Attention to Distinct Goal-relevant Features Differentially Guides Semantic Knowledge Retrieval

Gavin K. Hanson and Evangelia G. Chrysikou

Abstract

■ A critical aspect of conceptual knowledge is the selective activation of goal-relevant aspects of meaning. Although the contributions of ventrolateral prefrontal and posterior temporal areas to semantic cognition are well established, the precise role of posterior parietal cortex in semantic control remains unknown. Here, we examined whether this region modulates attention to goal-relevant features within semantic memory according to the same principles that determine the salience of task-relevant object properties during visual attention. Using multivoxel pattern analysis, we decoded attentional referents during a semantic judgment task, in which participants matched an object cue to a target according to concrete (i.e., color, shape) or abstract (i.e., function, thematic context) semantic features. The goal-relevant semantic feature participants attended to (e.g., color or shape, function or theme) could be decoded from task-associated cortical activity with above-chance accuracy, a pattern that held for both concrete and abstract semantic features. A Bayesian confusion matrix analysis further identified differential contributions to representing attentional demands toward specific object properties across lateral prefrontal, posterior temporal, and inferior parietal regions, with the dorsolateral pFC supporting distinctions between higher-order properties and the left intraparietal sulcus being the only region supporting distinctions across all semantic features. These results are the first to demonstrate that patterns of neural activity in the parietal cortex are sensitive to which features of a concept are attended to, thus supporting the contributions of posterior parietal cortex to semantic control. ■

INTRODUCTION

Goal-directed behavior involves the controlled retrieval of semantic memory, our conceptual knowledge about sounds, actions, facts, objects, and word meanings. Several studies with healthy participants and brain-lesioned patients have revealed that semantic memory follows a neural architecture that is organized by feature (e.g., color, shape, size, mode of manipulation) and distributed across many sensorimotor cortical areas (Binder et al., 2016; Yee & Thompson-Schill, 2016; Yee, Chrysikou, & Thompson-Schill, 2014), with the potential contribution of convergence regions within the anterior temporal cortex that can distinguish among highly overlapping representations (Martin, 2007; Patterson, Nestor, & Rogers, 2007; Chao & Martin, 2000). Among the most critical aspects of semantic cognition is the selective activation of goal-relevant aspects of meaning (Miller & Cohen, 2001). For example, when identifying a dime as an object that can tighten a screw, it is only the dime's small, flat shape and ridged surface that are relevant to the achievement of this goal; this information is highlighted during semantic retrieval, relative to other stored knowledge about the object that is not directly relevant to the task (e.g., that the dime is silver, shiny, and worth $0.10).

Much research has investigated the neural mechanisms that support the representation and selection of bottom–up, goal-relevant features in semantic retrieval via top–down modulation of feature-specific sensorimotor cortex (Kan & Thompson-Schill, 2004). These studies have primarily focused on an extended network of prefrontal and posterior temporal regions implicated in semantic control (Whitney, Kirk, O'Sullivan, Lambon Ralph, & Jefferies, 2011; Whitney, Grossman, & Kircher, 2009). Furthermore, prefrontal areas together with the left parietal cortex (i.e., angular gyrus [AG], intraparietal sulcus [IPS]) are thought to constitute a larger "multiple demand" network that may support increased executive requirements irrespective of specific task demands (Duncan, 2010; Nagel, Schumacher, Goebel, & D'Esposito, 2008; Dosenbach et al., 2006; Owen, Schneider, & Duncan, 2000). Most of these findings point to a critical role for the left ventrolateral pFC (VLPFC) in semantic cognition, according to which dissociable systems within the left inferior frontal gyrus (LIFG) support guided either semantic retrieval of information from the posterior middle temporal cortex when semantic associations are weak (anterior VLPFC) or postretrieval semantic selection of goal-relevant features after the activation of multiple competing representations (posterior VLPFC; Nagel et al., 2008; Badre & Wagner, 2007; Badre, Poldrack, Paré-Blagoev, Insler, & Wagner, 2005; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997).

University of Kansas

Journal of Cognitive Neuroscience 29:7, pp. 1178–1193
doi:10.1162/jocn_a_01121

Although the contributions of prefrontal cortical systems to semantic cognition are well established, much less is known about the precise role of posterior parietal cortex in semantic retrieval, despite the reliable recruitment of this region during feature-based semantic control (Gardner et al., 2012; Whitney et al., 2011; Binder, Desai, Graves, & Conant, 2009). A possible account for posterior parietal cortex engagement in semantic cognition is its general role within the "multiple demand" executive network (Nagel et al., 2008); on the other hand, the precise executive demands supported by this area and its relationship with prefrontal regions during semantic control are not well understood (Binder et al., 2009). Substantial evidence suggests a critical role of posterior parietal cortex in goal-directed attention in the visual domain (Bichot, Heard, DeGennaro, & Desimone, 2015; Bisley & Goldberg, 2010; Corbetta & Shulman, 2002; Kanwisher & Wojciulik, 2000), according to which left posterior parietal and ventral temporal regions support the top–down modulation of visual cortex activity, creating salience maps that allow for targeted and sustained attention toward object features that are behaviorally relevant. Recent work has revealed that activity in the left IPS reflects increased attentional demands toward context-dependent, task-relevant goals that alter the saliency of different object representations during visual object processing, thus warping the representational space toward attended and away from unattended features (Woolgar, Williams, & Rich, 2015; Harel, Kravitz, & Baker, 2014; Çukur, Nishimoto, Huth, & Gallant, 2013). Further coordination between prefrontal and posterior parietal regions imposes contextual biases during perception and action that shift attentional priorities and guide behavioral decisions according to particular internal goals and environmental demands (Waskom, Kumaran, Gordon, Rissman, & Wagner, 2014).

Despite its well-established role in goal-directed orienting during visual perception as discussed above, whether the involvement of lateral parietal cortex (especially the IPS) in attentional control extends to semantic tasks remains unknown. Past work has shown that, beyond the overall contribution of inferior parietal cortex to increased semantic selection and response demands (Whitney et al., 2009; Nagel et al., 2008), patterns of activity across different subpopulations of cells within this region generate neural priority maps that represent goal-relevant stimuli features not only according to spatial (e.g., motion) but also along different, nonspatial (e.g., color) feature dimensions. Specifically, using multivoxel pattern analysis (MVPA) in a visual attention task, Liu, Hospadaruk, Zhu, and Gardner (2011) have shown that such priority maps reflect a different distribution and organization of neural signals (or patterns of activity) in the IPS depending on whether the participant is preferentially attending to color or motion, differences that are independent of the spatial location of the stimuli. It has further been shown with a similar paradigm that the anterior IPS can encode attentional priorities to changing gratings or features (i.e., color, orientation, spatial frequency) within the same visual object (Liu, 2016).

Consistent with these results, interactions between prefrontal and inferior parietal cortices may also guide attention to goal-relevant object properties outside visual perception, exerting analogous top–down influences on attentional capture for task-relevant features within semantic memory. When selection demands are high and one has to focus on specific goal-relevant properties of an item in semantic memory while discarding others (e.g., a dime's small, flat shape and rigid texture that can be relevant for tightening a screw), frontoparietal attentional networks have been proposed to act as a mechanism for filtering out any goal-irrelevant semantic features (e.g., the dime's color or its value), even in the absence of visual stimuli (Rissman & Wagner, 2012; Serences, Ester, Vogel, & Awh, 2009). Consistent with a semantic memory architecture according to which features are distributed across many sensorimotor cortical areas (Binder et al., 2016; Yee & Thompson-Schill, 2016), frontoparietal attentional networks, including the dorsolateral PFC (DLPFC) and the IPS, can selectively guide attention to task-relevant sensorimotor representations (e.g., an object's color, shape, or size) according to particular retrieval goals. Moreover, the same mechanisms can bias attention to complex or abstract semantic characteristics (e.g., an object's function or the context in which it is typically used) that match task-relevant attentional priorities, and the representation of which is supported by the selective activation of particular neuronal ensembles within a distributed semantic memory system (Rissman & Wagner, 2012; Patterson et al., 2007). Accordingly, the same attentional selection brain systems involved in visual perception may similarly improve the resolution of retrieved conceptual memories during semantic control, thus sculpting the active representational space toward goal-relevant variables that are needed for performance (Yee & Thompson-Schill, 2016; Kan & Thompson-Schill, 2004).

