Decoding the internal focus of attention

Jarrod A. Lewis-Peacock [a,b,*], Bradley R. Postle [a]

[a] Department of Psychology, University of Wisconsin-Madison, United States
[b] Princeton Neuroscience Institute, Princeton University, United States

ARTICLE INFO

ABSTRACT

The significance of the recent introduction to cognitive neuroscience of multivariate pattern analysis (MVPA) is that, unlike univariate approaches which are limited to identifying *magnitudes of activity* in localized parts of the brain, it affords the detection and characterization of *patterns of activity* distributed within and across multiple brain regions. This technique supports stronger inferences because it captures neural representations that have markedly higher selectivity than do univariate activation peaks. Recently, we used MVPA to assess the neural consequences of dissociating the internal focus of attention from short-term memory (STM), finding that the information represented in delay-period activity corresponds only to the former (Lewis-Peacock, Drysdale, Oberauer, & Postle, in press). Here we report several additional analyses of these data in which we directly compared the results generated by MVPA vs. those generated by univariate analyses. The sensitivity of MVPA to subtle variations in patterns of distributed brain activity revealed a novel insight: although overall activity remains elevated in category-selective brain regions corresponding to unattended STM items, the multivariate patterns of activity within these regions reflect the representation of a *different* category, i.e., the one that is currently being attended to. In addition, MVPA was able to dissociate attended from unattended STM items in brain regions whose univariate activity did not appear to be sensitive to the task. These findings highlight the fallacy of the assumption of homogeneity of representation within putative category-selective regions. They affirm the view that neural representations in STM are highly distributed and overlapping, and they demonstrate the necessity of multivariate analysis for dissociating such representations.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Short-term memory refers to the ability to temporarily retain information when it is no longer present in the environment. The related and overlapping construct of working memory also incorporates the ability to manipulate or otherwise transform information, to protect it in the face of interference, and to use it to guide behavior. These abilities (from here on referred to as "STM") are of central importance in the study of human cognition, being implicated as critical contributors to such functions and properties as language comprehension, learning, planning, reasoning and general fluid intelligence (Baddeley, 1986; Conway, Kane, & Engle, 2003; Engle & Kane, 2003; Engle, Kane, & Tuholski, 1999; Unsworth & Engle, 2007). The brain structures and cognitive processes underlying STM are topics of intense investigation and debate (see Jonides et al., 2008; and Postle, 2006b for reviews).

Many contemporary accounts of the neural bases of cognition (e.g., Haxby et al., 2001; Rogers & McClelland, 2004) describe mental representations as emergent properties of coordinated and distributed neural activity. However, many traditional techniques for analyzing neuroimaging data are poorly suited for the investigation of distributed systems. This is because they are limited to identifying *magnitudes of activity* in localized parts of the brain and, in effect, assume a homogeneity of representation within contiguous clusters of voxels. The profound importance of the recent introduction to cognitive neuroscience of multivariate pattern analysis (MVPA) (Haynes & Rees, 2006; Kriegeskorte, Goebel, & Bandettini, 2006; Norman, Polyn, Detre, & Haxby, 2006; Pereira, Mitchell, & Botvinick, 2009), therefore, is that it affords the detection and characterization of *information* that is represented in patterns of activity distributed within and across multiple regions of the brain.

The research that we present here is motivated by the *embedded-component* theories of STM (Cowan, 1988, 1995; Ericsson & Kintsch, 1995; Oberauer, 2002), which characterize STM as an emergent property of the interaction of long-term memory and attention. They postulate a distinction between a capacity-limited central component of STM (referred to as the *focus of attention*) and a more

* Corresponding author at: Princeton University; Green Hall; Princeton, NJ 08540, United States. Tel.: +1 609 258 4442; fax: +1 609 258 1113.
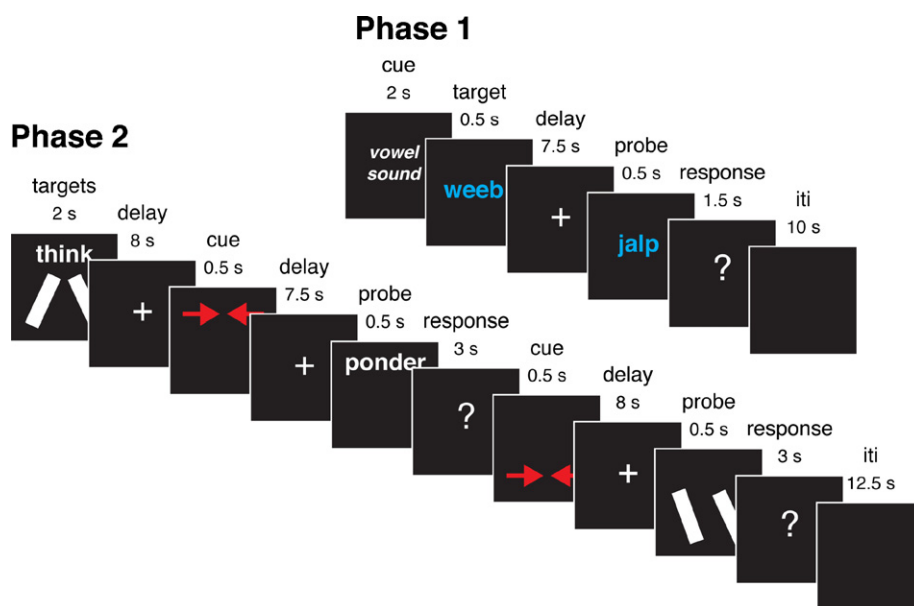 E-mail address: jalewpea@princeton.edu (J.A. Lewis-Peacock).

Phase 1

Phase 2

Fig. 1. Task Diagrams. In Phase 1, participants performed short-term recognition of a single stimulus. In Phase 2, they performed short-term recognition of two stimuli.

peripheral component (outside the focus, commonly referred to as *activated long-term memory*). To date, we have leveraged MVPA to generate stronger neural evidence than had previously existed for the idea that reactivated long-term memory representations are the basis of STM (Lewis-Peacock & Postle, 2008), and to generate some of the first evidence (see also Nee & Jonides, 2008, 2011) that the distinction between attended and unattended representations *within* STM, which has been proposed on the basis of behavioral evidence (Cowan, 1988; Oberauer, 2002), has a neural basis (Lewis-Peacock et al., in press). Independent of the embedded-component model, the results from this latter study have demonstrated that the active neural signature of information being remembered across a brief delay can be disrupted by redirecting attention, without sacrificing the short-term retention of that information. This finding raises questions about the common view in cognitive neuroscience that the maintenance of persistently elevated neural activity is required for the short-term retention of information, and supports an alternative model: the sustained activation of a stimulus representation is not necessary for its short-term retention; this activity, instead, corresponds to the focus of attention.

Here, we present several additional analyses of the data from Lewis-Peacock et al. (in press) that compare the inferences that can be drawn from MVPA vs. from univariate approaches based on the general linear model (GLM). We began by attempting to decode brain activity, not from multivariate patterns of activity throughout the brain, but from the average activity inside category-selective regions of interest (ROI). The successes and failures of this approach in replicating the results from MVPA are instructive. The successes provide confirmatory evidence that neural representations inside and outside the focus of attention are neurally dissociable. The failures of the GLM approach, however, illustrate how MVPA can provide additional insights into the neural bases of cognition. For example, MVPA alone was able to verify that STM representations, like perceptual ones (Haxby et al., 2001), are widely distributed and overlapping, and that they can be observed in brain regions which fail to show elevated activity during the delay period (e.g., Serences, Ester, Vogel, & Awh, 2009; Harrison & Tong, 2009). Also, the multivariate results highlight potentially misleading interpretations of univariate results (e.g., Postle, 2006a) that are based on the (faulty) assumption of homogeneity of representation in category-selective brain regions.

