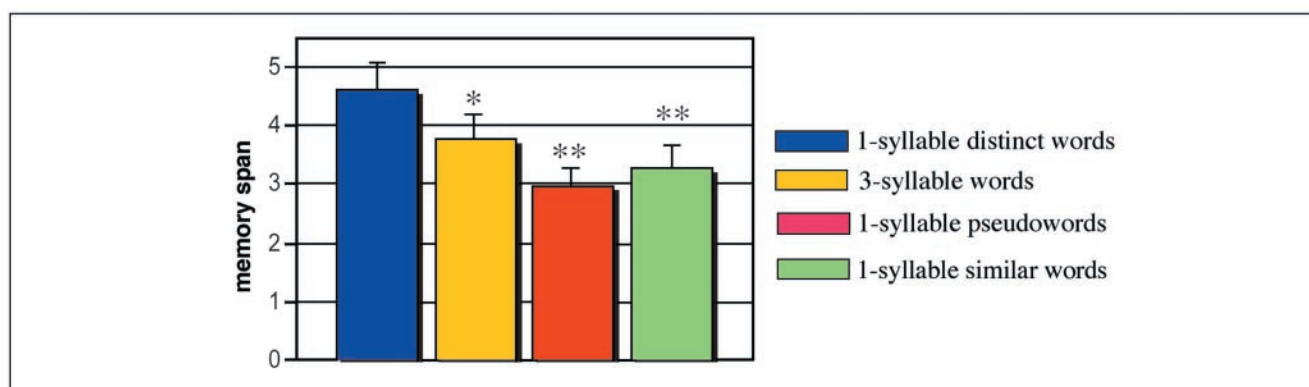


Figure 2

Figure 1. Schematic diagram of the experimental protocol, showing the sequence and timing for trial events and data acquisition. Each trial began with the presentation of five items, followed by a 20 s rehearsal interval. During the rehearsal interval, subjects maintained the presented items by covert rehearsal. At the end of the rehearsal interval, subjects were cued to recall the items in their presented order, and were given 4 s for overt recollection. In the final 20 s of each trial, subjects rested while viewing a fixation cross. Thirteen volumetric image acquisitions, each lasting 4 s, were collected during each trial. Acquisitions shown with shading comprised the specific epochs used for stage-based analyses, and are labeled accordingly.

Figure 2. Average trial span measures for one-syllable phonologically distinct words, three-syllable words, pseudowords, and phonologically similar words. Statistically significant span reductions were associated with three-syllable words, pseudowords and phonologically similar words, when compared to one-syllable phonologically distinct words. * $P < 0.01$, ** $P < 0.001$.

$P = 0.005$). However, weak effects of stimulus condition ($P < 0.05$, but above 0.005) were identified in left premotor (BA 6) and right inferior frontal (BA 45) regions.

Effects of Difficulty

In our second ROI analysis we investigated the effects of task difficulty, also using regions identified in the maintenance to baseline comparison. The results from these analyses are included in Table 2. In this analysis, several regions were found to be more active during the maintenance interval for trials with difficult stimuli than with the easier stimuli. Specifically, this effect was statistically significant in the dorsal left inferior frontal gyrus (BA 44), the SMA (BA 6), the left premotor cortex (BA 6), and the basal ganglia. Additional bilateral ventral inferior frontal regions showed weak difficulty effects ($0.005 < P < 0.05$). Although the difficulty effect in the right DLPFC (BA 46) showed a trend toward significance ($P = 0.063$), the magnitude of the difference in this region was much smaller than in other regions

where a significant difficulty effect was identified. None of the identified maintenance regions were more active for 'easy' trials than for 'difficult' trials.

Results from Post Hoc Analyses

Our maintenance versus baseline comparison, which collapsed data across stimulus types, did not reveal activation in the inferior parietal cortex. However, a voxel-wise analysis looking for a main effect of stimulus condition during maintenance, showed that the inferior parietal region (bilaterally) was active for some stimulus conditions, but not for others. Specifically, the left inferior parietal cortex was active for phonologically similar and pseudoword conditions, but not for three-syllable and one-syllable phonologically distinct conditions (see Fig. 3B). This pattern of results accounts for the lack of an overall (pooled across all stimulus conditions) maintenance-based activation in parietal cortex.

The analysis of stimulus condition effects during encoding and

Table 1

Brain areas showing increases for maintenance, encoding and retrieval relative to baseline

Regions of interest	BA	Maintenance <i>x, y, z</i>	Encoding <i>x, y, z</i>	Retrieval <i>x, y, z</i>	(E)ncoding versus (M)aintenance	(R)etrieval versus (M)aintenance
Maintenance, encoding and retrieval						
Medial frontal g.	6	0, 5, 49	1, 6, 49	0, 4, 50		
L. precentral g.	6	-32, -6, 42	-42, -8, 53	-49, -11, 37	E > M	
L. inferior frontal g.	45	-42, 14, 5	-48, 16, 0	-49, 16, 3		R > M
R. inferior frontal g.	45	40, 21, 5	48, 21, 0	47, 17, 1		R > M
Maintenance and encoding						
L. inferior frontal g.	44	-37, 13, 24	-41, 2, 27		E > M	
L. caudate nucleus		-14, -5, 16	-4, 2, 9		E > M	
Maintenance only						
R. middle frontal g.	46	36, 33, 13				M > R
L. anterior insula	45	-28, 22, 9				
R. inferior frontal g.	45	34, 17, 20				
R. caudate nucleus		19, -11, f24			M > E	
Encoding and retrieval						
Anterior cingulate g.	24,32		0, 2, 27	0, 18, 36	E > M	R > M
L. lateral cerebellum			-32, -49, -25	-21, -67, -23	E > M	R > M
R. lateral cerebellum			44, -55, -23	39, -68, -22	E > M	R > M
Encoding only						
L. inferior parietal	40,7		-27, -58, 43		E > M	
R. precentral g.	6		47, 10, 27		E > M	
L. inferior occipital g.	18,17		-21, -90, -11		E > M	
R. inferior occipital g.	18,17		23, -94, -8		E > M	
Retrieval only						
L. precentral g.	6			-57, 2, 17		R > M
R. precentral g.	4,6			42, -11, 36		R > M
L. thalamus				6, -11, 16		R > M
L. sup. temporal g.	42			-53, -21, 10		R > M
	41			-38, -32, 11		R > M
	22			-62, -1, 3		R > M
R. sup. temporal g.	42,22			47, -23, 6		R > M
Medial cerebellum				0, -54, -8		R > M