In line with this prediction, in this study, we employed MVPA of brain activity during a semantic judgment task to examine whether frontoparietal systems flexibly guide attention to different conceptual properties during semantic memory retrieval according to shifting attentional templates that prioritize behaviorally relevant semantic features. Participants were asked to match a cue word denoting an object to a target on the basis of a specific object property (i.e., color, shape, function, or thematic context; Figure 1A). We hypothesized that, during semantic control, frontoparietal cortical systems would respond in a manner consistent with a biased competition model of visual attention (Desimone & Duncan, 1995), with the left posterior parietal cortex dynamically modulating neural activity to accommodate attentional shifts across both concrete (i.e., color, shape) and abstract (i.e., function, or thematic context) goal-relevant

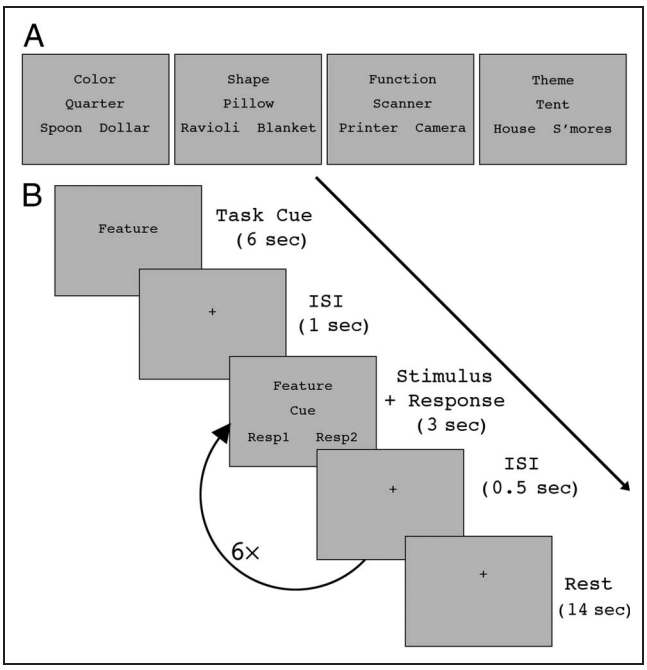

Figure 1. Experimental paradigm. (A) Sample semantic judgment tasks for each of the four object features presented in this study (color, shape, function, and theme). Seventy-four unique semantic judgment tasks were created for each object feature, for a total of 296 stimuli. (B) Sequence of events in a single block. The task-relevant feature was presented for 6 sec before stimulus onset to ready participants for the task, followed by a fixation cross to cue the onset of semantic judgment tasks. Each trial was presented for 3 sec, followed by a 0.5-sec fixation cross. Six trials were presented in each block, followed by 14 sec of rest to complete the block. Each of the 12 runs was composed of one block for each semantic feature of interest.

semantic features and the left DLPFC regulating attention toward higher-order (i.e., concrete vs. abstract) property distinctions. We used MVPA to decode the neural signatures associated with participants prioritizing particular goal-relevant semantic features during the semantic decisions task. In addition, we employed a novel Bayesian confusion matrix analysis (Olivetti, Greiner, & Avesani, 2012; Olivetti, Veeramachaneni, & Nowakowska, 2012) to examine differential neural encoding of attentional referents among frontoparietal and posterior temporal regions.

METHODS

Participants

Twelve (n = 12) right-handed native English speakers (five women, mean age = 25 years) participated in the study for $50 in compensation. Participants were recruited from the University of Kansas Lawrence and Medical Center campuses from a pool of volunteers who indicated an interest in participating in neuroimaging studies. All participants reported no neurological disorders or history of neurological illness. Three additional participants were included in the experiment; however, their data were excluded from all analyses because of excessive head movement (displacement greater that one voxel width) over the course of the scanning session. Informed consent was obtained for each participant before the experiment, in accordance with the guidelines of the University of Kansas Medical Center institutional review board.

Stimuli and Task

A semantic judgment task was used to examine the neural mechanisms that allow for decisions of semantic relatedness based on a specific, goal-relevant item property (Whitney et al., 2011; Badre et al., 2005; Thompson-Schill et al., 1997). Four item properties, or features, were selected from within two larger categories of concrete (or physical) properties and abstract (or amodal) properties. Shape and color were selected as concrete item properties. Function and theme were selected as abstract item properties. Function was defined broadly to refer to what an item is used for, whereas theme was defined as the typical context or environment that an item is found in (Figure 1A). The task followed a cue–target–distractor paradigm. In each trial, participants were instructed to view a cue word, which referred to an object. They then selected the response that best matched the cue along a given property, from a set of two response options. For example, if the property of interest was "shape," then the cue object could be "pillow," with response options of "ravioli" and "blanket." In this case, the correct answer would be "ravioli" because it matches the cue with respect to shape. The distractor stimulus was overall semantically associated with the cue but not along the property relevant to the trial. Seventy-four stimulus triplets of the kind discussed above were created for each of the four properties considered. For the neuroimaging study, an identity task, wherein both the cue and response options consisted of repeats of a single character (e.g., XXXXX), was also included to act as a baseline task. Participants simply selected the response letter string that was identical to the cue.

Stimulus Preparation

Stimuli triplets were matched for word length (mean number of syllables per word) across conditions (color = 1.74, shape = 1.62, function = 1.68, theme = 1.57). Stimuli triplets were further matched on word familiarity (mean rating: color = 528.50, shape = 541.03, function = 533.43, theme = 545.34) across conditions according to the norms obtained from the Medical Research Council psycholinguistic database (Coltheart, 1981, www.psych.rl.ac.uk). The differences between conditions across these measures were not statistically significant (all ps > .10). Because of the nature of the semantic judgment task, stimuli were not matched on imageability according to the Medical Research Council database norms (mean

rating: color = 592.91, shape = 581.47, function = 574.94, theme = 572.74); specifically, color stimuli were more imageable relative to function ($p = .02$) and theme ($p = .01$) stimuli. No other differences in imageability were statistically significant (all ps $> .23$). To quantify the strength of the relatedness of cue–target and cue–distractor pairs with respect to the property they were associated with in the behavioral task, an independent group of participants from the University of Kansas undergraduate population ($n = 40$, 15 men, mean age = 20 years) were asked to rate how well stimuli were related along their intended features. Participants performed the rating task online using Qualtrics (Qualtrics, Provo, UT) for class credit. Participants judged the level of relatedness along that property on a 7-point Likert scale, ranging from "not at all related" to "identical." Results of this norming procedure showed that participants found cue–target pairs to be significantly more closely related along the relevant property ($M = 4.95$, $SD = 0.71$) relative to the cue–distractor pairs ($M = 3.66$, $SD = 0.97$), Welch's two-sample $t(592) = 18.42$, $p < .001$. This trend held within each property set. In addition, participants were presented with cue–target pairings and were asked to select which of the five properties of interest the two objects were related by, with a sixth "none of the above" option included as well. A higher proportion of respondents chose the intended feature as the feature by which cue and target were best related ($M = 55.10\%$, $SD = 3.58\%$) significantly more than any other single feature ($M = 9.05\%$, $SD = 0.70\%$; $t(302) = 11.31$, $p < .001$).

Although we employed a block fMRI design with the aim to limit stimulus-specific effects, to ensure that our stimuli would elicit decisions of semantic relatedness based on attention to specific object properties (relative, e.g., to more global, semantic associations), we presented all stimulus triplets to an independent group of participants ($n = 57$, mean age = 19.52 years, 21 men). Half of the participants ($n = 28$) were simply asked to select which of the two response options (e.g., "ravioli" and "blanket") best matched the cue word (e.g., "pillow") overall. Reflective of our experimental task, the other half of the participants ($n = 29$) were asked to select which of the two response options (e.g., "ravioli" and "blanket") best matched the cue word (e.g., "pillow") with respect to a specific given property (e.g., shape). Participants performed the rating task online using Qualtrics for class credit. We measured the percentage of responses in which participants selected the distractor response over the property-matched response by the property of interest. The difference in the percentage of distractor responses selected within each property type was assessed with logistic regression, with the fraction of distractors selected as the dependent variable and the experimental group as the independent variable. In the absence of instructions to focus on an object property, participants selected the distractor as the correct answer significantly more often (77%) than they did when the object property was given (41%; $t(32.86) = 9.52$, $p < .001$, Cohen's $d = 2.50$; see Table 1), which confirmed the suitability of our experimental task to elicit attention to specific object properties as opposed to global semantic associations.