2. Methods

2.1. Behavioral task

A full description of the design and analysis of this experiment is presented in Lewis-Peacock et al. (in press). Here, we provide a brief overview. Nine healthy young adults were scanned in one session performing two different tasks. In the first, they performed short-term recognition of a stimulus drawn from one of three categories: words, pronounceable pseudo-words, or line segments (Fig. 1, Phase 1). Subjects indicated whether the probe stimulus matched the target stimulus ($p = 0.5$) according to a domain-specific judgment: synonym (words), rhyme (pseudo-words), and orientation (line segments). Foils for the three categories were, respectively, conceptually unrelated words, single-syllable pseudo-words with a non-matching vowel sound, or line segments in which one of the segments differed in orientation from the targets by at least 30°. The stimuli and task demands were designed to encourage domain-specific encoding in a primary dimension for each trial (semantic, phonological, and visuospatial, respectively). In the second task, subjects performed a two-step short-term recognition task in which two target items, each drawn from a separate category, are presented as targets, followed by a brief delay, followed by a cue indicating the target item for which memory would be tested by the first recognition probe. As in the first task, subjects indicated whether this first probe stimulus matched the cued target item according to a domain-specific judgment. Trials were configured such that there was a probability of 0.5 that the probe stimulus satisfied the criterion, with foils chosen as before. After the first probe, a second cue appeared which indicated the target item for which memory would be tested by the second probe, with equal probability of cuing the same item (*repeat* trials) or the other item (*switch* trials). Thus, until the onset of the second cue, all items needed to be maintained in STM for successful task performance.

2.2. Analysis plan

The logic of the analyses presented here was to demonstrate the benefits of using MVPA over a conventional univariate approach for addressing questions relating to the neural bases of STM. The structure of the analyses was first to analyze neural data from Phase 1 (the "training" data), and then to use the results of that analysis to decode neural data from Phase 2 (the "testing" data). We used two methodologies – univariate analysis (GLM), and multivariate analysis (MVPA) – in order to compare the results obtained from each. For Phase 1, a GLM on whole-brain fMRI data identified voxels whose activity remained elevated during the delay periods of the three categories of STM trials. Such elevated delay-period activity is widely considered to be the neural basis of STM (e.g., Curtis & D'Esposito, 2003; Fuster & Alexander, 1971; Kubota & Niki, 1971). These voxels were used to create a ROI for each stimulus category. For Phase 2, we decoded the moment-to-moment contents of STM from the neural data, at every time point and in every trial, using two different approaches: (1) by inspecting the average signal intensity within category-specific ROIs; and (2) by assessing the multivariate patterns of signal intensity within these same ROIs. The former is modeled on traditional univariate fMRI analysis which assumes a homogeneity of representation within contiguous clusters of voxels (e.g., Worsley & Friston, 1995); the latter on MVPA which assumes that neural representations are

both distributed and overlapping (e.g., Norman et al., 2006). Finally, we also applied MVPA to brain regions outside the putative "task-sensitive network" identified by the GLM to demonstrate the heightened sensitivity and inferential power of the method.

The continuous decoding of data from the entirety of the Phase 2 trials allowed for a complete characterization of the evolution of brain states corresponding to category-specific information inside and outside the focus of attention. If sustained brain activity reflected the contents of the focus of attention, but not all of STM, one would expect that the delay-period brain activity would reflect only that information that had most recently been cued. Based on behavioral evidence (Oberauer, 2001, 2005), we expected that during the initial delay period both target items should be maintained in the focus of attention because both were potentially relevant for the first response. Following the first cue, the uncued item would be removed from the focus of attention after about 1–2 s. (This timing also holds for the cross-category stimulus sets used in the present study; Drysdale, Lewis-Peacock, Oberauer, & Postle, 2010.) Such removal of task-irrelevant information from the focus would be indicated by an attenuation of neural evidence for the active representation of that target item. Whether the strength of the evidence was to drop to an intermediate level, or to baseline, would have implications for what it means for information to be "in STM" but outside the focus of attention. On switch trials, retrieval of information into the focus of attention would be indicated by the restrengthening of neural evidence for the target item cued as relevant for the second decision. In contrast, if sustained brain activity reflected the full contents of STM, we would expect that, regardless of cueing, evidence for the active neural representation for both target items should remain strong throughout the trial (at least until the second cue, because both stimuli had to be remembered up to that point). Each component of our analyses will now be described in more detail.

2.3. GLM: Phase 1

A traditional mass-univariate analysis based on the GLM was performed on the Phase 1 data using AFNI's *3dDeconvolve*. All trial events were modeled with boxcar regressors of different lengths: cue (1 s), target (0.5 s), delay (7.5 s), probe (2 s), and feedback (0.5 s). A third-order polynomial was used for the null hypothesis, and all basis functions for trial events were normalized to have an amplitude of 1. For each participant, three thresholded ($p < 0.01$, uncorrected) sets of voxels were extracted from the GLM based on *t*-tests (with respect to baseline) of delay-period regressors from phonological, semantic, and visual trials, respectively. No clustering algorithm was used, thus voxels were not forced to be spatially contiguous at an arbitrary threshold. Nonetheless, the activation maps showed high levels of spatial clustering based on interrogation of subject-level and group-level maps. Because the extracted ROIs were not mutually exclusive (i.e., they shared voxels that were identified as active for multiple trial categories), we refer to them as the "inclusive" ROIs. Three "exclusive" ROIs were created by removing, for each category, any voxels that were shared by at least one other category. These ROIs were subsequently used for both univariate and multivariate decoding of the neural data from Phase 2. Four more ROIs were created solely for use with MVPA. A concatenation of the three inclusive ROIs formed an aggregate *Inclusive* ROI, and a concatenation of the three exclusive ROIs formed an aggregate *Exclusive* ROI. The use of these aggregate ROIs provided MVPA access to the all of the voxels that were used by the GLM decoding analysis (described below). Finally, the *Overlap* ROI contained only those voxels that were identified as active for all three categories, and the *Removed* ROI contained all voxels that remained active after removing those identified as active for any category. These final two ROIs allowed MVPA to perform additional hypothesis testing in brain regions for which the GLM could not. For display purposes, the ROIs for each participant were transformed into standardized space using AFNI's *@auto_tlrc*, blurred with a full-width half-max of 8 mm, averaged across all participants with *3dmerge*, and mapped onto an inflated anatomical version of the N27 brain dataset (Holmes et al., 1998) using AFNI's surfacing mapping utility (SUMA).