L., left; R., right; sup., superior; g., gyrus.

early maintenance scans produced additionally informative results. This analysis showed that during this early phase, sites in the left and right lateral cerebellum were more highly activated for the three-syllable word condition than for the other conditions (see Fig. 3C). These cerebellar regions did not appear in the maintenance to baseline comparison because activity in these regions was not present for all stimulus conditions, and was comparatively transitory.

Analysis of time-based effects during the maintenance interval provided further information about the dynamics of regional processing. This analysis showed that in BA 44 of the dorsal left inferior frontal cortex decreases in the magnitude of activation across the maintenance interval were statistically significant [$F(3,11) = 37.9, P < 0.001$]. This profile of activation could be contrasted with that in other regions, such as the ventral left inferior frontal gyrus (BA 45) [$F(3,11) = 0.28, P > 0.05$], the medial frontal gyrus (SMA, BA 6) [$F(3,11) = 1.14, P > 0.05$] and

the middle frontal gyrus (DLPFC, BA 46) ($F(3,11) = 1.85, P > 0.05$), where levels of activation were sustained throughout the maintenance interval (see Fig. 4). Accordingly, paired ANOVA tests for interactions between region (dorsal inferior frontal, ventral inferior frontal, medial frontal, middle frontal) and time (scans 4–7) yielded significant interactions only when the *dorsal* inferior frontal ROI was contrasted with the other ROIs [$F(3,11) > 8.76, P < 0.05$].

Negative Activation Regions

The results that we have reported above focus specifically on those regions for which active conditions produced an increased signal compared to the baseline condition. There were, however, several regions that consistently demonstrated a signal decrease for the active, as compared to baseline, states. When baseline was compared to encoding, maintenance, and retrieval, each of these comparisons showed negative changes in signal magnitude

Figure 3. Regions of activity resulting from the (A) *a priori* voxel-based comparison of maintenance (scans 6 and 7) to baseline (scans 12 and 13), (B) *post hoc* voxel-based test for stimulus condition effects across the maintenance interval (scans 4–7) and (C) *post hoc* voxel-based test for stimulus condition effects in encoding and early maintenance (scans 2–4). Voxels shown surpass a statistical probability threshold of $P < 0.0001$ in the group composite data, and comprise a cluster of at least five contiguous significant voxels. The statistical images are shown overlaid onto horizontal sections, at 2.5 mm increments, from the structural MRI image used for inter-subject registration, which has been warped into stereotaxic space (Talairach, 1988). The numbers to the left and right of each row represent the distance, in millimeters, above the anterior commissure to posterior commissure plane. Peak activation locations for regions detailed in Figure 4 are identified with labels 'a'–'f'.

Figure 4. Activation profiles for regions commonly associated with verbal working memory. Regions identified in the present study in the maintenance to baseline comparison included, a = medial frontal gyrus, BA 6 (SMA); b = dorsal left inferior frontal gyrus, BA 44; c = ventral left inferior frontal gyrus, BA 45; and d = left middle frontal gyrus, BA 46 (DLPFC). *Post hoc* testing for stimulus condition effects during maintenance produced an ROI in the (e) left inferior parietal lobule, while a *post hoc* test for stimulus condition effects in encoding and early maintenance produced an ROI in the (f) cerebellum. From left to right, the shaded panels in the background of each graph represent the encoding, maintenance, retrieval and baseline epochs of data acquisition used in the stage-based analyses. The bar running horizontally across the bottom of each graph tracks the timing of presentation, rehearsal, recall and fixation trial events, respectively.