Design and Procedure

Before the scanning session, participants were presented with example stimuli for each semantic property they would be asked to consider during the semantic judgment task (i.e., color, shape, function, and theme) and were given an opportunity to familiarize themselves with the operational definitions of those object properties as used in this study (e.g., function was defined broadly to refer to what an object is used for; theme was defined as the typical context or environment that an object is found in). Example stimuli used during this practice session were presented in the same format as they would be in the scanner, and participants were provided with feedback

Table 1. Behavioral Results from Norming Study Showing the Effect of the Property-specific Instructions on Distractor Selection

	Color	*Shape*	*Function*	*Theme*
% Distractor Selected (SE)				
Global task ($n = 29$)	78.33 (4.41)	78.15 (4.09)	77.91 (3.58)	71.81 (3.23)
Property task ($n = 28$)	22.97 (3.30)	23.65 (2.72)	48.94 (1.35)	68.68 (2.35)
Main Effect of Feature Direction on Distractor Selection				
t statistic ($df = 56$)	−33.73	−33.28	−19.13	−2.223
p value	<.001	<.001	<.001	.10

t statistics corresponding to the main effect of task type from the logistic regression and associated Bonferroni-corrected p values are presented in addition to response percentages. Participants who did not receive instructions to focus on any specific property but performed overall or global similarity judgments chose the distractor significantly more so than participants who received the property-specific instructions at $p < .001$, for color, shape, and function. The difference was not statistically significant for theme; however, this finding is not unanticipated given that this object property is compatible with the general context in which the object is found.

on their responses to each trial to ensure that they understood the nature of the task.

Participants were scanned over 12 short runs within one scanning session. The semantic judgments were performed in blocks to allow any individual stimulus-specific effects to be averaged out of the analysis and maximize sensitivity to the attentional demands directed toward a specific object feature (Figure 1B). Each trial type (color, shape, function, or theme) was administered within one block per run. For each block, six stimulus sets were randomly selected without replacement from the list of stimulus sets for each property. Each block began with a 2-repetition time (TR; 7 sec) display informing participants of the nature of the task in the next block, followed by six trials, 3.5 sec each. Each trial was composed of a cue word and two response options. The trial type was also displayed at the top of the screen throughout the block to limit working memory demands for the maintenance of trial type for each block. Participants had 3 sec to select the correct response, which was the option that matched the cue word with respect to the trial-relevant property (color, shape, function, or theme). Each trial was followed by a 500-msec fixation. Four TRs (14 sec) of rest were included in between each block to allow for hemodynamic response normalization before the beginning of the new block. Block order was randomized within each of the 12 runs. Stimuli were designed and displayed using E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA), and they were presented against a gray background to decrease eyestrain over the 1.5-hr-long experimental session. Participants viewed the stimuli via a mirror reflecting each stimulus frame as it was projected onto a screen located in the back of the scanner. Participants indicated their response to each question by pressing one of two buttons with either the index or middle finger of the left (nondominant) hand, with the correct response buttons counterbalanced within and between blocks. E-Prime was used to record participant responses.

fMRI Data Acquisition

Participants were scanned using a Siemens (Erlangen, Germany) 3-T Skyra MRI scanner at the Hoglund Brain Imaging Center at the University of Kansas Medical Center campus, using a 12-channel surface coil. Functional scanning was divided into 12 runs, each consisting of 96 acquisitions (image matrix = 70 × 64, slice number = 54, repetition time = 3500 msec, echo time = 25 msec, flip angle = 40°, slice thickness = 3 mm, with 3 × 3 mm in-plane resolution). A high-resolution ($1 \times 1 \times 1$ mm^3) T1-weighted magnetization prepared rapid gradient echo sequence was acquired at the beginning of the session (sagittal slice orientation, centric phase encoding, image matrix = 256 × 256 mm, 176 slices with 1-mm thickness, repetition time = 2300 msec, echo time = 2.01 msec, TI = 900 msec, 9° flip angle).

Data Preprocessing and Univariate Analysis

All preprocessing and univariate analyses were performed with the FMRIB Software Library (FSL) Version 5 (Jenkinson, Beckmann, Behrens, Woolrich, & Smith, 2012; Woolrich et al., 2009; Smith et al., 2004). The first four volumes of each run were discarded to allow for the magnetic field to reach a steady state. Per-run data were first corrected for head motion and registered to the first run using the MCFLIRT motion correction tool (Jenkinson, Bannister, Brady, & Smith, 2002), followed by slice timing correction using Hanning-windowed sinc interpolation, before each voxel was high-pass filtered at 126 sec to remove low-frequency signal drift. Functional images were registered to the high-resolution anatomical image using a boundary-based registration implemented within FSL's FLIRT (Jenkinson et al., 2002), and anatomical images were registered nonlinearly to the Montreal Neurological Institute (MNI) 2-mm template brain using FNIRT within the FSL package (Andersson, Jenkinson, & Smith, 2007).

A general linear model was used to derive per-run activity estimates for each voxel. One hemodynamic response predictor was calculated for each block, based on the convolution of a single boxcar function corresponding to the duration of each block with a double-gamma hemodynamic response function. The preparatory phase of each block was included as a regressor of no interest, along with RTs. Additional confound regressors, corresponding to individual volumes corrupted by excessive motion as calculated by FSL Motion Outliers, were included. Motion outliers were calculated by taking the root mean square intensity difference between each volume and the following volume (Power, Barnes, Snyder, Schlaggar, & Petersen, 2012), with threshold for exclusion calculated via the boxplot method, where outliers were defined as those volumes with root mean square intensity differences greater than the 75th percentile plus 1.5 times the interquartile range. The resulting beta estimates were divided by their standard error to yield maps of t statistics, quantifying each voxel's activity to each condition. t statistics were chosen to maximize sensitivity to voxel-wise patterns across blocks, while simultaneously de-weighting noisy voxels (Misaki, Kim, Bandettini, & Kriegeskorte, 2010). This procedure was repeated for each run, resulting in 12 sets of t statistic images, corresponding to voxel activity for each of the object properties used to guide judgments of semantic similarity in each run.

Multivariate Searchlight Analysis

A cortical mask was constructed in native space using the automatic parcellation calculated during FreeSurfer reconstruction (Desikan et al., 2006; Fischl, 2004). Subcortical structures, including the brain stem and cerebellum, were excluded, isolating cortical structures and white matter. A spherical searchlight (SL) with a radius of four voxels was moved across a participant-specific cortical

mask, voxel by voxel, and activity patterns were extracted from within SL ROIs for each stimulus presentation block (Kriegeskorte, Goebel, & Bandettini, 2006). SL ROIs were further refined by their intersection with the whole brain mask, ensuring that only cortical voxels (including both gray and white matter) were included in the analysis. Patterns from within each SL ROI were tested for the presence of information via multivariate analysis. The choice of a four-voxel SL ROI was chosen to ensure that an adequate amount of useful information was included to maximize pattern discrimination, while avoiding overfitting due to the high dimensionality of the feature space associated with larger SL sizes. We tested both a three- and five-voxel SL on the four-way analysis, which produced qualitatively similar maps; the four-voxel analyses produced the greatest number of separate clusters after thresholding. The three-voxel SL produced comparable results, whereas the five-voxel SL created a sufficiently diffuse map in which single clusters were large and often covering much of the cortical surface, thus precluding the identification of discrete SL ROIs. The four-voxel radius SL presented the proper balance of sufficient local information and discrete spatial localization.

Classification was performed using a linear support vector machine (SVM) with a trainable c term (scaled within each training set in accordance with the norm of the data), within a leave-one-out cross-validation scheme across all 12 runs. Each run acted in turn as the testing set, whereas the remaining runs were used for training the SVM. We chose the linear SVM because it appears to perform as well as, if not better than, other classifiers in neuroimaging, especially in cases where the number of samples is relatively small (Mahmoudi, Takerkart, Regragui, Boussaoud, & Brovelli, 2012; Misaki et al., 2010; Mur, Bandettini, & Kriegeskorte, 2008). Leave-one-out cross-validation was carried out on the four-way discrimination of the semantic task attentional referent (i.e., shape, color, function, or theme) simultaneously, with accuracy calculated within each fold. Accuracies were then averaged across folds and assigned to the central ROI voxel to produce participant-specific information maps. To assess the importance of multivariate patterns for classification accuracy, we used a univariate version of the multivariate SL. To accomplish this, the mean activity was taken across all voxels with each SL ROI, and mean activation was used as the single factor driving classification within the linear SVM. We also performed a version of the multivariate SL in which features were mean-centered within each trial, removing any contributions of mean activation and retaining only pattern information. The per-SL accuracy maps associated with the univariate and mean-centered analyses were then entered into the second level analysis as described below. The rationale for this approach was to show that (1) univariate effects do not achieve comparable decoding accuracy and (2) mean-centered data still retain sufficient information to support above-chance decoding accuracy; thus, through these analyses, we can verify that we have identified a region that encodes task-relevant information in patterns of cortical activity (Coutanche, 2013).