2.4. GLM: Phase 2

For each participant, we calculated the average signal intensity in category-specific ROIs to assess the extent to which STM representations in the Phase 2 task exhibited an active neural trace during the delay period. Preprocessed (see Lewis-Peacock et al. (in press) for details) and z-scored fMRI data at intervals of TR = 2 s from every trial were classified by the GLM. At each time point in the trial, the strength of representation for each of the two target items was estimated by averaging the signal intensity within the appropriate category-specific ROI. For example, for a trial that contained a semantic and a visual stimulus (as depicted in Fig. 1, Phase 2), the moment-to-moment signal strength of the semantic item representation was estimated by calculating the average activity within the *semantic* ROI, whereas the signal strength of the visual item representation was estimated separately from the average activity within the *visual* ROI. To combine results from all trials, the GLM estimates for phonological, semantic, and visual were relabeled and collapsed across all trials into three new categories: *1st* (the category of the target item selected by the first cue), *2nd* (the category of the other target item), and *irrel* (the trial-irrelevant category). For display purposes, the signal intensity values at each time point in the trial-averaged data for each participant were normalized by removing the resting state baseline level of activity from the ROIs being tested. Per standard

GLM procedure, this baseline activity was used as reference for assessing whether a category-specific representation was "above baseline" throughout the trial. Finally, the recoded data were averaged across all participants for hypothesis testing, and the group-averaged data were spline interpolated across the 23 discrete data points in each trial to create smooth waveforms for display.

2.5. MVPA: Phase 1

MVPA was performed on the Phase 1 data, separately, in all ROIs identified by the GLM. The classification procedure used was modeled on the whole-brain analyses from Lewis-Peacock et al. (in press). All classification was carried out using penalized logistic regression, using L_2 regularization with a penalty parameter of 50. Regularization prevents over-fitting by punishing large weights during classifier training (Duda, Hart, & Stork, 2001). Results were fairly insensitive to the penalty strength (sampled between 0.1 and 1000), although performance was markedly improved over unpenalized classification using the backpropagation algorithm. A unique classifier was created for each participant and applied only to that participant's data. A feature selection analysis of variance (ANOVA) was applied to the preprocessed images to select those voxels whose activity varied significantly ($p < 0.05$) between the categories over the course of entire task. This standard machine learning procedure reduces noise in the classification by removing uninformative voxels. (Note: all classification analyses were repeated without feature selection and the results were qualitatively similar). Data from the final 6 s of the 7.5-s delay period in the Phase 1 task, at intervals of TR = 2 s, were used to train a classifier to distinguish patterns of brain activity corresponding to the short-term retention of information encoded primarily in a phonological (pseudoword trials), semantic (word trials), or visual (line trials) form. All data were shifted back in time by 4 s to account for hemodynamic lag of the BOLD signal. Therefore, the data that were used from each trial were actually recorded between 8 and 14 s after the beginning of the trial. This adjustment, although crude, reasonably accommodates the slow hemodynamic response and is standard practice in MVPA. As a check on validity, we retrained the classifier using a 6 s lag adjustment, and this did not significantly alter the results. We evaluated classifier training accuracy by using the method of k-fold cross-validation, i.e., training on k-1 blocks of data and testing on the kth block, and then rotating and repeating until all trials had been classified. For each 2-s TR of fMRI data, the classifier produced an estimate (from 0 to 1) of the extent to which the brain activity matched the pattern of activity corresponding to the categories it had been trained on. These estimates reflected the classifier's *evidence* for each category. The classifier's *prediction* at each TR corresponded to the category with the most evidence. Prediction accuracy was calculated as the proportion of TRs in which the classifier correctly predicted the actual category of the trial from which that TR was sampled.

To improve the interpretability of the whole-trial decoding of the Phase 2 data, we trained the classifier on a fourth category: resting state brain activity during the unfilled inter-trial interval (ITI). Resting activity served as a "ground reference" for the classifier, analogous to how the Earth serves as a zero-voltage ground reference for electrical circuits. Training the classifier with rest activity did not alter the classifier's assessment of the relative differences between the three stimulus categories during the task-portion of the trial. It did, however, normalize the classifier's assessment such that evidence for the stimulus categories was low during the rest periods (during which time the participants were not performing a STM task). Data from the ITI were randomly sampled so that, within each block of trials, the classifier was trained on the same number of exemplars for all four categories (72 total TRs each of phonological, semantic, visual, and ITI across the whole experiment).

To assess the relative importance of different brain areas to the classification of the stimulus categories, we determined, from a classifier trained using all brain voxels, which voxels were important for (correctly) identifying patterns of brain activity corresponding to each of the three categories. We applied the voxel importance formula (from McDuff, Frankel, & Norman, 2009): $imp_{ij} = 100 \times w_{ij} \times avg_{ij}$, where w_{ij} is the weight between input unit i and output unit j, and avg_{ij} is the average activity of input i during the short-term retention of category j. Positive importance was assigned to a voxel whose average activity was positive (indicating that it was more active than usual), negative importance was assigned to a voxel whose average activity was negative (indicating that it was less active than usual), and voxels where the sign of w_{ij} differed from the sign of avg_{ij} (indicating a net negative contribution of that voxel to detecting that task state) were assigned an importance value of zero. Importance maps for the three categories were calculated for each participant. For display purposes, these maps were then transformed into standardized space, averaged across participants, thresholded at an absolute value of importance of 0.075, and mapped onto an inflated brain (as described above for the GLM maps).

2.6. MVPA: Phase 2

MVPA decoding was performed on the Phase 2 data, separately, in all ROIs identified by the GLM from the Phase 1 data. A pattern classifier for each participant, trained on all four blocks of Phase 1 data, was used to assess the extent to which category-specific patterns of brain activity reappeared during the Phase 2 task. Preprocessed fMRI data at intervals of TR = 2 s, masked by the feature-selected set of voxels within the ROI being tested, were classified from every trial. Classifier evidence values for phonological, semantic, and visual representations were relabeled and collapsed across all trials into *1st*, *2nd*, and *irrel* categories (as described above).
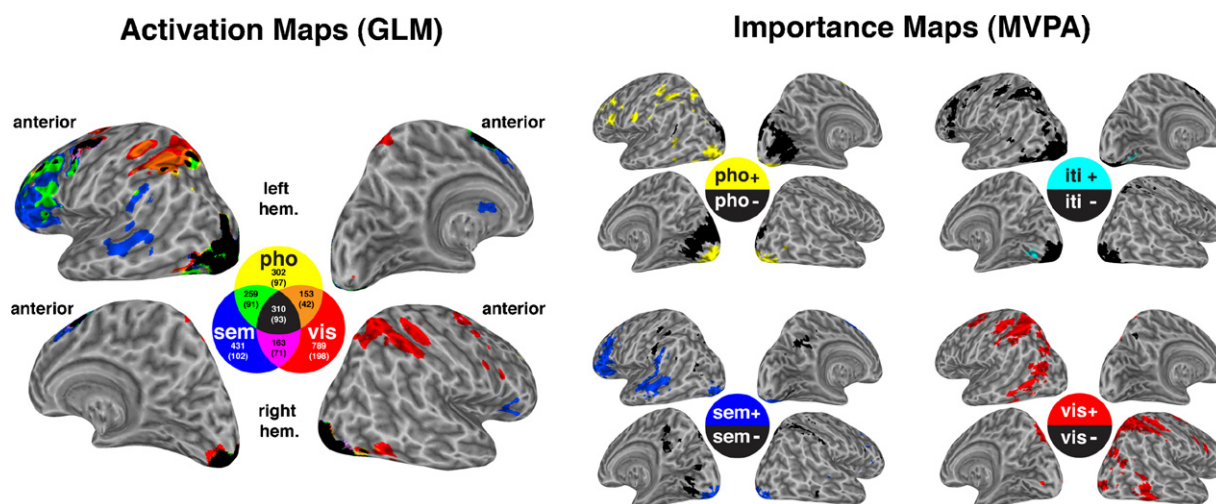
Activation Maps (GLM)　　　　　　Importance Maps (MVPA)