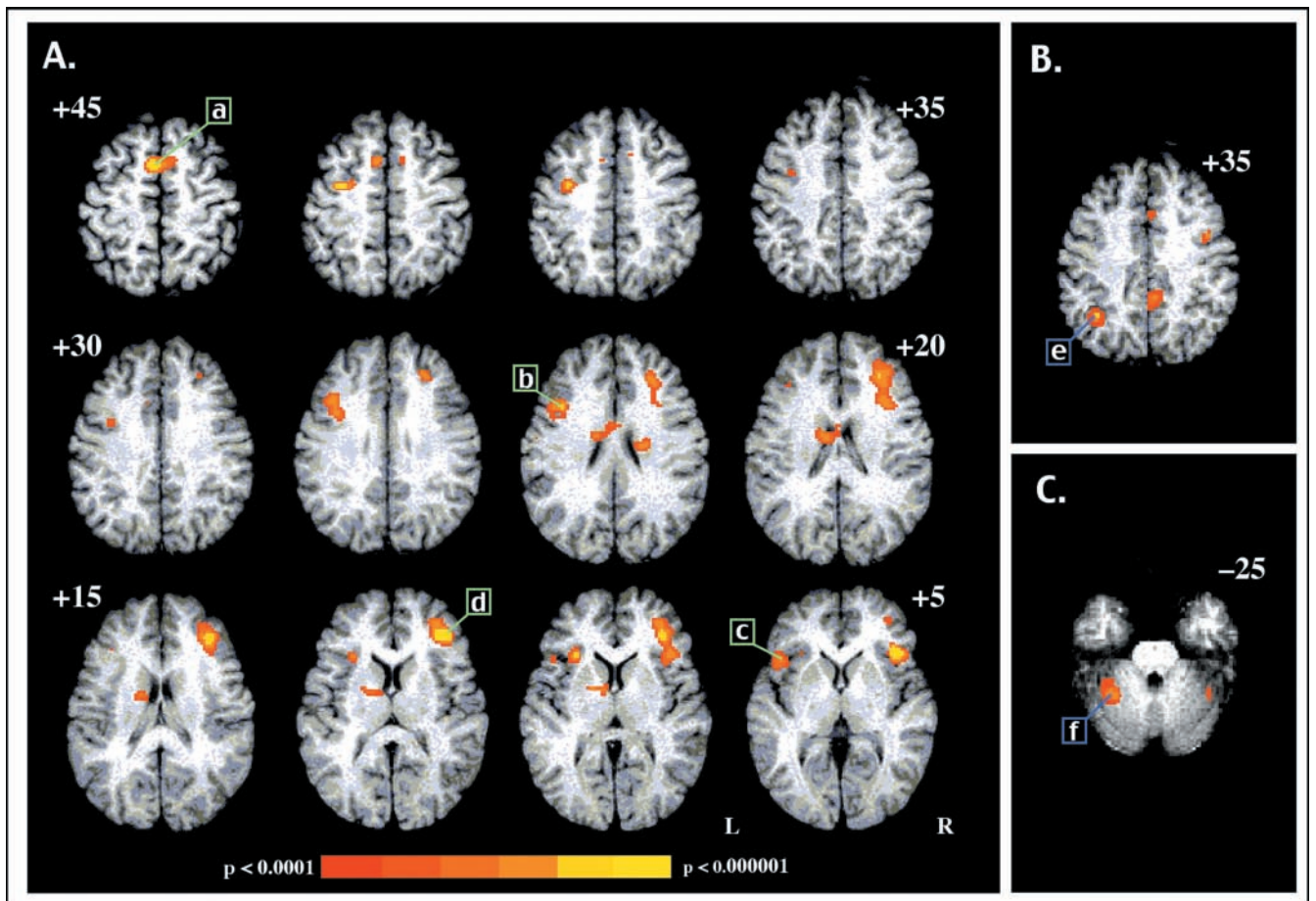


Figure 3

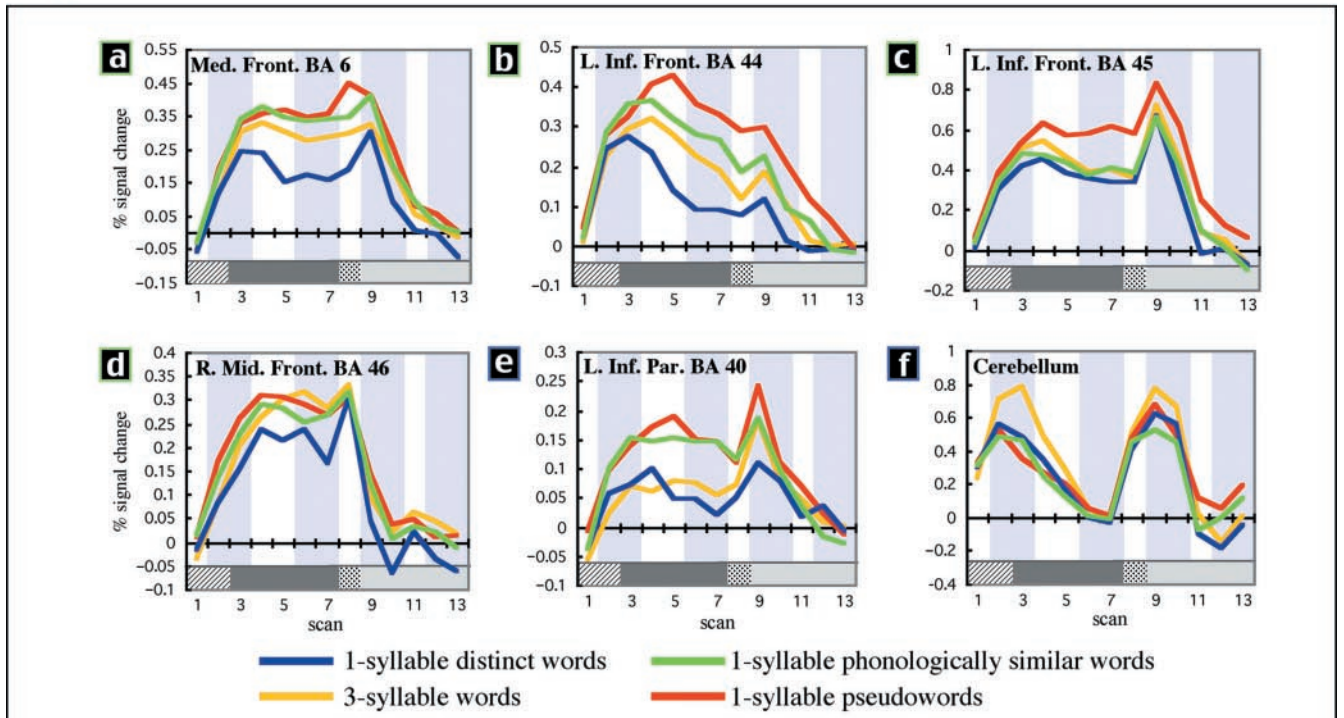


Figure 4

Table 2
Maintenance regions sensitive to difficulty and stimulus condition manipulations

Regions	BA	x, y, z	Difficulty		Stimulus condition		
			t-test	ANOVA	t-test		
					Sm versus Lg	Pw versus Lg	Pw versus Sm
Specific effect of stimulus condition							
Medial frontal g.	6	0, 5, 49	**	**	Sm > Lg	Pw > Lg	
L. inferior frontal g.	44	-37, 13, 24	**	**		Pw > Lg	Pw > Sm
	45	-42, 14, 5	*	**		Pw > Lg	Pw > Sm
Non-specific effect of difficulty							
L. precentral g.	6	-32, -6, 42	**	*			
R. inferior frontal g.	45	40, 21, 5	*	*			
	45	34, 17, 20	*				
L. caudate nucleus		-14, -5, 16	**				

Sm = One-syllable phonologically similar words, Lg = three-syllable words, Pw = one-syllable pseudowords.