To identify areas that consistently encode attention to task-relevant semantic information across participants, participant-specific information maps from the first level analysis were entered into a second-level analysis, the purpose of which was to detect regions that supported above-chance decoding accuracy across our study population. This procedure was implemented for the multivariate analysis, the univariate analysis, and the mean-centered multivariate analysis. To exploit undistorted high-frequency spatial information, all analyses were carried out in native space without any smoothing. The information maps were then registered to MNI space using a nonlinear transformation and tested for above-chance information content (accuracy > chance) using a nonparametric voxel-wise one-sample t test across participants (Winkler, Ridgway, Webster, Smith, & Nichols, 2014). The null distribution was obtained via exhaustive ($n = 2^{12}$) sign-flip permutation testing. A conservative voxel-wise threshold, based on the distribution of the maximum test statistic across all permuted images, was used to correct for multiple comparisons to maximize specificity to highly significant regions, with a family-wise error (FWE) corrected at $p < .05$ (Holmes, Blair, Watson, & Ford, 1996). All multivariate analyses were carried out using the MVPA in Python package (Hanke et al., 2009). Preprocessing and statistical testing of information maps were carried out using the FSL (Jenkinson et al., 2012). FreeSurfer was used for the surface reconstruction of the MNI standard brain, and both FreeSurfer and PySurfer (Ramachandran & Varoquaux, 2011, pysurfer.github.io) were used for visualization of information maps on the MNI brain surface.

Property-specific SL Analysis

To assess those regions that supported the higher-level distinction in establishing attentional priorities between abstract and concrete properties, a property-specific SL analysis following the same protocol outlined in the information-based mapping analysis section above was used to construct separate information maps specific to above-chance classification of color versus shape properties, function versus theme properties, and concrete (color and shape combined) versus abstract (function and theme combined) properties. The abstract versus concrete distinction was established by separately averaging across samples for each property within the higher-order classes, within each run. Each of these information maps was entered into a group level, nonparametric one-sample t test relative to chance classification (accuracy > 0.50), as above. Thresholding was performed with the same voxel-wise correction method used in the four-way SL analysis. However, this conservative method failed to yield any significant activation in the function versus theme classification, so a less restrictive cluster-mass-based

threshold was used, with cluster significance determined by the null distribution of cluster size across all permutations and a cluster forming threshold, F, of 2.3 (Nichols & Holmes, 2002).

Bayesian Confusion Matrix Analysis

A limitation of the classification accuracy approach in MVPA is that, in a multiway classification scheme such as the one employed in this study, above-chance discrimination could hypothetically be driven by exceptional classification of any one category, any subset of categories, or all categories equally. We used Bayesian confusion matrix analysis post hoc to address this limitation and evaluate in detail the confusion patterns within the four-way classification (Olivetti, Greiner, et al., 2012). In a Bayesian confusion matrix analysis, the contribution of each possible category pattern to multinomial classification performance constitutes a hypothesis (H_i). The posterior probability $p(H_i|C)$ represents the likelihood that a given hypothetical set of category patterns (H_i) drove classification given the observed relationship between predicted and true category labels arising from the classification. Specifically, the confusion analysis works by quantifying the statistical independence between true class labels and predicted class labels during the multinomial classification (Olivetti, Veeramachaneni, et al., 2012). In the case of perfect classification, predicted class labels are statistically dependent on true classes. If only a subset of classes are driving classification, then only part of the predicted class labels will be dependent on true class labels, whereas the remaining classes will show statistical independence. Given true and predicted class labels in the form of a confusion matrix, the probability that full dependence exists between predicted and true classes can be assessed with Bayesian methods, as can the probability of any arbitrary partially independent pattern. The Bayesian confusion matrix analysis does not rely on the detection of stronger decoding for comparison across brain regions, thus offering the possibility of an unbiased examination of differing patterns of coding of attentional referents across different regions in this study. An additional benefit of the Bayesian approach, relative to simple classification accuracy, is that, within the Bayesian analysis, all properties and their higher-order combinations can be examined simultaneously, whereas in the simple property-based analysis discussed above, information maps specific to above-chance classification of properties were constructed separately (i.e., color vs. shape properties, function vs. theme properties, and concrete vs. abstract properties).

Participant-specific Cluster Processing

The implementation of the Bayesian confusion analysis requires per-cluster confusion matrices from a participant level classification analysis. To obtain these clusters, SL-based ROIs were defined by projecting significant group level clusters from MNI space back into the native space of each participant using previously calculated transformations. These cluster-defined masks were then dilated using a spherical kernel equal in radius to the original SL ROI (four voxels) to ensure each cluster contained all voxels associated with the original SL ROIs. These expanded clusters were then entered into a participant level classification analysis to yield the per-cluster confusion matrices required for the Bayesian confusion analysis, and features for each cluster were mean-centered within conditions to rule out any contribution of univariate effects.

The classification procedure used to generate the per-cluster confusion matrices employed a linear SVM classifier within a leave-one-out cross-validation scheme, as implemented in the SL. However, unlike in the SL analysis and its fixed cluster radius, participant-specific clusters varied in size both across participants and clusters, with some clusters containing as many as 600 voxels. To prevent overfitting in larger clusters and normalize the information content driving the Bayesian confusion analysis, a voxel-based feature selection strategy was employed. Within each fold of the cross-validation, each voxel-based feature was assigned an F statistic to quantify variability between conditions based on the data within the training set. To select voxel-based features for testing within the fold without making an arbitrary selection of feature number, features were eliminated iteratively within the training set, followed by a nested cross-validation step to assess the classification performance associated with each iteration. The five voxel-based features with the lowest F statistic were discarded at each iteration step, followed by an updated assessment of classification accuracy within the training set. This procedure was carried out until the number of voxel-based features remaining reached 75. The feature set associated with the highest classification performance in the nested cross-validation was then used to assess classification accuracy based on classification of the testing set associated with the fold (Pereira, Mitchell, & Botvinick, 2009). The feature selection process was carried out only on the training data within each fold, ensuring no circularity during the within-participant feature selection step. Results of the classification were saved as confusion matrices at the per-participant level for use in the Bayesian confusion analysis. We also report the mean decoding accuracy associated with the per-cluster classification analysis.

Bayesian Confusion Analysis

When implementing the Bayesian confusion analysis, an uninformative prior probability was assigned to each of the 15 possible patterns of statistical dependencies, $p(H_i) = .067$. Probabilities associated with each confusion pattern were updated one participant at a time, with the updated assessment of probability resulting from a participant's confusion pattern being used as the prior probability for the next

participant. This strategy allowed us to take advantage of the Bayesian framework's ability to update posterior probability estimates with each additional piece of information. This resulted in a sample level estimation for the most likely pattern of dependence in each SL-based cluster. The order in which participants are entered does not have an effect on this analysis and was determined randomly. We extracted the posterior probabilities associated with four possible confusion patterns considered relevant to this study: perfect statistical dependence, partial dependence between abstract properties only, partial dependence between concrete properties only, and partial dependence between concrete and abstract properties. Note that dependence between abstract properties only or dependence between concrete properties only implies that the classifier can distinguish between abstract and concrete object properties as higher-order categories, as those models are implicitly nested within the higher-order distinction. The hypothesis that can explain the maximum statistical dependence between actual and predicted labels is given priority in the assessment of posterior probability. No other pattern of partial dependence was found to be likely given the data, $p(H_i|C) < .001$, so the remaining confusion patterns, including total statistical independence (no pattern discrimination), were aggregated into a null category by summing across their posterior probability estimates together to yield the null posterior probability, which represented the posterior probability associated with the hypothesis that no relevant confusion pattern existed in our participant population.

Variability in Bayesian Estimation

A limitation of this novel Bayesian approach to assessing class contributions to classifier performance is an inability to directly quantify the uncertainty in the estimate or the extent to which an outlier participant might influence the data. We therefore performed an ad hoc jackknife resampling procedure to enable us to estimate the stability of the posterior probability estimates for each hypothesis within each ROI. Jackknife resampling is known to be a general tool to estimate variance and bias in a wide range of situations; as such, it was the most appropriate tool for obtaining readily interpretable variance estimates in this analysis. To this end, the data from each participant in turn were excluded from the Bayesian analysis; final posterior probabilities for each pattern of class contribution to decoding performance were calculated as the average of the posterior probabilities estimated after excluding each participant in turn from the analysis. The standard error of the estimates was calculated by taking the sum of the squared deviations of each resampled parameter estimate from the mean resampled estimate, multiplied by a correction factor tied to sample size:

$$\mathrm{Var}(x) = \frac{n-1}{n}\sum_{i=1}^{n}(\bar{x}_i - \bar{x}_{\mathrm{M}})$$

where n is the sample size, $\bar{x}_i$ is the parameter estimate without the ith sample included, and $\bar{x}_{\mathrm{M}}$ is the mean of all $\bar{x}_i$ or the jackknife estimate of the parameter (Efron & Stein, 1981). The square root of this value was then taken to yield a standard error measure, rather than variance. We note that the variance estimates derived from this procedure are more useful as qualitative indications of stability in the Bayesian estimation than for any purely quantitative interpretation.