Fig. 2. Brain Maps. Delay-period activation maps are shown on the left for phonological (*pho*), semantic (*sem*), and visual (*vis*) trials in Phase 1. The mean number of voxels for each category (plus the overlapping voxels between categories) are shown in the color-coded venn diagram. Delay-period classifier importance maps are shown on the right, separately for each category (including the inter-trial interval, *iti*). For each category, positive importance is indicated with a category-specific color and negative importance is indicated with black. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Unlike the GLM analyses described above, the information estimates from the classifier were not normalized by removing from each time point the classifier's estimates from the rest period. Whereas resting state levels of BOLD signal are typically interpreted as a meaningful baseline, resting state levels of classifier evidence values are not. Therefore, the "baseline" reference used for the MVPA analyses was the classifier's evidence for the trial-irrelevant category at each time point throughout the trial. (Although we also show mean signal intensities from the trial-irrelevant category's ROI in the GLM decoding results, we do not interpret them as baseline because such data are not intuitively meaningful and are not typically, if ever, used to normalize univariate activation results.) Finally, the recoded trial-averaged MVPA results were averaged across all participants for hypothesis testing, and spline interpolated across the 23 discrete data points in each trial to create smooth waveforms for display.

3. Results

3.1. Behavior

Full behavioral results for the Phase 1 and Phase 2 tasks are reported in Lewis-Peacock et al. (in press). The mean accuracy and response time across all participants in the Phase 1 task were 94% (SEM = 1) and 933 ms (SEM = 22). The mean accuracy and response time across all participants in the Phase 2 task were 91% (SEM = 1) and 936 ms (SEM = 10). Participants were more accurate ($F(1,8) = 27.18$, $p < 0.001$) and faster to respond ($F(1,8) = 7.86$, $p = 0.023$) on repeat trials (93%, SEM = 1; 898 ms, SEM = 13) than on switch trials (88%, SEM = 1; 975 ms, SEM = 15). Here, all subsequent analyses focus on switch trials only.

3.2. Phase 1: GLM vs. MVPA

Group-averaged locations of voxels in GLM-defined ROIs are displayed on inflated brain hemispheres in Fig. 2 (*left*). The category-specific regions identified here are broadly consistent with previous findings related to phonological (e.g., Buchsbaum & D'Esposito, 2008), semantic (e.g., Shivde & Thompson-Schill, 2004), and visuospatial (e.g., Postle & D'Esposito, 1999) STM. (Note that group averaging obscured the presence of supra-threshold voxels that were anatomically heterogeneous across subjects, such as in posterior superior temporal gyrus in many individual phonological ROIs.) Importance maps for each category (distinguishing positive from negative voxels; and based on whole-brain classification) are also shown in Fig. 2 (*right*). Positive voxels are those for which increases in activity were important for classification, and negative

voxels are those for which decreases in activity were important for classification. Although importance maps do not indicate *where* information is stored in the brain, but rather *which* voxels the classifier found to be important for classification of each category, visual comparison of these maps to the GLM activation maps reveals a high-degree of correspondence between supra-threshold voxels (GLM) and positively informative voxels (MVPA) for each category.

MVPA performance (classification accuracy & classification evidence) for the Phase 1 data is shown in Fig. 3 for seven different ROIs defined by the GLM. The assumption that the GLM identifies localized areas (i.e., voxels) that are specific for a particular kind of information would make the following predictions: classification in the *Inclusive* and *Exclusive* ROIs should be excellent for all categories; in the *Phonological*, *Semantic*, and *Visual* ROIs it should be excellent for that ROI's category and at chance for all others; and in the *Overlap* and *Removed* ROIs it should be at chance for all categories. However, group-averaged results show that classification succeeded in *all* ROIs. That is, delay-period activity from every ROI were reliably classified as matching the stimulus category of the trial. This result indicates that the classifier successfully differentiated visuospatial from phonological (Baddeley, 1986) from semantic (Cameron, Haarmann, Grafman, & Ruchkin, 2005; Haarmann & Usher, 2001; Martin, Wu, Freedman, Jackson, & Lesch, 2003; Shivde & Thompson-Schill, 2004) STM, and all three from the resting state activity recorded during the ITI. Prediction accuracy for each category in all ROIs was significantly above chance based on independent-sample *t*-tests across participants. This was true if we considered chance-level performance to be 25% (considering all four categories that the classifier was trained on; $p < 0.001$) or 33% (considering only the three task-related categories (ignoring ITI); $p < 0.05$). The heightened sensitivity to neural representations of MVPA compared to the GLM is clearly demonstrated by the fact that MVPA was able to neurally distinguish phonological, semantic, and visual STM in regions that the GLM identified as: (a) task-sensitive but not category-selective (*Overlap*); (b) task-sensitive but exclusive to one category (*Phonological*, *Semantic*, and *Visual*); and (c) task-insensitive (*Removed*).

The mean classifier evidence values in all ROIs showed strong category-selectivity, supported by a significant interaction of trial type (*phonological/semantic/visual/ITI*) × evidence type (*phonological/semantic/visual/ITI*) from a 4×4 repeated measures ANOVA on the classifier evidence values ($p < 0.001$). In all ROIs except *Overlap*,
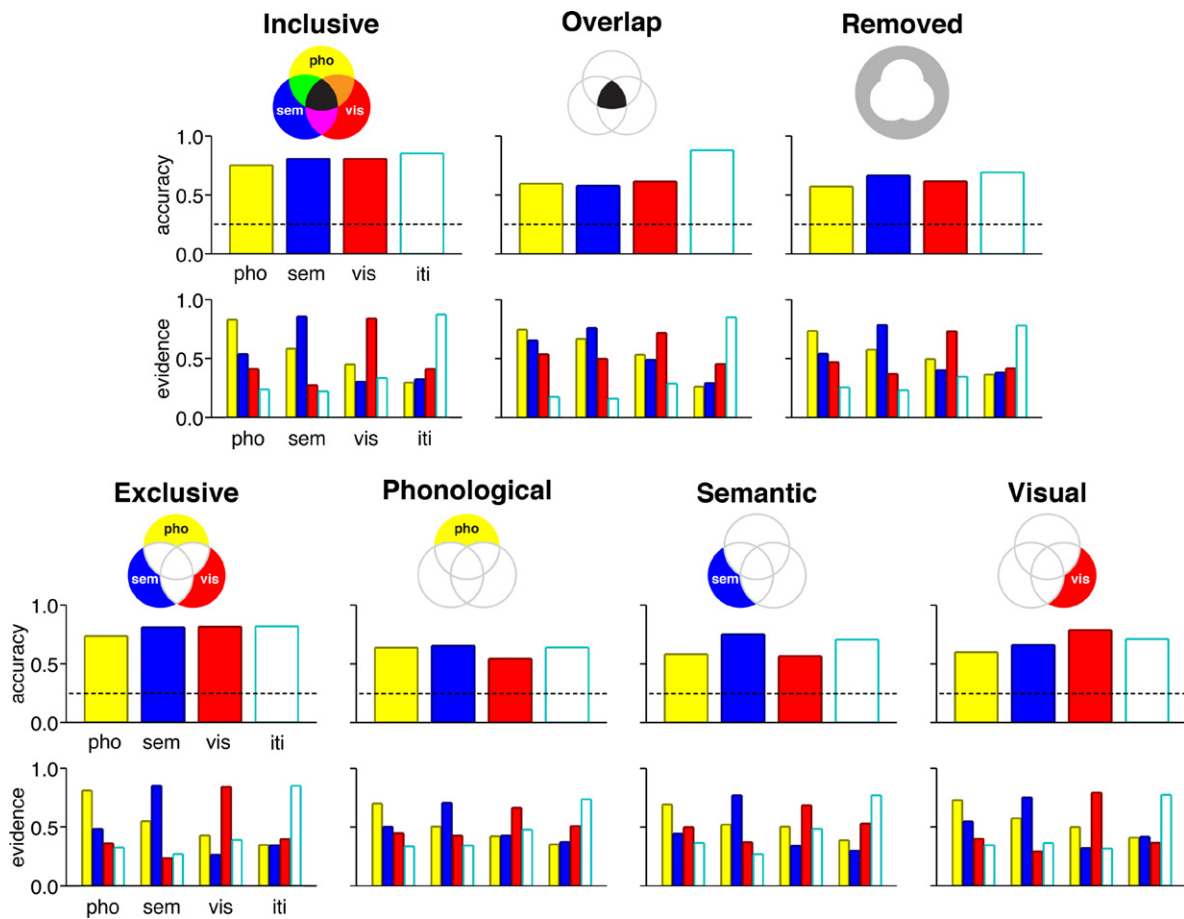
Fig. 3. MVPA Training. Classifier training performance on Phase 1 data is shown for seven different ROIs. For each ROI, the mean classifier accuracy (*top*) and mean classifier evidence (*bottom*) across all trials and participants are shown for phonological (*pho*), semantic (*sem*), and visual (*vis*) trials, and for the inter-trial interval (*iti*).