* $P < 0.05$ (above Bonferroni corrected threshold).

** $P < 0.005$ (below Bonferroni corrected threshold).

in a large region around posterior cingulate cortex (BA 31) (-1, -53, 27), in a region near the junction of the temporal, occipital, and parietal lobes bilaterally (BA 39, 19) (-34, -80, 28; 50, -61, 23), and in the rostral cingulate (BA 32) (2, 44, 8). These regions showing negative activations are consistent with those documented in a review of negative activations found in PET imaging studies (Shulman *et al.*, 1997).

Discussion

The results from our comparison of maintenance to baseline largely replicate findings from prior neuroimaging studies of verbal working memory, although there were some notable differences. Consistent with previous findings, we have implicated a network of regions in the frontal cortex, including the DLPFC, the inferior frontal gyrus, the supplementary motor area, and the premotor cortex, in verbal working memory processing. Among the regions not identified in our maintenance to baseline comparison, but implicated in previous studies, were the left inferior parietal cortex and the cerebellum. In the following discussion we consider results from the present study together with previous findings, and show how these findings can be synthesized into a modified account of the nature of verbal maintenance and its neural underpinnings.

The Involvement of Maintenance Regions in Encoding and Retrieval

In the present study we were interested in determining if regions commonly associated with maintenance-based processing are also integral to encoding and retrieval processes, in the context of a working memory task. As expected, the encoding and retrieval analyses revealed some areas of activation that were distinct from those identified in maintenance, primarily in regions associated with sensory and motor demands of task performance. For example, we found activation during encoding in a region centered in extrastriate cortex, which most likely reflected early visual processing of the presented stimuli.

Perhaps more interesting is the high degree of correspondence in brain regions engaged during encoding, maintenance, and retrieval. Across these stages of task performance, a core network of brain regions including the SMA, premotor cortex, and bilateral inferior frontal regions was activated. These findings partially replicate results from two previous verbal working memory studies, in which both medial and lateral prefrontal

areas showed encoding-, maintenance-, and retrieval-based activity (Rypma and D'Esposito, 1999; Kruggel *et al.*, 2000). There are at least two contrasting ways to explain this correspondence in stage-based regional activation. One explanation is that these brain areas are specialized for *maintenance-based* processes that are also engaged during encoding and retrieval stages of task performance. For example, covert maintenance may continue throughout the recall stage, to insure that the contents of working memory are not lost before recollection is completed. An alternative explanation is that these brain regions mediate computational processes (e.g. phonological processing, accessing of long-term memory) that are recruited recurrently across encoding, maintenance, and retrieval. Because neuroimaging studies of reading and phonological processing reveal many of the same brain areas as are typically found in verbal maintenance (Price *et al.*, 1996; Rumsey *et al.*, 1997; Fiez and Petersen, 1998), we favor the latter alternative. This is also consistent with results from a prior PET study conducted by Jonides and colleagues (Jonides *et al.*, 1998). Although they reported little evidence for common regions of activation between encoding, maintenance, and retrieval, the control condition that they used to identify encoding and retrieval regions (but not maintenance regions) involved viewing nonwords, a control task that may recruit the same shared processes.

Characterizing the Components of Verbal Maintenance

It has been argued that verbal maintenance may be mediated by separate storage and rehearsal components. To examine the dissociability of these components we investigated the specific patterns of stimulus condition effects and the temporal profiles of regional activity in maintenance-related regions.

Storage in the Verbal Maintenance Subsystem

Inspection of the neuroimaging literature produces only one candidate region as the site of phonological storage, the left inferior parietal cortex (Awh *et al.*, 1996; Jonides *et al.*, 1998; Henson *et al.*, 2000). However, in review articles (Becker *et al.*, 1999; Fiez, 2001) examining both neuroimaging and neuropsychological evidence, we have argued that the inferior parietal region may contain two distinct subregions contributing to maintenance (dorsal and ventral), and that neither one of these subregions behaves in a way that is consistent with the phonological store, as it is portrayed in the Baddeley model. For

instance, the Baddeley model suggests that the phonological store is refreshed repeatedly during extended maintenance. We should therefore expect that the associated brain region would be active across a maintenance interval for any phonologically codable material. However, prior PET studies using task designs similar to that used in the present study (Fiez *et al.*, 1996; Jonides *et al.*, 1998) have failed to find reliable left parietal activation during maintenance. Our current study, which failed to produce significant activity in this region for the pooled (across stimuli) maintenance to baseline comparison, further supports these previous findings. These null findings should be interpreted cautiously, as they may have resulted from limitations in statistical power. However, in the present study, several other areas showed sufficient effects to be detected during maintenance, and the left inferior parietal region produced detectable activation during encoding under the same power limitations.