RESULTS

Behavioral Results

Mean RT for the semantic judgment task was 1897.89 msec (SD = 170.25 msec; Table 2). RTs differed significantly among conditions, $F(1.44, 15.81$, Greenhouse–Geisser correction) = 10.31, $p = .003$; post hoc comparisons (Bonferroni correction) revealed higher RTs in the shape condition relative to the color ($p < .001$), function ($p = .004$), and theme ($p = .02$) conditions. No other comparisons reached statistical significance (all ps > .23). Performance accuracy for the task was 83% (SD = 0.07%; Table 2). Accuracy also differed significantly among conditions, $F(1.88, 20.7$, Greenhouse–Geisser correction) = 20.34, $p < .001$; post hoc comparisons (Bonferroni correction) revealed lower accuracy for the theme condition relative to the color ($p < .001$), shape ($p = .001$), and function ($p = .02$) conditions. No other comparisons reached statistical significance (all ps > .39). The fMRI analysis included RTs as regressors of no interest.

Multivariate SL

We used an SL with a four-voxel radius to measure local patterns for strength of discrimination, on the basis of which object property (i.e., shape, color, function, or theme) participants were asked to attend to in making their semantic similarity judgments. Individual information maps were submitted to a nonparametric one-sample t test to identify regions that support above-chance classification (accuracy > 0.25) across all participants (Figure 2A).

Table 2. Behavioral Results across Stimulus Conditions from the Neuroimaging Study

	RT (msec)		*Proportion Accuracy*	
Condition	*M*	*SD*	*M*	*SD*
Color	1856	185	0.88	0.07
Shape	2004	164	0.86	0.06
Function	1882	172	0.84	0.07
Theme	1850	160	0.74	0.10

RT for shape was significantly higher than the other properties, and accuracy for theme was significantly lower than the other properties ($p < .05$).

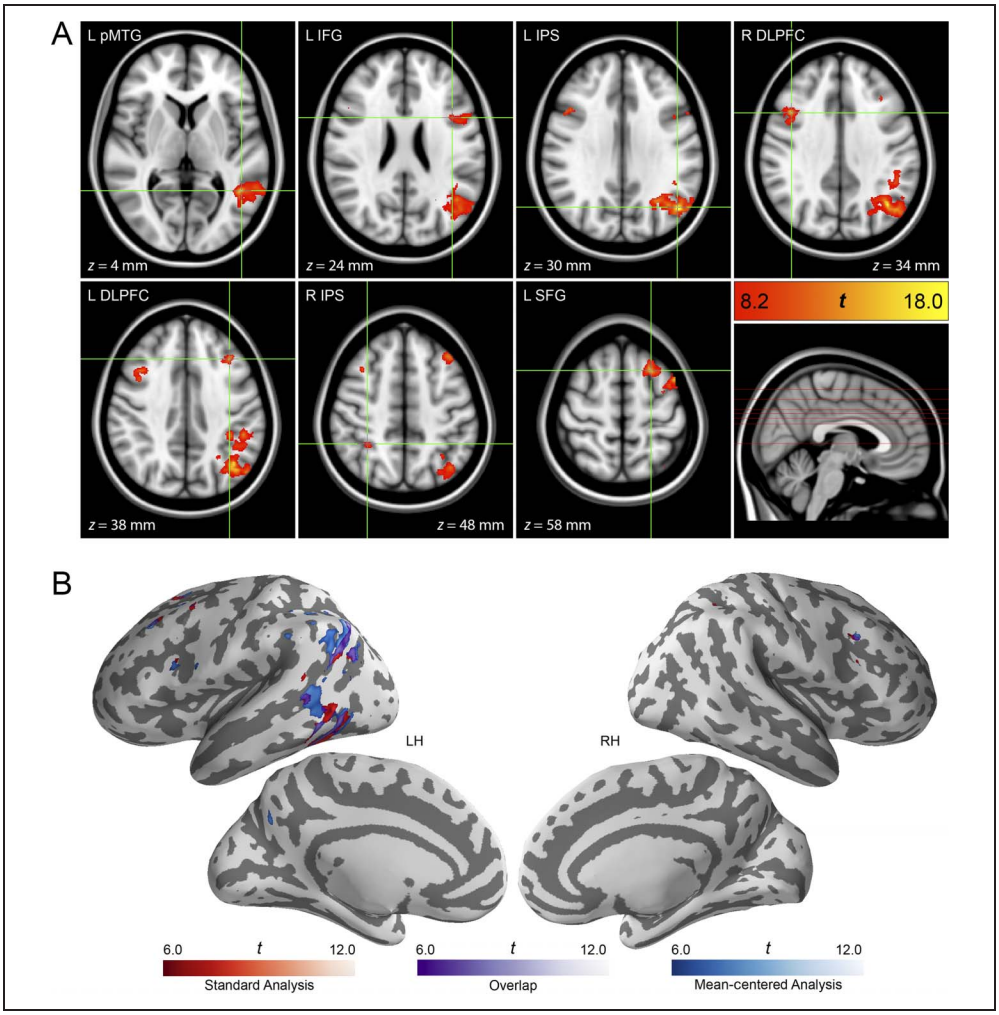

Figure 2. Group-level information maps based on task-relevant object properties. These maps were established using an SL with a radius of four voxels and represent the discriminability of goal-relevant semantic properties during the semantic judgment task. (A) This light box view shows each significant cluster in the four-way analysis. The slices represent the peak sensitivity of the seven clusters identified in that analysis. Voxel-wise FWE-corrected $p < .05$. (B) A surface projection of the four-way analysis presented in A has been included to facilitate comparisons between the four-way analysis and the property-specific SL analyses. In addition, multivariate SLs with raw (red heat map) and mean-centered (blue heat map) features are presented in conjunction. Areas of overlap are presented with a purple heat map. This conjunction shows the importance of multivariate patterns to classification, over univariate information. As in A, only significant clusters are shown after correction of the voxel-wise FWE at $p < .05$.

SL mapping revealed nine clusters sensitive to goal-relevant object properties (Table 3): two located in the left IPS, two located in the left DLPFC, one in the posterior LIFG centered in the pars opercularis, one in the left posterior middle temporal gyrus (pMTG) extending into the posterior IFG, one located in the left superior frontal gyrus (SFG), one located in the right IPS, and one within the right DLPFC. Although multivariate patterns and univariate effects need not to be mutually exclusive, to ensure that our findings were primarily due to multivariate factors and not univariate changes in average activity within SL ROIs, we repeated this analysis using per-sample mean activity within each SL ROI as the sole feature guiding classification, following the same procedures used for the multivariate SL. The results of this analysis showed univariate-driven clusters in the left frontal pole and left OFC as well as clusters that minimally overlapped with multivariate-driven clusters in the left pMTG and left DLPFC (Table 4), hence supporting the significance of multivariate pattern information over univariate changes for the encoding of attention to goal-relevant object properties in the IFG and IPS. The importance of multivariate information for classification was further reinforced by the results of a mean-centered SL analysis, in which features extracted by the SL kernel were mean-centered within each condition before classification. The clusters identified in this SL analysis were qualitatively indistinguishable from the standard SL, with the same clusters identified (see Figure 2B).