follow-up pairwise comparisons indicated significant differences ($p < 0.05$, Bonferroni corrected) between the relevant evidence value and the other evidence values (e.g., the evidence for the *phonological* category vs. evidence values for *semantic, visual, and ITI* on phonological trials). A qualitatively similar pattern of results was observed in the *Overlap* ROI: the overall ANOVA on classifier evidence scores was significant ($F(9,72) = 40.42$, $p < 0.001$), and all follow-up pairwise comparisons were significant ($p < 0.05$, Bonferroni corrected) except for the comparisons between phonological and semantic evidence scores on phonological and semantic trials ($p > 0.05$, uncorrected). Despite the relatively weaker category-selectivity for phonological and semantic representations in this ROI, however, the classifier's prediction accuracy was well above chance for both categories ($p < 0.001$). Importantly, although classification is a discriminative procedure, the pattern of evidence values in these results demonstrates that the categories were not anti-correlated by the classifier. That is, the increasing strength of one category representation did not necessarily decrease the strength for another category. This claim is supported by the graded evidence values for a given category across the three trial types (e.g., *phonological* evidence was highest in phonological trials, moderately high in semantic trials, and lowest in visual trials). If the categories were anti-correlated, one would expect to find high category evidence for trials from that category and uniformly low evidence for that category for trials from other categories. Therefore, we interpret the classifier evidence values as (reasonably) independent indicators of brain activity for each category. This interpretation is supported by the classifier's detection of superimposed patterns of brain activity corresponding to two memory

items from different categories during the initial delay period in the Phase 2 task. These data will now be described in more detail.

3.3. Phase 2: GLM vs. MVPA

Brain data from every time point in all Phase 2 switch trials were decoded from the *Inclusive* and the *Exclusive* ROIs, separately for each participant, using two different methods: GLM decoding (Fig. 4, *top row*); and MVPA decoding (Fig. 4, *bottom row*). The initial overall conclusion that one can draw from the results is that whereas the BOLD response is markedly different in these two ROIs – with signal strength being higher in the Inclusive ROI and waveforms less discriminable – MVPA decoding performance is effectively identical in the two. At a finer grain of detail, in both ROIs, group-averaged decoding across all trials revealed an initial rise in mean BOLD signal (GLM) and mean classifier evidence (MVPA) corresponding to the two categories of stimuli presented at the beginning of each trial. Thus, both methods indicated that the two target items were encoded and sustained in the focus of attention across the initial memory delay, while it was equiprobable that either would be relevant for the first memory response. Following onset of the first cue, both methods revealed a strengthening of the neural representation of the cued item. The pattern for the uncued item, however, differed across methods and, for the BOLD data, across ROIs. With MVPA, classifier evidence for the uncued item dropped to baseline. For BOLD data in the *Inclusive* ROI, the uncued signal also increased in response to the first cue, but remained weaker than the cued signal until the time of the second cue. In the *Exclusive* ROI, although the uncued BOLD signal
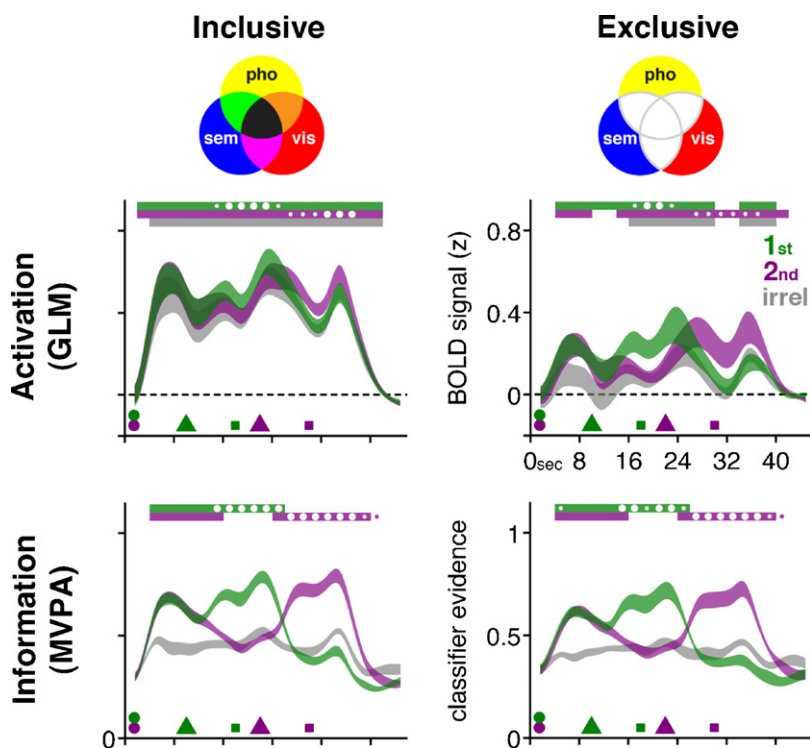
Fig. 4. Decoding: GLM vs. MVPA. Trial-averaged decoding of Phase 2 switch trials across all participants is shown for the *Inclusive* ROI (*left column*) and *Exclusive* ROI (*right column*), using the GLM method (*top row*) and the MVPA method (*bottom row*). Data for each category (*1st*, the first cued category; *2nd*, the second cued category; *irrel*, the trial-irrelevant category) are shown as ribbons whose thickness indicate ±1 SEM across participants. The colored shapes along this horizontal axis indicate the onset of the targets (green and purple circles), the first cue (green triangle), the first probe (green square), the second cue (purple triangle), and the second probe (purple square). Statistical comparisons focused on within-subject differences: for every 2-s interval throughout the trial, color-coded bars along the top of each graph indicate that the signal intensity (GLM) or classifier evidence (MVPA) for each category was above baseline. Activation baseline is mean signal intensity during rest, whereas information baseline is mean classifier evidence for *irrel* at each time point (see Section 2). Circles inside and outside these bars indicate that the value for one trial-relevant category was stronger than the value for the other trial-relevant category (small circles: $p < 0.05$; big circles: $p < 0.002$, Bonferroni corrected). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

did decline relative to the cued signal, it nonetheless remained elevated above baseline throughout the first delay period. Because we present results only for switch trials here, in these data the second cue always selected the previously uncued item as relevant for the second response. Following the second cue, all four plots show the neural dissociation between the two target items inverted, such that the previously uncued item exhibited the stronger representation both in terms of BOLD activity and classifier evidence. The BOLD signal for the now-uncued item in the *Inclusive* ROI is different from the other three, however, in that it remained elevated above baseline throughout the second delay period. The pairwise comparisons shown for both ROIs at each time interval in Fig. 4 are validated by significant ($p < 0.001$) 3×23 repeated measures ANOVAs on trial-averaged BOLD signal (GLM) and classifier evidence values (MVPA), with stimulus type (*1st/2nd/irrel*) and time (*TRs 1-23*) as within-subjects factors.