We had additionally expected that if the store uses phonological representations, and the phonological similarity effect is due to interference between these representations, then the site of storage would show a particular sensitivity to phonological similarity among list items. The pattern of results found in the left inferior parietal region was only partially consistent with this expectation. Although activity in the left inferior parietal region was high for phonologically similar items, it was also high for the pseudoword items. This pattern of activity may still be consistent with that of a phonological store if, as suggested by Jonides and colleagues (Jonides *et al.*, 1998), we interpret high pseudoword trial activity as a reflection of heavy reliance on phonological storage when semantic representations are lacking. Considered in its entirety, the available evidence does not allow us to draw strong conclusions regarding the role of the inferior parietal region in phonological storage, but incites us to consider alternative interpretations of parietal activity in working memory tasks.

One interesting clue to inferior parietal function is that, along with the present study, all of the verbal working memory studies that have failed to find significant activation in the left inferior parietal cortex (Becker *et al.*, 1993; Grasby *et al.*, 1993; Fiez *et al.*, 1996; Jonides *et al.*, 1998) have utilized recall, rather than recognition-based (in which the contents of memory are assessed by recognition of a probe stimulus; e.g. N-back, delayed match-to-sample) paradigms. One potentially relevant aspect of recognition-based tasks is that they can be successfully performed by either a prospective (e.g. active rehearsal) or retrospective (e.g. familiarity) strategy (Lovett *et al.*, 1999), while recall-based tasks are likely to be performed primarily through active rehearsal. Inferior parietal activation may therefore reflect the use of retrospective strategies, or the utilization of an executive process coordinating the use of alternative strategies. This latter interpretation is consistent with suggestions that the left inferior parietal region may serve an executive, rather than storage-based, role in working memory (Collette *et al.*, 1999; Marshuetz *et al.*, 2000).

Rehearsal in the Verbal Maintenance Subsystem

Numerous previous studies have linked left inferior frontal, supplementary motor, premotor, and cerebellar regions to articulatory rehearsal. We had predicted that regions involved in rehearsal would demonstrate sensitivity to stimulus factors, including articulatory length and lexicality, that can influence the articulatory control process. Of the regions we found to be active during maintenance, two left inferior frontal regions (dorsal BA 44 and ventral BA 45) and the SMA (medial BA 6) showed significant stimulus condition effects potentially linking

them with rehearsal. Interestingly, the pattern of these stimulus condition effects was distinct across regions, as were the effects of difficulty. Specifically, sensitivity to lexical status dissociated the left inferior frontal regions from the SMA, while sensitivity to difficulty dissociated dorsal BA 44 and the SMA from ventral BA 45 (which exhibited only a marginally significant difficulty effect that was specifically driven by increases during pseudo-word trials). Additionally, while the dorsal inferior frontal region (BA 44) showed a significant decrease in activity across the maintenance interval [also observed by Cohen *et al.* (Cohen *et al.*, 1997)], the ventral inferior frontal (BA 45) and supplementary motor regions both showed sustained levels of activation throughout maintenance.

Given the differential patterns of response exhibited by these regions, we can speculate about their specific contributions to rehearsal. In doing so, we refer to a two-stage model of rehearsal proposed by Naveh-Benjamin and Jonides (1984), which nicely accommodates our results. According to their model, during a first stage of rehearsal a novel articulatory rehearsal program is effortfully assembled and initiated. In a second stage, this rehearsal program is automatically and repetitively executed.

The region in left dorsal BA 44 was the only region to exhibit a decreasing level of activity across the delay interval. This decreasing pattern of activity suggests that this region may contribute specifically to the first, organizational, stage of rehearsal. Further, this region's sensitivity to difficulty is consistent with the proposal that this first stage is effortful. Lexicality effects in this region can be explained as the consequence of additional organizational demands associated with the assembly of novel pseudoword items (which lack prior articulatory representations) into a subvocal rehearsal program.

The ventral inferior frontal (BA 45) region exhibited sustained activation across the delay, and a significant lexicality effect. This demonstration of lexicality effects in ventral left inferior frontal cortex replicates our previous PET results (Fiez *et al.*, 1996), and is consistent with the results from Jonides *et al.* (Jonides *et al.*, 1998). Functional imaging studies of single-word reading have also reported activation localized in left inferior frontal regions to be greater for pseudowords than for words (Herbster *et al.*, 1997; Fiez *et al.*, 1999; Hagoort *et al.*, 1999). One interpretation of this pattern of results is that this region contributes to sub-lexical processes, most heavily utilized during pseudoword trials, used throughout rehearsal for coding and storage of the articulatory rehearsal program.

The SMA has been previously linked with the automatic implementation of internally cued sequential motor gestures (Goldberg, 1985; Jenkins *et al.*, 1994). This association with sequential motor processes, and the exhibition of sustained activity throughout maintenance, suggest that the SMA may be particularly instrumental in mediating the second stage of rehearsal, involving the repetitive execution of the coded and assembled articulatory rehearsal program (although, this fails to explain why the SMA showed sensitivity to task difficulty, given that this latter stage of rehearsal is thought to be automatic).

In several previous studies, the cerebellum was also found to be active during verbal maintenance. Desmond and colleagues (Desmond *et al.*, 1997) have further demonstrated that cerebellar regions may be sensitive to the amount of information being maintained. Such findings are consistent with the notion that cerebellar activity in a working memory task reflects motor imagery associated with subvocal rehearsal (Decety *et al.*, 1990; Thach, 1996). Based on this motor-imagery interpretation, we expected that in the present study cerebellar activity would be greatest for the three-syllable words. This expectation

derives from the presumed relationship between the duration of physical articulation and the magnitude of the associated motor imagery. Indeed, in bilateral lateral cerebellar regions we found increased signal for the three-syllable words. However, activity in these cerebellar regions was temporally limited to scans acquired immediately following encoding, and did not persist into the maintenance interval. Our results are therefore partially inconsistent with what we had expected to find, and stand in contrast to previous results showing activity of the cerebellum during maintenance. One possible explanation for this inconsistency is that prior studies may over-emphasize cerebellar contributions to maintenance by not having sufficiently separated maintenance from encoding or retrieval (which strongly activated the cerebellum in the present study). Additionally, the fact that cerebellar activity persists briefly during maintenance could explain why PET studies, which weight heavily on the beginning of the acquisition period, have detected activity in this area.