Property-specific SL Analysis

To observe areas that may participate directly in the neural encoding of information pertinent to within-category property discrimination, we employed SL analyses specific to distinctions between concrete properties (i.e., color vs. shape), abstract properties (i.e., function vs. theme), and higher-order property categories (i.e., concrete [color and shape together] vs. abstract [function and theme together] properties), following the same SL analysis procedure outlined above. To minimize influences from univariate effects, which were shown to partially guide classification in the multivariate SL analysis, all samples were mean-centered across voxel-based features within each SL ROI before classification. Statistical significance was determined at alpha = .007, two-tailed (Bonferroni correction). The results of the group level information maps showed that the same areas identified in

Table 3. Areas Sensitive to Object Properties during the Semantic Judgment Task in the Multivariate SL Analysis, Along with the Decoding Accuracy within Each Cluster

	Coordinates (Peak) in mm					
Region	*x*	*y*	*z*	*Voxels*	*Peak t Value*	*Decoding Accuracy (SE)*
L IPS, posterior	−46	−66	30	1164	18.1	0.62 (0.03)
L IPS, anterior	−36	−38	38	20	11.2	
R IPS	32	−46	48	27	10.6	0.33 (0.01)
L SFG	−24	16	58	132	15	0.40 (0.03)
L DLPFC, anterior	−36	26	38	123	14.5	0.39 (0.02)
L DLPFC, posterior	−44	8	58	75	14.6	
R DLPFC	42	14	34	135	15.3	0.36 (0.02)
L IFG	−40	10	24	75	11.1	0.46 (0.03)
L pMTG	−46	−52	4	1040	15.6	0.57 (0.03)

These areas are the left and right IPS, left SFG, left and right DLPFC, left IFG (pars opercularis), and the left pMTG. Coordinates are presented in MNI 152 nonlinear template space. L = left; R = right.

the four-way SL were largely present when considering each of the specific property categories (Figure 3). The results of these property-specific SLs are also presented, along with per-cluster decoding accuracy, in Table 5.

Areas sensitive to concrete item properties (shape vs. color) were identified in the left pMTG ($x = -44$, $y = -56$, $z = 0$; 383 voxels; peak t value = 17.7) and left IPS ($x = -48$, $y = -34$, $z = 38$; 263 voxels; peak t value = 16.6; all coordinates are reported in MNI 152 nonlinear template space). The abstract property SL failed to yield significant results after the conservative voxel-wise correction for multiple comparisons; however, a more permissive corrected cluster-based threshold ($F = 2.3$, $p_{corr} < .05$) revealed a network within the posterior parietal cortex, including the left IPS, that was sensitive to the distinction between abstract properties (function vs. theme; Figure 3B). Specifically, areas sensitive to abstract item properties (function vs. theme) were identified in the left IPS ($x = -36$, $y = -50$, $z = 44$; 1744 voxels; peak t value = 5.86), left superior parietal lobe ($x = -28$, $y = -54$, $z = 60$; 438 voxels; peak t value = 5.67), right lateral occipital cortex ($x = 30$, $y = -66$, $z = 28$; 333 voxels; peak t value = 5.65), and cuneus bilaterally ($x = -2$, $y = -78$, $z = 26$; 1694 voxels; peak t value = 8.98). Areas sensitive to concrete versus abstract item properties were identified: left AG ($x = -46$, $y = -60$, $z = 30$; 3000 voxels; peak t value = 21), left pMTG ($x = -52$, $y = -46$, $z = 2$; 1407 voxels; peak t value = 22.9), left ($x = -34$, $y = 18$, $z = 58$; 1386 voxels; peak t value = 14.6) and right ($x = 38$, $y = 18$, $z = 44$; 477 voxels; peak t value = 14.5) DLPFC, left precuneus ($x = -10$, $y = -68$, $z = 40$; 476 voxels; peak t value = 13.4), and right IPS ($x = 41$, $y = -54$, $z = 42$; 297 voxels; peak t value = 14.9). According to these results, the left IPS was the single region that was sensitive to distinctions in attention within both concrete (color vs. shape) and abstract (function vs. theme) item properties but not higher-order properties (abstract vs. concrete properties). In contrast, clusters in the right and left DLPFC were present for the higher-order category level (i.e., abstract vs. concrete properties) classification map (Figure 3C) but not for the concrete (color vs. shape; Figure 3A) or the abstract (theme vs. function; Figure 3B) classification maps.[1] Overall, the results of the property-specific SL analysis were in line with those of the multivariate SL analysis, establishing the differential contribution of frontoparietal regions to different aspects of feature-based semantic control.

Table 4. Areas with Univariate Patterns of Activity Sensitive to Object Properties during the Semantic Judgment Task

	Coordinates (Peak) in mm				
Region	*x*	*y*	*z*	*Voxels*	*Peak t Value*
L pMTG	−64	−50	0	37	6.77
L DLPFC	−32	16	54	28	6.38
L FP	−42	52	2	49	6.93
L OFC	−50	32	−14	33	7.02

These areas are the left pMTG, left DLPFC, left fontal pole (FP), and left OFC. Coordinates are presented in MNI 152 nonlinear template space.

Bayesian Confusion Matrix Analysis

The results of the Bayesian confusion matrix analysis revealed that the left pMTG, the left IFG, and the left SFG were maximally sensitive to the distinction between the concrete properties (i.e., color vs. shape; $p(H_i|C) = 0.99$, $SE = 0.01$, for the left pMTG; $p(H_i|C) = 1.00$, $SE < 0.001$,

for the left IFG; and $p(H_i|C)$ = .85, SE = 0.27, for the left SFG), whereas the left DLPFC was maximally sensitive to the distinction between abstract (i.e., function and theme together) versus concrete (i.e., color and shape together) higher-level categories, $p(H_i|C)$ = .88, SE= 0.20, while confusing exact property distinctions within these categories (i.e., color vs. shape, function vs. theme). The left IPS was equally sensitive to the distinction between concrete properties only, $p(H_i|C)$ = .56, SE = 0.62, and the distinction across all properties, $p(H_i|C)$ = .44, SE = 0.62. Table 6 presents the posterior probabilities and error estimates for each hypothesis within each mask. The right DLPFC and right IPS did not seem to encode any semantic property distinction consistently across participants, $p(H_{null}|C) > .99$.

Relative to the simple classification accuracy SL analyses discussed above, the Bayesian approach allowed us to examine all four attentional referents (i.e., color, shape, function, and theme) and their higher-order combinations (color and shape, function and theme) simultaneously, without artificially altering the representational space of the data by separately constructing maps specific to the above-chance classification of properties. Overall, the results of the Bayesian confusion matrix analysis revealed that only the left IPS was equally likely to contribute to the discrimination between basic (color vs. shape) and higher-order (concrete vs. abstract) item properties, whereas only the left DLPFC was sensitive to the distinction between higher-order (concrete vs. abstract) item properties, thus clarifying the results of the SL analyses (Figure 4).

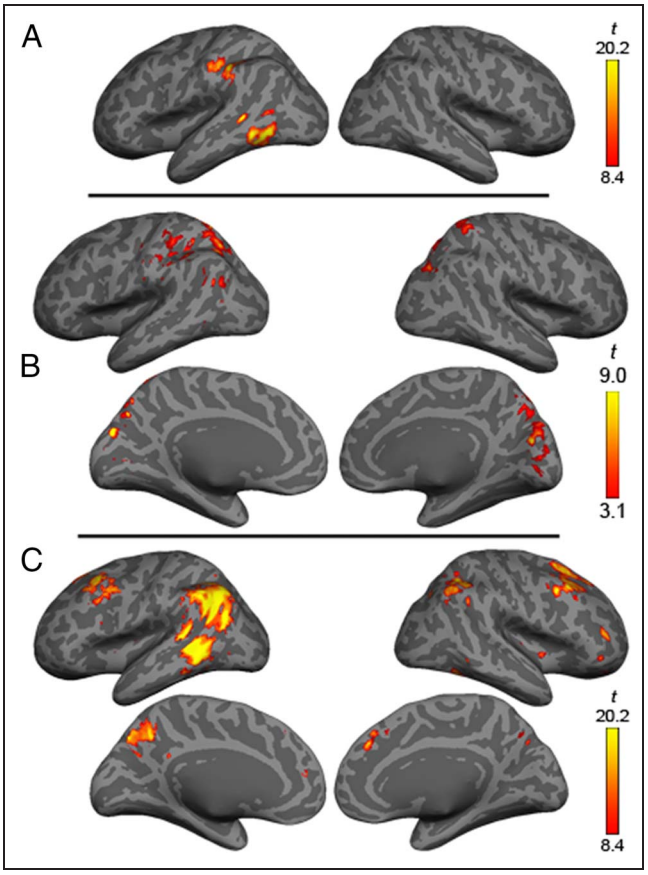

Figure 3. Group classification maps for property-specific classifications of (A) concrete properties (i.e., color vs. shape), (B) abstract properties (i.e., theme vs. function), and (C) higher-order property category (i.e., concrete vs. abstract). Both A and C have been corrected for multiple comparisons using a voxel-wise correction based on the distribution of the max statistic across permutations; for B, a more liberal, cluster-based threshold (cluster-forming threshold, F = 2.3) was used. Maps represent classification accuracy for multivariate data with mean activity for each sample removed, ensuring that the classification does not include any effects of univariate activation.