There are two highly salient observations that emerge from the analyses performed on these two ROIs. The first is that whereas the behavior of the BOLD signal is highly sensitive to the ROI from which it is extracted, the MVPA appears to be relatively stable. (This will be reinforced as additional ROIs are interrogated; Fig. 5.) The second observation, which has important implications for both cognitive and neurobiological models of STM, relates to the neural fate of the uncued (i.e., unattended) STM item after the first cue. During this portion of the trial, this item remained potentially relevant for the second half of the trial and thus could not be forgotten. According to the GLM analysis, in both the *Inclusive* and *Exclusive* ROIs, the BOLD signal remained elevated above baseline for voxels corresponding to the uncued category, although it was lower than in the voxels corresponding to the cued category. This result could

be taken as evidence for an "intermediate" state of activation for STM representations outside the focus of attention. MVPA decoding of this elevated activity, however, refuted this interpretation. Instead, MVPA showed the multivariate patterns in both regions to reflect the representation of the category that had been cued, and that the active neural representation of the uncued item effectively disappeared.

Further demonstrations of the inferential strength of MVPA are shown in Fig. 5. Successful decoding of Phase 2 data is shown from within voxel regions that the GLM identified as: (a) task-sensitive but not category-selective (*Overlap*); (b) task-sensitive but exclusive to one category (*Phonological*, *Semantic*, and *Visual*); and (c) task-insensitive (*Removed*). The MVPA results obtained in these regions are consistent with those obtained in the putative task-sensitive regions (Fig. 4), in that the delay-period brain activity is shown to reflect the focus of attention, but not the contents of STM per se. Whereas the GLM results show that category-selective voxels remain activated above baseline even when their category is uncued, MVPA applied to the same voxels (i.e., *Phonological, Semantic,* and *Visual* ROIs) reveals that that above-baseline activity actually reflects a representation of the cued stimulus (even if that stimulus is from a different category). Therefore, MVPA shows that it would be misleading to interpret the activation of category-selective voxels as reflecting a sustained representation of an item from that category. The pairwise comparisons shown for all ROIs at each time interval in Fig. 5 are validated by significant ($p < 0.001$) 3×23 repeated measures ANOVAs on trial-averaged classifier evidence values, with stimulus type (*1st/2nd/irrel*) and time (*TRs 1-23*) as within-subjects factors. This pattern of results was extremely robust throughout the neocortex: it was also observed in prefrontal
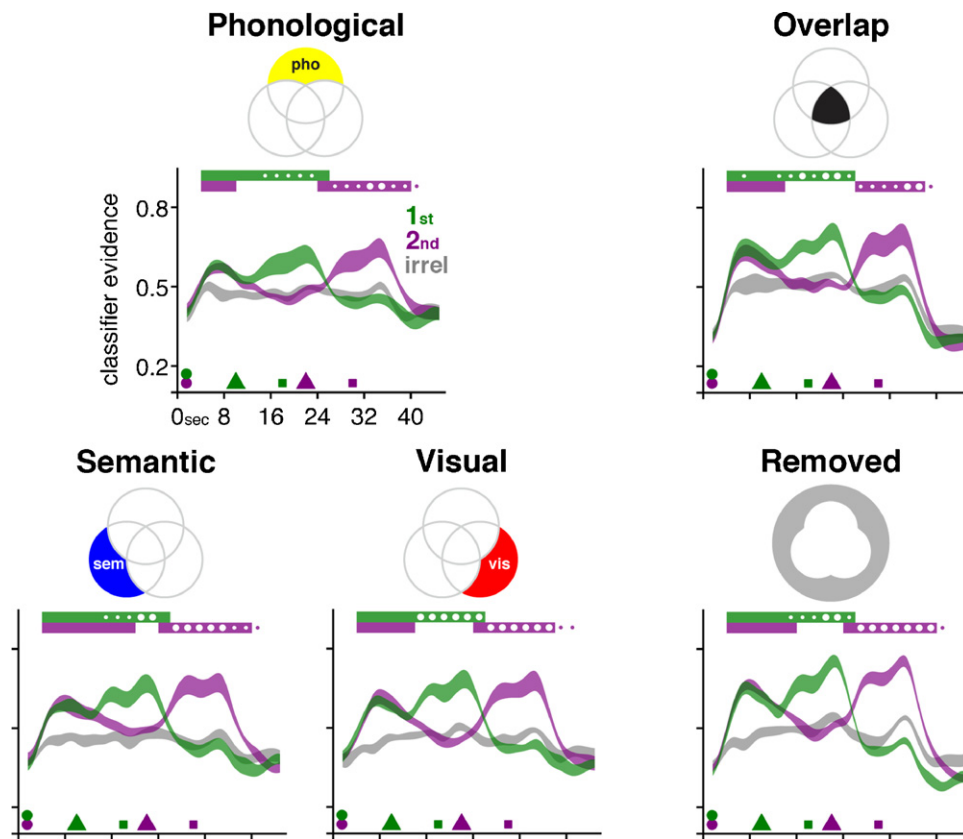
Fig. 5. More MVPA Decoding. Trial-averaged MVPA decoding of Phase 2 switch trials across all participants is shown for five ROIs from which the GLM could not decode. All graph conventions are as described in Fig. 4.

cortex, a brain region that is known to be important for STM but one whose activity patterns are notoriously difficult to classify (e.g., Lewis-Peacock & Postle, 2008); and it was even observed in primary motor cortex which is an area not thought to be responsive to tests of STM (data not shown). However, these results could not be replicated when we restricted decoding to sub-cortical regions.

4. Discussion

Our recent findings (Lewis-Peacock et al., in press) derived with MVPA, have provided some of the first neural evidence for a distinction between attended and unattended representations in STM, and they have also revealed a novel insight: the active neural signature of information being remembered across a brief delay can be disrupted by redirecting attention, without sacrificing the short-term retention of that information. This finding raises questions about the common view in cognitive neuroscience that the maintenance of persistently elevated neural activity is required for the short-term retention of information, and supports an alternative model: the sustained activation of a stimulus representation is not necessary for its short-term retention; this activity, instead, corresponds to the focus of attention. Therefore, we posited that two complementary forms of retention underlie STM: (1) the active retention of information *inside* the focus of attention via sustained neural firing, and (2) the passive retention of information *outside* the focus via some other neural mechanism (e.g., changes in synaptic potentiation) from which it can be reactivated with cue-based retrieval. The latter idea is anticipated in neural-network models of serial order recall (Botvinick & Plaut, 2006; Burgess & Hitch, 1999, 2006; Farrell & Lewandowsky, 2002), and in "retrieved context" models of memory search (Howard & Kahana, 2002; Polyn, Norman, & Kahana, 2009; Sederberg, Howard, & Kahana, 2008).