The Effects of Difficulty on the Central Executive

Previous research has shown that increases in task difficulty can influence activity in brain regions associated with the central executive. We were interested in examining how a manipulation of task difficulty associated specifically with verbal stimuli might influence activity in executive regions. Our analyses of the effects of difficulty and stimulus condition revealed several areas that showed a non-specific effect of difficulty (an effect of difficulty, but not an effect of length, lexicality or similarity). These areas included the left premotor cortex, the basal ganglia, and the right inferior frontal gyrus. However, given that most neuroimaging studies of executive processing have focused on the properties of the DLPFC, we were surprised to find that this region showed only a weak and statistically insignificant effect of difficulty.

The lack of a difficulty effect in DLPFC can be interpreted in the context of previous studies of executive processing. For example, studies using manipulations of memory load (Braver *et al.*, 1997; Cohen *et al.*, 1997; Rypma *et al.*, 1999) have implicated the DLPFC as a component of the central executive. Load effects, however, are necessarily confounded with performance factors (i.e. it is more difficult to keep more items in memory). This confounding makes it hard to determine if regions showing modulation in these studies are sensitive specifically to load, or more generally to levels of difficulty. In a study by Barch and colleagues (Barch *et al.*, 1997), it was shown that varying the length of a maintenance delay interval, a manipulation of difficulty that is not based on memory load, influenced the magnitude of DLPFC activation. Barch and colleagues further demonstrated that modulation of DLPFC activity was specifically associated with manipulation of mnemonic task demands (degradation of stimuli did not impact on DLPFC activation). In the present study, with memory load held constant while difficulty was manipulated, we did not find significant difficulty effects in DLPFC. One possible way to reconcile our findings with those from Barch *et al.* (Barch *et al.*, 1997) is to suppose that while domain-independent mnemonic factors that effect difficulty (such as the manipulation of delay duration) influence DLPFC activity, this region is minimally sensitive to verbal domain-specific factors that effect difficulty (such as those employed in the present study). The insensitivity of this region to differences among verbal stimuli further supports the view that this region plays a more prominent role in non-verbal processes (Goldman Rakic, 1998).

Relating Imaging Results to the Cognitive Model

Based on our results, we formulate a tentative theory of verbal maintenance that is consistent with existing evidence, but that contrasts in some subtle but potentially important ways from the dominant theoretical framework of verbal working memory. One distinction relates to the buffering of information into the phonological store. As discussed above, although the inferior parietal cortex remains a possible center for storage, no brain region behaves in such a way that is entirely consistent with the notion of storage as it is conceived in the Baddeley model. The lack of a strong candidate for storage across a delay interval points to a potential shortcoming of the behavioral model. An alternative account, in which phonological encoding is a temporally limited event that allows input to be integrated into a speech-output based rehearsal stream that ultimately supports both rehearsal and storage, may be more consistent with the available evidence.

Our results further suggest a need to extend the characterization of verbal rehearsal beyond its portrayal in Baddeley's model. These results suggest that articulatory rehearsal should not be thought of as a single component operating as a continuous loop, but that it may be further fractionated into subcomponents that show changing levels of engagement as rehearsal proceeds. Recognizing that rehearsal is based on multiple subprocesses may facilitate interpretation of behavioral, neuropsychological, and neuroimaging evidence.

Lastly, investigation of the effects of task difficulty yielded further insights into the nature of the interaction between executive and maintenance systems. Our failure to find robust difficulty effects in the DLPFC leads us to speculate that this region may contribute to domain-independent executive processes such as response inhibition and attentional selection, but may not support domain-specific executive processes integral to the operation of the verbal maintenance subsystem. This interpretation fits with the notion that the central executive system embodied in Baddeley's model may be further fractionated (Baddeley, 1998), and suggests that domain-specificity is one dimension upon which such fractionation is contingent.

Notes

This work was supported by a grant from the National Institute of Mental Health, MH59256. The authors wish to thank Kate Fissell for technical support and advice, and Steve Wilson for assistance in behavioral data collection and analysis.

Address correspondence to Jason M. Chein, 604 LRDC Building, Department of Psychology, University of Pittsburgh, Pittsburgh, PA 15260, USA. Email: jchein@pitt.edu.

References

- Awl E, Jonides J, Smith EE, Schumacher EH, Koeppel RA, Katz S (1996) Dissociation of storage and rehearsal in verbal working memory: evidence from positron emission tomography. *Psychol Sci* 7:25-31.
- Baddeley AD (1966) Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Q J Exp Psychol* 18: 362-365.
- Baddeley AD (1986) Working memory. Oxford: Oxford Science Publications.
- Baddeley AD (1998) Recent developments in working memory. *Curr Opin Neurobiol* 8:234-238.
- Baddeley AD, Logie RH (1999) Working memory: the multiple-component model. In: Models of working memory: mechanisms of active maintenance and executive control (Miyake A, Shah P, eds), pp. 28-61. New York: Cambridge University Press.
- Baddeley AD, Thomson N, Buchanan M (1975) Word length and the structure of short-term memory. *J Verbal Learn Verbal Behav* 14: 575-589.