Table 5. Areas with High Decoding Accuracy for Specific Property Pairs, in the Post Hoc Two-Way Classification SL Analysis, Along with the Decoding Accuracy within Each Cluster

	Coordinates (Peak) in mm					
Region	*x*	*y*	*z*	*Voxels*	*Peak t Value*	*Decoding Accuracy (SE)*
Color vs. Shape						
L pMTG	−44	−56	0	383	17.7	0.79 (0.02)
L IPS	−48	−34	38	263	16.6	0.77 (0.02)
Theme vs. Function						
L IPS	−36	−50	44	1744	5.86	0.58 (0.01)
L SPL	28	−54	60	438	5.67	0.59 (0.02)
R LOC	30	−66	28	333	5.65	0.59 (0.01)
CUN	−2	−78	26	1694	8.98	0.59 (0.02)
Abstract vs. Concrete						
L AG	−46	−60	30	3000	21	0.69 (0.01)
L pMTG	−52	−46	2	1407	22.9	0.71 (0.01)
L DLPFC	−34	18	58	1386	14.6	0.68 (0.01)
R DLPFC	38	18	44	477	14.5	0.76 (0.02)
L PC	−10	−68	40	476	13.4	0.80 (0.02)
R IPS	41	−54	42	297	14.9	0.79 (0.01)

These areas are the left IPS, left superior parietal lobe (SPL), right lateral occipital cortex (LOC), and cuneus bilaterally (CUN) in the "theme versus function" classification; the left pMTG and left IPS in the "color versus shape" classification; and the left AG, left pMTG, left and right DLPFC, left precuneus (PC), and right IPS in the "abstract versus concrete" classification. Coordinates are presented in MNI 152 nonlinear template space.

DISCUSSION

Despite their importance for several aspects of goal-directed behavior, the neural mechanisms supporting attention to

Table 6. Estimates of Posterior Probabilities Associated with Each Hypothesis within Each Mask

	Null	*AvC*	*TvF*	*CvS*	*All*
Left IPS	<0.001 (<0.001)	<0.001 (<0.001)	<0.001 (<0.001)	0.560 (0.618)	0.440 (0.618)
Left pMTG	<0.001 (<0.001)	<0.001 (<0.001)	<0.001 (<0.001)	0.992 (0.010)	0.008 (0.010)
Left IFG	<0.001 (<0.001)	<0.001 (<0.001)	<0.001 (<0.001)	0.999 (<0.001)	<0.001 (<0.001)
Left SFG	<0.001 (<0.001)	<0.001 (0.004)	<0.001 (0.001)	0.846 (0.274)	0.153 (0.272)
Left DLPFC	<0.001 (<0.001)	0.883 (0.197)	0.115 (0.196)	0.002 (0.002)	<0.001 (<0.001)

Values have been determined via a jackknife resampling procedure, where each of the 12 participants was left out of the Bayesian confusion hypothesis estimation procedure in turn, and results were averaged across samples. The values in parentheses are the jackknife estimated standard errors for the posterior probabilities, demonstrating the relative instability between different hypotheses. AvC = abstract vs. concrete properties; TvF = theme vs. function; CvS = color vs. shape.

goal-relevant properties during semantic cognition are not fully understood. Here, we employed a semantic judgment task to investigate whether the cortical networks that have been shown to guide attention to task-relevant stimulus properties during visual perception also shape attentional priorities toward behaviorally relevant object attributes during semantic retrieval. Because of the subtlety of feature-guided attentional shifts that would make their capturing with conventional fMRI methods elusive, we used MVPA to quantify the discriminability of multivoxel patterns pertaining to participants' attention toward different, goal-relevant semantic properties, with classification accuracy reflecting the relative strength of neural coding for attention to a particular property within each region. Our results implicate an extensive network within the left prefrontal, posterior parietal, and ventral temporal cortices in semantic control but also reveal diverse sensitivity profiles among these regions in guiding attention to different conceptual features during semantic retrieval. In line with our predictions, our four-way SL classification scheme revealed the contributions of several areas within the "multiple demand" network (Duncan, 2010; Owen et al., 2000) in this task, particularly within the left IFG and the left IPS. A property-specific SL analysis offered further support for these results, highlighting a key role for the left posterior parietal cortex in feature-guided semantic retrieval: The left IPS and the left pMTG were sensitive to the distinction between concrete item properties (shape vs. color); a network within the posterior parietal cortex, including the left IPS, was further sensitive to the distinction between abstract properties (function vs. theme), whereas the left AG and areas within the left pMTG and DLPFC bilaterally showed sensitivity to the distinction between concrete versus abstract item properties. Overall, the results of the multivariate and property-specific SL analyses support the conclusion that the left IPS is sensitive to distinctions in attention within both concrete (color vs. shape; Figure 3A) and abstract (function vs. theme; Figure 3B) item properties, but not higher-order properties (abstract vs. concrete properties; Figure 3C). In contrast, clusters within the DLPFC are sensitive to distinctions in attention only within the higher-order category level (i.e., abstract

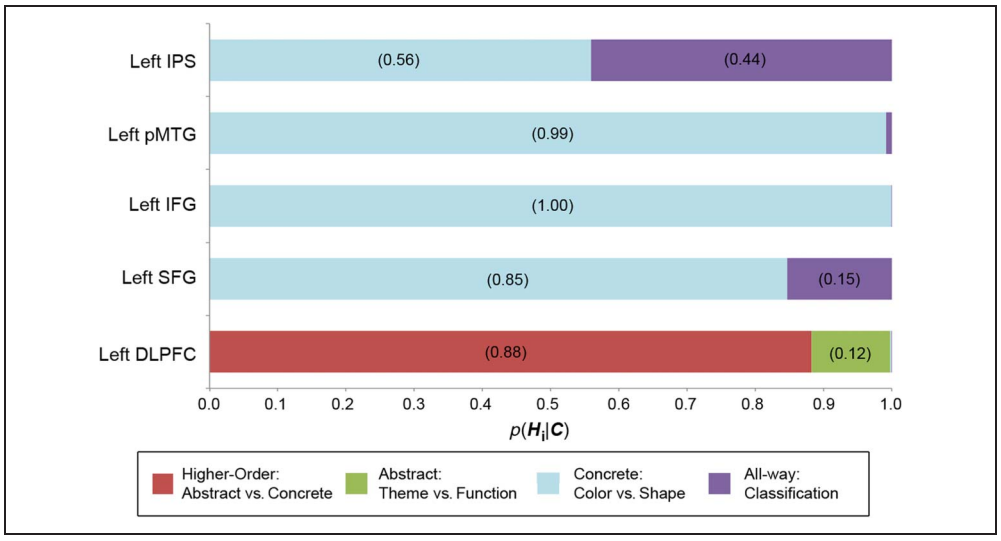

Figure 4. Results of the Bayesian confusion matrix analysis. Jackknifed posterior probabilities are shown for the five relevant classification schemes across SL ROIs. Median posterior probabilities are presented in the figure for classification schemes with nontrivial probability ($p(H_i|C) > .01$) in each ROI.

vs. concrete) properties (Figure 3C), but not within the concrete (color vs. shape; Figure 3A) or the abstract (theme vs. function; Figure 3B) properties.

In line with these results, a post hoc Bayesian confusion matrix analysis showed evidence for activity patterns reflecting attention across all four semantic features (i.e., color, shape, function, and theme) only within the left IPS. The left IFG and pMTG were sensitive primarily to attentional priorities toward concrete object properties (i.e., color, shape), whereas the left DLPFC showed increased discriminability between attention to concrete and abstract properties, only when each higher-order property category was considered as a whole. As hypotheses were inherently nested within each other in this analysis (i.e., dependence between concrete properties implied that the classifier could distinguish between abstract and concrete object properties as higher-order categories), the classifier was able to distinguish between color and shape while confusing theme and function but was also concurrently able to distinguish between concrete and abstract features. The results of the Bayesian confusion analysis clarify the outcome of the multivariate and property-based SL analyses but also offer additional insights regarding the role of the IPS in feature-based semantic control. Although the property-based SL offered some evidence for IPS involvement in attention to abstract feature distinctions (i.e., function vs. theme), the Bayesian confusion matrix analysis was more agnostic on the point, in that both the "concrete only" and "equal contributions" hypotheses were similarly plausible for this region. One possible interpretation of this finding is that it reflects larger individual variability in the engagement of these attentional spotlight mechanisms for abstract, relative to concrete, semantic features during semantic control. An alternative possibility is that neural encoding of attention to abstract semantic features is more distributed or heterogeneous relative to that of concrete features (at least for the features of function and theme used here). Overall, the results of the SL and Bayesian confusion matrix analyses suggest that the IPS may support attention to abstract semantic properties but that its contributions may be more variable or nuanced relative to the engagement of this area during attention to concrete, goal-relevant semantic features.