The suggestion that long-term memory (LTM) mechanisms support performance during a test of short-term retention is not novel. In dual-store models (Atkinson & Shiffrin, 1968; Waugh & Norman, 1965), the contribution of LTM is thought to supplement (and not replace) a STM system that is capable of holding several items. Neural evidence for this idea comes from neuroimaging and neuropsychological studies which have demonstrated that medial temporal lobe structures (known to be essential for LTM) also contribute to performance on tests of short-term retention (Hannula, Tranel, & Cohen, 2006; Jeneson, Mauldin, & Squire, 2010; Jeneson, Mauldin, Hopkins, & Squire, 2011; Nichols, Kao, Verfaellie, & Gabrieli, 2006; Olson, Moore, Stark, & Chatterjee, 2006; Olson, Page, Moore, Chatterjee, & Verfaellie, 2006). All theories of STM assume a capacity of more than one item, and typical estimates are around four (Cowan, 2001; Luck & Vogel, 1997). In our study, we deliberately held the overall memory load so small (2 items maximum) that the capacity limits of STM would not be exceeded. Therefore, based on the ubiquitous assumption that sustained activity is the neural correlate of maintenance in STM, one would expect to observe persistent neural representations for all memory items in our task. However, our results demonstrate that only the item in the focus of attention retained its active representation during the delay period. In fact, the focus demonstrably held two items at the same time, as shown by high classifier evidence for both target items after encoding, so it was not for lack of attentional capacity that only one representation was actively represented after the cue. Rather it was the behavioral relevance of the memory item that determined its activity status.

Through the additional analyses of these data presented here, we demonstrated that this discovery could not have been made with conventional mass-univariate analysis. Additionally, the direct comparisons presented here highlight the susceptibility of

the univariate approach to voxel selection bias, and the remarkable insensitivity of MVPA to this bias. In this section we will consider, in turn, the implications of these two findings. Although both methods dissociated neural representations for attended and unattended items in STM, the divergent results of GLM and MVPA with regard to the neural fate of unattended information have important implications for both cognitive and neurobiological models of STM. According to the GLM analysis, brain activity remained elevated above baseline in voxels presumed to be representing the unattended information. This result could be taken as evidence for an "intermediate" state of activation for STM representations outside the focus of attention, a hypothetical state which resides in between the highly activated state of information inside the focus of attention and the latent state of the vast majority of representations in episodic and semantic LTM. Such an intermediate state is predicted by many theoretical models (e.g., "activated long-term memory" in Cowan, 1988 and Oberauer, 2002; "long-term working memory" in Ericsson & Kintsch, 1995; "working memory" in McElree, 2006), and putative neural evidence for it has been described in BOLD data (Nee & Jonides, 2008). However, our MVPA results directly challenge this view, and they do so in two ways. First, is the "negative" finding from ROI-based analyses (and whole-brain analysis; Lewis-Peacock et al., in press) that unattended STM representations are not actively sustained during a brief memory interval, and yet they are not forgotten. Further, these representations can be restored to an active state if they are cued as relevant for subsequent behavior. (We interpret this as the retrieval of this information back into the focus of attention; Oberauer, 2005). Second, is the "positive" finding that voxels within ROIs defined by GLM to be category specific will flexibly represent the information that is currently in the focus of attention, rather than being restricted to only representing the category used to define them. These results are consistent with previous reports of representational flexibility in category-specific processing regions (e.g., Carlson, Schrater, & He, 2003; Chao, Martin, & Haxby, 1999; Cox & Savoy, 2003; Walther, Caddigan, Fei-Fei, & Beck, 2009). In addition to MVPA's performance in category-selective ROIs, it was also able to successfully decode STM representations from voxels that the GLM identified as either (a) insensitive to the task, or (b) task-sensitive but category-insensitive.

This highly stable pattern of classification results highlights at least two points. From the methodological perspective, it shows that MVPA, unlike GLM, is remarkably unsusceptible to voxel-selection bias (or, alternatively, that it lacks the regional specificity of GLM). Therefore, from the theoretical perspective, it reaffirms the need to dissociate brain regions whose information is actually "used" for task performance from those brain regions whose activity is epiphenomenal. For example, one might question the extent to which, despite successful decoding from its anatomical ROI, primary motor cortex represents the three categories in our task. Such a dissociation could be made, for example, by assessing correlations between ubiquitous category information and behavioral performance (Williams, Dang, & Kanwisher, 2007), or by disrupting this information in a specific brain region (e.g., by using transcranial magnetic stimulation; Mattavelli, Cattaneo, & Papagno, 2011) to test the necessity of that region for performance. Such approaches are being considered in our ongoing research.

In conclusion, the present findings highlight the fallacy of the assumption of homogeneity of representation within GLM-defined category-selective regions; one that, whether implicitly or explicitly acknowledged, is inherent in this approach. Instead, they provide strong evidence for highly distributed and overlapping neural representations in STM. Our analyses affirm and extend the view that MVPA is much more sensitive and robust than traditional measures of BOLD (see also Harrison & Tong, 2009; Kriegeskorte, Formisano, Sorger, & Goebel, 2007; Serences et al., 2009), and they

highlight the necessity of multivariate approaches for addressing theoretical questions pertaining to the neural bases of STM and attention. One final, but important, note is that our findings do not indict the broad array of neuropsychological and neuroscientific findings in support of localization of cognitive function in the brain. Rather, they suggest that multivariate and univariate fMRI approaches are complementary forms of analysis that should be applied depending on the question one is addressing and interpreted in accord with the aspects of the physiology that each is sensitive to.

Acknowledgements

References

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *The Psychology of Learning and Motivation: Advances in Research and Theory*, 2, 89–195.

Baddeley, A. D. (1986). *Working memory*. London: Oxford University Press.

Botvinick, M. M., & Plaut, D. C. (2006). Short-Term memory for serial order: A recurrent neural network model. *Psychological Review*, 113(2), 201–233.

Buchsbaum, B. R., & D'Esposito, M. (2008). The search for the phonological store: From loop to convolution. *Journal of Cognitive Neuroscience*, 20(5), 762–778.

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review*, 106(3), 551.

Burgess, N., & Hitch, G. J. (2006). A revised model of short-term memory and long-term learning of verbal sequences. *Journal of Memory and Language*, 55(4), 627–652.

Cameron, K. A., Haarmann, H. J., Grafman, J., & Ruchkin, D. S. (2005). Long-Term memory is the representational basis for semantic verbal short-term memory. *Psychophysiology*, 42(6), 643–653.

Carlson, T. A., Schrater, P., & He, S. (2003). Patterns of activity in the categorical representations of objects. *Journal of Cognitive Neuroscience*, 15(5), 704–717.

Chao, L. L., Martin, A., & Haxby, J. V. (1999). Are face-responsive regions selective only for faces? *Neuroreport*, 10(14), 2945–2950.

Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, 7(12), 547–552.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, 104(2), 163–191.

Cowan, N. (1995). *Attention and memory: An integrated framework*. New York: Oxford University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity (vol. 23, p. 87, 2001). *Behavioral and Brain Sciences*, 24(3).