- Baddeley AD, Lewis V, Vallar G (1984) Exploring the articulatory loop. *Q J Exp Psychol: Hum Exp Psychol* 36a:233-252.
- Barch DM, Braver TS, Nystrom LE, Forman SD, Noll DC, Cohen JD (1997) Dissociating working memory from task difficulty in human prefrontal cortex. *Neuropsychologia* 35:1373-1380.
- Becker JT, Mintun MA, Diehl DJ, Dobkin J, Martidis A, Madoff DC, DeKosky ST (1993) Functional neuroanatomy of verbal free recall: a replication study. *Hum Brain Mapp* 1:284-292.
- Becker JT, MacAndrew DK, Fiez JA (1999) A comment on the functional localization of the phonological storage subsystem of working memory. *Brain Cogn* 41:27-38.
- Braver TS, Cohen JD, Nystrom LE, Jonides J, Smith EE, Noll DC (1997) A parametric study of prefrontal cortex involvement in human working memory. *Neuroimage* 5:49-62.
- Bunge SA, Klingberg T, Jacobsen RB, Gabrieli JD (2000) A resource model of the neural basis of executive working memory. *Proc Natl Acad Sci USA* 97:3573-3578.
- Cabeza R, Nyberg L (2000) Imaging cognition. II. An empirical review of 275 PET and fMRI studies. *J Cogn Neurosci* 12:1-47.
- Chein JM, Fissell K, Jacobs S, Fiez JA (2001) Functionally distinct subdivisions within Broca's area during verbal working memory: results from a novel method for meta-analysis of neuroimaging studies. *Soc Neurosci Abstr* (in press).
- Cohen JD, Perlstein WM, Braver TS, Nystrom LE, Noll DC, Jonides J, Smith EE (1997) Temporal dynamics of brain activation during a working memory task. *Nature* 386:604-608.
- Collette F, Salmon E, Van der Linden M, Chicherio C, Belleville S, Degueldre C, Delfiore G, Franck G (1999) Regional brain activity during tasks devoted to the central executive of working memory. *Cogn Brain Res* 7:411-417.
- Coltheart V (1993) Effects of phonological similarity and concurrent irrelevant articulation on short-term-memory recall of repeated and novel word lists. *Mem Cogn* 21:539-545.
- Conrad R, Hull AJ (1964) Information, acoustic confusion and memory span. *Br J Psychol* 55:429-432.
- Cox RW (1996) AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res* 29:162-173.
- Decety J, Sjoeholm H, Ryding E, Stenberg G, Ingvar DH (1990) The cerebellum participates in mental activity: tomographic measurements of regional cerebral blood flow. *Brain Res* 535:313-317.
- Desmond JE, Gabrieli JD, Wagner AD, Ginier BL, Glover GH (1997) Lobular patterns of cerebellar activation in verbal working-memory and finger-tapping tasks as revealed by functional MRI. *J Neurosci* 17:9675-9685.
- D'Esposito M, Detre JA, Alsop DC, Shin RK, Atlas S, Grossman M (1995) The neural basis of the central executive system of working memory. *Nature* 378:279-281.
- D'Esposito M, Aguirre GK, Zarahn E, Ballard D, Shin RK, Lease J (1998) Functional MRI studies of spatial and nonspatial working memory. *Brain Res Cogn Brain Res* 7:1-13.
- D'Esposito M, Postle BR, Ballard D, Lease J (1999) Maintenance versus manipulation of information held in working memory: an event-related fMRI study. *Brain Cogn* 41:66-86.
- Fiez JA (2001) Bridging the gap between neuroimaging and neuropsychology: using working memory as a case-study. *J Clin Exp Neuropsychol* 23:19-31.
- Fiez JA, Petersen SE (1998) Neuroimaging studies of word reading. *Proc Natl Acad Sci USA* 95:914-921.
- Fiez JA, Raife EA, Balota DA, Schwarz JP, Raichle ME, Petersen SE (1996) A positron emission tomography study of the short-term maintenance of verbal information. *J Neurosci* 16:808-822.
- Fiez JA, Balota DA, Raichle ME, Petersen SE (1999) Effects of lexicality, frequency, and spelling-to-sound consistency on the functional anatomy of reading. *Neuron* 24:205-218.
- 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.
- Francis WN, Kucera H (1982) Frequency analysis of english usage: lexicon and grammar. Boston: Houghton Mifflin.
- Gainotti G, Cappa A, Perri R, Silveri MC (1994) Disorders of verbal and pictorial memory in right and left brain-damaged patients. *Int J Neurosci* 78:9-20.
- Gathercole SE (1994) Neuropsychology and working memory: A review. *Neuropsychology* 8:494-505.
- Goldberg G (1985) Supplementary motor area structure and function: review and hypotheses. *Behav Brain Sci* 8:567-616.
- Goldman Rakic PS (1998) The prefrontal landscape: implications of functional architecture for understanding human mentation and the central executive. In: *The prefrontal cortex: executive and cognitive functions* (Roberts AC, Robbins TW, eds), pp. 87-102. New York: Oxford University Press.
- Grasby PM, Frith CD, Friston KJ, Bench C, Frackowiak RS, Dolan RJ (1993) Functional mapping of brain areas implicated in auditory-verbal memory function. *Brain* 116:1-20.
- Hagoort P, Indefrey P, Brown C, Herzog H, Steinmetz H, Seitz RJ (1999) The neural circuitry involved in the reading of German words and pseudowords: a PET study. *J Cogn Neurosci* 11:383-398.
- Henson RNA, Burgess N, Frith CD (2000) Recoding, storage, rehearsal and grouping in verbal short-term memory: an fMRI study. *Neuropsychologia* 38:426-440.
- Herbster AN, Mintun MA, Nebes RD, Becker JT (1997) Regional cerebral blood flow during word and nonword reading. *Hum Brain Mapp* 5:84-92.
- Hulme C, Maughan S, Brown GD (1991) Memory for familiar and unfamiliar words: evidence for a long-term memory contribution to short-term memory span. *J Mem Lang* 30:685-701.
- Jenkins IH, Brooks DJ, Nixon PD, Frackowiak RS, Passingham RE (1994) Motor sequence learning: a study with positron emission tomography. *J Neurosci* 14:3775-3790.
- Jonides J, Schumacher EH, Smith EE, Koeppel RA, Awh E, Reuter-Lorenz PA, Marshuetz C, Willis CR (1998) The role of parietal cortex in verbal working memory. *J Neurosci* 18:5026-5034.
- Kruggel F, Zysset S, von Cramon DY (2000) Nonlinear regression of functional MRI data: an item recognition task study. *Neuroimage* 12:173-183.
- Levelt WJM, Roelofs A, Meyer AS (1999) A theory of lexical access in speech production. *Behav Brain Sci* 22:1-75.
- Longoni AM, Richardson JT, Aiello A (1993) Articulatory rehearsal and phonological storage in working memory. *Mem Cognit* 21:11-22.
- Lovett MC, Reder LM, Lebiere C (1999) Modeling working memory in a unified architecture: an ACT-R perspective. In: *Models of working memory: mechanisms of active maintenance and executive control* (Miyake A, Shah P, eds), pp. 135-182. New York: Cambridge University Press.
- Marshuetz C, Smith EE, Jonides J, DeGutis JM, Krishnan SA, Chenevert TL, Noll DC (2000) fMRI evidence for prefrontal and parietal involvement in the sequencing of information in verbal working memory. *Soc Neurosci Abstr* 26:263-264.
- Multhaup KS, Balota DA, Cowan N (1996) Implications of aging, lexicality, and item length for the mechanisms underlying memory span. *Psychon Bull Rev* 3:112-120.
- Naveh Benjamin M, Jonides J (1984) Maintenance rehearsal: a two-component analysis. *J Exp Psychol Learn Mem Cogn* 10:369-385.
- Noll DC, Cohen JD, Meyer CH, Schneider W (1995) Spiral K-space MR imaging of cortical activation. *J Magn Reson Imaging* 5:49-56.
- Noll DC, Genovese CR, Nystrom LE, Vazquez AL, Forman SD, Eddy WF, Cohen JD (1997) Estimating test-retest reliability in functional MR imaging. II: Application to motor and cognitive activation studies. *Magn Reson Med* 38:508-517.
- Paulesu E, Frith CD, Frackowiak RS (1993) The neural correlates of the verbal component of working memory. *Nature* 362:342-345.
- Postle BR, Zarahn E, D'Esposito M (2000) Using event-related fMRI to assess delay-period activity during performance of spatial and nonspatial working memory tasks. *Brain Res Brain Res Protoc* 5:57-66.
- Price CJ, Wise RJ, Frackowiak RS (1996) Demonstrating the implicit processing of visually presented words and pseudowords. *Cereb Cortex* 6:62-70.
- Roe CM (1993) Presentation rate, articulation rate, word length, and working memory theory. *Dissert Abstr Int* 53:5468-5469.
- Rumsey JM, Horwitz B, Donohue BC, Nace K, Maisog JM, Andreason P (1997) Phonological and orthographic components of word recognition: a PET-rCBF study. *Brain* 120:739-759.
- 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.
- Rypma B, Prabhakaran V, Desmond JE, Glover GH, Gabrieli JD (1999)