Cox, D. D., & Savoy, R. L. (2003). Functional magnetic resonance imaging (fMRI) "brain reading": Detecting and classifying distributed patterns of fMRI activity in human visual cortex. *Neuroimage*, 19(2), 261–270.

Curtis, C. E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences*, 7(9), 415–423.

Drysdale, A. T., Lewis-Peacock, J. A., Oberauer, K., & Postle, B. R. (2010). Removing irrelevant information from working memory: Domain-Specific prioritization takes time. In *Poster presented at the cognitive neuroscience society annual meeting* Montreal, Quebec, Canada.

Duda, R. O., Hart, P. E., & Stork, D. G. (2001). *Pattern classification* (2nd ed.). New York: Wiley.

Engle, R. W., & Kane, M. J. (2003). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In *Psychology of learning and motivation*. Elsevier., pp. 145–199.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake, & P. Shah (Eds.), *Models of working memory* (pp. 102–134). Cambridge, U.K.: Cambridge University Press.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102(2), 211–244.

Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review*, 9(1), 59–79.

Fuster, J. M., & Alexander, G. E. (1971). Neuron activity related to short-term memory. *Science*, 173(3997), 652–654.

Haarmann, H., & Usher, M. (2001). Maintenance of semantic information in capacity-limited item short-term memory. *Psychonomic Bulletin & Review*, 8(3), 568–578.

Hannula, D. E., Tranel, D., & Cohen, N. J. (2006). The long and the short of it: Relational memory impairments in amnesia, even at short lags. *Journal of Neuroscience*, 26(32), 8352–8359.

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, 458(7238), 632–635.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293(5539), 2425–2430.

Haynes, J. D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, 7(7), 523–534.

Holmes, C. J., Hoge, R., Collins, L., Woods, R., Toga, A. W., & Evans, A. C. (1998). Enhancement of magnetic resonance images using registration for signal averaging. *Journal of Computer Assisted Tomography*, 22(2), 324–333.

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology*, 46(3), 269–299.

Jeneson, A., Mauldin, K. N., & Squire, L. R. (2010). Intact working memory for relational information after medial temporal lobe damage. *Journal of Neuroscience*, 30(41), 13624–13629.

Jeneson, A., Mauldin, K. N., Hopkins, R. O., & Squire, L. R. (2011). The role of the hippocampus in retaining relational information across short delays: The importance of memory load. *Learning & Memory*, 18(5), 301–305.

Jonides, J., Lewis, R. L., Nee, D. E., Lustig, C. A., Berman, M. G., & Moore, K. S. (2008). The mind and brain of short-term memory. *Annual Review of Psychology*, 59, 193–224.

Kriegeskorte, N., Formisano, E., Sorger, B., & Goebel, R. (2007). Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 104(51), 20600–20605.

Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-Based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America*, 103(10), 3863–3868.

Kubota, K., & Niki, H. (1971). Prefrontal cortical unit activity and delayed alternation performance in monkeys. *Journal of Neurophysiology*, 34(3), 337–347.

Lewis-Peacock, J. A., & Postle, B. R. (2008). Temporary activation of long-term memory supports working memory. *Journal of Neuroscience*, 28(35), 8765–8771.

Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K., & Postle, B. R. Neural evidence for a distinction between short-term memory and the focus of attention. Journal of Cognitive Neuroscience, in press.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657), 279–280.

Martin, R. C., Wu, D., Freedman, M., Jackson, E. F., & Lesch, M. (2003). An event-related fMRI investigation of phonological versus semantic short-term memory. *Journal of Neurolinguistics*, 16(4–5), 341–360.

Mattavelli, G., Cattaneo, Z., & Papagno, C. (2011). Transcranial magnetic stimulation of medial prefrontal cortex modulates face expressions processing in a priming task. *Neuropsychologia*, 49(5), 992–998.

McDuff, S. G., Frankel, H. C., & Norman, K. A. (2009). Multivoxel pattern analysis reveals increased memory targeting and reduced use of retrieved details during single-agenda source monitoring. *Journal of Neuroscience*, 29(2), 508–516.

McElree, B. (2006). Accessing recent events. *Psychology of Learning and Motivation*, 46, 155–200.

Nee, D. E., & Jonides, J. (2008). Neural correlates of access to short-term memory. *Proceedings of the National Academy of Sciences of the United States of America*, 105(37), 14228–14233.

Nee, D. E., & Jonides, J. (2011). Dissociable contributions of prefrontal cortex and the hippocampus to short-term memory: Evidence for a 3-state model of memory. *Neuroimage*, 54(2), 1540–1548.

Nichols, E. A., Kao, Y. C., Verfaellie, M., & Gabrieli, J. D. (2006). Working memory and long-term memory for faces: Evidence from fMRI and global amnesia for involvement of the medial temporal lobes. *Hippocampus*, 16(7), 604–616.

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: Multi-Voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, 10(9), 424–430.

Oberauer, K. (2001). Removing irrelevant information from working memory: A cognitive aging study with the modified sternberg task. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 27(4), 948–957.

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 28, 411–421.

Oberauer, K. (2005). Control of the contents of working memory – A comparison of two paradigms and two age groups. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 31(4), 714–728.

Olson, I. R., Moore, K. S., Stark, M., & Chatterjee, A. (2006). Visual working memory is impaired when the medial temporal lobe is damaged. *Journal of Cognitive Neuroscience*, 18(7), 1087–1097.

Olson, I. R., Page, K., Moore, K. S., Chatterjee, A., & Verfaellie, M. (2006). Working memory for conjunctions relies on the medial temporal lobe. *Journal of Neuroscience*, 26(17), 4596–4601.

Pereira, F., Mitchell, T., & Botvinick, M. (2009). Machine learning classifiers and fMRI: A tutorial overview. *Neuroimage*, 45(Suppl. 1), S199–S209.

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review*, 116(1), 129–156.

Postle, B. R. (2006a). Distraction-spanning sustained activity during delayed recognition of locations. *Neuroimage*, 30(3), 950–962.

Postle, B. R. (2006b). Working memory as an emergent property of the mind and brain. *Neuroscience*, 139, 23–38.

Postle, B. R., & D'Esposito, M. (1999). "What" – then – "where" in visual working memory: An event-related fMRI study. *Journal of Cognitive Neuroscience*, 11, 585–597.

Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cam-bridge, MA: MIT Press.

Sederberg, P. B., Howard, M. W., & Kahana, M. J. (2008). A context-based theory of recency and contiguity in free recall. *Psychological Review*, 115(4), 893–912.

Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science*, 20(2), 207–214.

Shivde, G. S., & Thompson-Schill, S. L. (2004). Dissociating semantic and phonological maintenance using fMRI. *Cognitive, Affective, & Behavioral Neuroscience*, 4, 10–19.

Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin*, 133(6), 1038–1066.

Walther, D. B., Caddigan, E., Fei-Fei, L., & Beck, D. M. (2009). Natural scene categories revealed in distributed patterns of activity in the human brain. *Journal of Neuroscience*, 29(34), 10573–10581.

Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological Review*, 72(2), 89.

Williams, M. A., Dang, S., & Kanwisher, N. G. (2007). Only some spatial patterns of fMRI response are read out in task performance. *Nature Neuroscience*, 10(6), 685–686.

Worsley, K. J., & Friston, K. J. (1995). Analysis of fMRI time-series revisited—Again. *Neuroimage*, 2(3), 173–181.