- Load-dependent roles of frontal brain regions in the maintenance of working memory. *Neuroimage* 9:216-226.
- Shallice T, Warrington EK (1970) Independent functioning of verbal memory stores: a neuropsychological study. *Q J Exp Psychol* 22:261-273.
- Shulman GL, Fiez JA, Corbetta M, Buckner RL, Miezin FM, Raichle ME, Petersen SE (1997) Common blood flow changes across visual tasks. II. Decreases in cerebral cortex. *J Cogn Neurosci* 9:648-663.
- Smith EE, Jonides J (1995) Working memory in humans: neuropsychological evidence. In: *The cognitive neurosciences* (Gazzaniga MS, ed.), pp. 1009-1020. Cambridge, MA: The MIT Press.
- Smith EE, Jonides J (1998) Neuroimaging analyses of human working memory. *Proc Natl Acad Sci USA* 95:12061-12068.
- Smith EE, Jonides J (1999) Storage and executive processes in the frontal lobes. *Science* 283:1657-1661.
- Smith EE, Jonides J, Koeppel RA (1996) Dissociating verbal and spatial working memory using PET. *Cereb Cortex* 6:11-20.
- Smith EE, Jonides J, Marshuetz C, Koeppel RA (1998) Components of verbal working memory: evidence from neuroimaging. *Proc Natl Acad Sci USA* 95:876-882.
- Talairach J, Tournoux P (1988) *Co-planar stereotaxic atlas of the human brain*. New York: Thieme.
- Thach WT (1996) On the specific role of the cerebellum in motor learning and cognition: clues from PET activation and lesion studies in man. *Behavioral and Brain Sciences* 19:411-431, 503-527.
- Vallar G, Shallice T, eds (1990) *Neuropsychological impairments of short-term memory*. New York: Cambridge University Press.
- Van der Linden M, Coyette F, Seron X (1992) Selective impairment of the 'central executive' component of working memory: a single case study. *Cogn Neuropsychol* 9:301-326.
- Woods RP, Cherry SR, Mazziotta JC (1992) Rapid automated algorithm for aligning and reslicing PET images. *J Comput Assist Tomogr* 16: 620-633.
- Woods RP, Mazziotta JC, Cherry SR (1993) MRI-PET registration with automated algorithm. *J Comput Assist Tomogr* 17:536-546.