We note that the stimuli employed in this study were matched on word length and familiarity across conditions. Inherent to the nature of our semantic task were differences in imageability (i.e., stimuli in the color condition were more imageable than stimuli in the function and theme conditions); nevertheless, these differences are unlikely to have influenced our results given that our classifier was able to distinguish between conditions that did not differ on imageability (e.g., color vs. shape, theme vs. function). Stimuli may have varied on other dimensions (e.g., concreteness, frequency in written text), which might have influenced the results. We attempted to curtail the possibility of stimulus-specific effects with the block design of the study, which allowed any individual stimulus influence to be averaged out of the analysis, thus prioritizing only effects due to the attentional demands directed toward a specific object property. In addition, the inclusion of RTs as regressors of no interest limits the likelihood that our findings are guided by the behavioral differences observed for certain condition pairs. Although it is possible that stimulus-specific effects, to an extent, may have influenced our results, the block design, RT regressors, and variability in our stimulus set to ensure lack of any systematic confounding relationships among items render any contributions of stimuli effects to our univariate and multivariate classifications negligible. The inclusion of the attentional cue in the trials, which was employed to limit working memory demands for each block, may have influenced, to some degree, decoding performance. Although it is possible that classification accuracy was driven, in part, by the cue itself, the high accuracy in participant performance on the semantic judgment task would imply that participants used the cues to perform the task, as opposed to focusing on the cues per se.

Our findings are consistent with a substantial body of work linking VLPFC regions to semantic control (Whitney et al., 2009, 2011; Badre & Wagner, 2007; Badre et al., 2005; Thompson-Schill et al., 1997). Critically, however, this study is the first to demonstrate the involvement of frontoparietal attentional systems, typically implicated in visual perception, in a purely semantic memory task: Our analyses show that neuronal clusters within the left posterior parietal cortex support the modulation of attention beyond spatial or perceptual object properties and along both concrete and abstract semantic features during conceptual processing (Liu et al., 2011). We note that, although our results offer support for a domain-general system that is active in shifting attention toward concrete (or perceptually based) and abstract semantic features, conclusions regarding the involvement of this system in distinguishing among abstract semantic features from these data were less definitive. Additional work is required to clarify precisely the kinds of abstract semantic features represented within this hypothesized attentional system. Moreover, although our current resolution parameters did not allow us to test this hypothesis directly, future work employing high-resolution scanning protocols may be able to identify feature-selective attentional signals that would support some domain specificity according to a particular topographical organization within the left IPS. Such studies may also be in better position to distinguish the exact contributions among multiple regions within the inferior parietal cortex (e.g., AG, supramarginal gyrus, IPS) to context integration and semantic control (Binder et al., 2009). Similarly, our property-specific SL and Bayesian confusion matrix analyses revealed sensitivity to different semantic properties across lateral frontal, posterior temporal, and lateral parietal regions. On the other hand, the precise interactions

among these different areas for semantic control still remain unclear. Future studies employing methodologies with higher temporal resolution may be able to identify any hierarchical involvement of different regions within the frontoparietal network in guiding attention to goal-relevant properties during semantic tasks.

Our results point to a critical role of frontoparietal regions, especially the left IPS, for feature-based semantic control. Adaptive representations within the human frontoparietal cortex may adjust dynamically to task demands and thus serve as a source of bias within a distributed semantic memory system to sharpen attention toward goal-relevant features (Woolgar, Hampshire, Thompson, & Duncan, 2011). The pattern of activity in the inferior parietal cortex during semantic control mirrors evidence from studies of visual attention in nonhuman primates showing that neurons within this region may be sensitive to general distinctions among categorical outcomes (e.g., Fitzgerald, Freedman, & Assad, 2011). Our results are further consistent with functional connectivity analyses of human functional neuroimaging data revealing that selectivity in the recruitment of visual cortical areas for the processing of specific features during visual attention is functionally coupled with activity in frontoparietal regions when these features are task relevant (Chadick & Gazzaley, 2011). On the basis of this past work and the patterns of activity observed during our semantic judgment task, we hypothesize that neurons within the left IPS flexibly shift their activity according to task goals and transform their selectivity toward particular semantic features in line with more abstract representations of behavioral relevance and irrespective of the characteristics of individual stimuli (Woolgar, Thompson, Bor, & Duncan, 2011; Whitney et al., 2009; Freedman & Assad, 2006). According to this prediction, activity in this region during semantic control can support attention to specific task-relevant object attributes (e.g., a dime's shape when used to tighten a screw or its value when used to pay a parking meter), thus guiding the particular instantiation of the object in one's active conceptual space in working memory (see Hindy, Solomon, Altmann, & Thompson-Schill, 2015). As such, they comprise a flexible attentional mechanism that supports independent instantiations of retrieved semantic representations for the same concept, each highlighting different, context-dependent, and goal-relevant semantic dimensions.

Contrary to past work primarily focusing on the role of frontal and temporal networks in semantic retrieval, in this study, we used MVPA to decode—for the first time within semantic memory—the neural signatures of attentional priorities toward particular semantic features. The results of multivariate and property-specific SL analyses, in conjunction with a novel Bayesian confusion matrix analysis (Olivetti, Greiner, et al., 2012; Olivetti, Veeramachaneni, et al., 2012) revealed, according to our predictions, differential encoding of attentional referents among frontoparietal and posterior temporal regions, with DLPFC encoding attention to higher-order property classes (abstract vs. concrete features), inferior frontal and posterior temporal cortices encoding attention to concrete properties (color vs. shape), and left posterior parietal cortex (IPS) encoding attention across all object properties. These results strongly support a highly dynamic view of semantic control, according to which frontoparietal systems specify current attentional priorities and carve different instantiations of the same stimulus during semantic retrieval, depending on the salience status of particular goal-relevant semantic features.

Reprint requests should be sent to Evangelia G. Chrysikou, Department of Psychology, University of Kansas, 1415 Jayhawk Blvd., m 426 Fraser Hall, Lawrence, KS 66049, or via e-mail: lilachrysikou@gmail.com.

Note

1. We note that, at the more permissive corrected cluster-based threshold ($F = 2.3, p_{corr} < .05$) employed for the distinction between abstract properties (theme vs. function), the distinction between the concrete properties (color vs. shape) also revealed activity in the left IFG, a finding in parallel with the results of the Bayesian analysis.

REFERENCES

Andersson, J., Jenkinson, M., & Smith, S. (2007). Non-linear registration, aka spatial normalisation (Technical Report TR07JA2). Oxford Centre for Functional Magnetic Resonance Imaging of the Brain, Department of Clinical Neurology, Oxford University, Oxford, UK.

Badre, D., Poldrack, R. A., Paré-Blagoev, E. J., Insler, R. Z., & Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron, 47,* 907–918.

Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia, 45,* 2883–2901.

Bichot, N. P., Heard, M. T., DeGennaro, E. M., & Desimone, R. (2015). A source for feature-based attention in the prefrontal cortex. *Neuron, 88,* 832–844.

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex, 19,* 2767–2796.

Binder, J. R., Conant, L. L., Humphries, C. J., Fernandino, L., Simons, S. B., Aguilar, M., et al. (2016). Toward a brain-based componential semantic representation. *Cognitive Neuropsychology,* 1–45.

Bisley, J. W., & Goldberg, M. E. (2010). Attention, intention, and priority in the parietal lobe. *Annual Review of Neuroscience, 33,* 1–21.

Chadick, J. Z., & Gazzaley, A. (2011). Differential coupling of visual cortex with default or frontal-parietal network based on goals. *Nature Neuroscience, 14,* 830–832.

Chao, L. L., & Martin, A. (2000). Representation of manipulable man-made objects in the dorsal stream. *Neuroimage, 12,* 478–484.

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology, 33A,* 497–505.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3,* 215–229.

Coutanche, M. N. (2013). Distinguishing multi-voxel patterns and mean activation: Why, how, and what does it tell us? *Cognitive, Affective & Behavioral Neuroscience, 13,* 667–673.
Çukur, T., Nishimoto, S., Huth, A. G., & Gallant, J. L. (2013). Attention during natural vision warps semantic representation across the human brain. *Nature Neuroscience, 16,* 763–770.
Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage, 31,* 968–980.
Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18,* 193–222.
Dosenbach, N. U. F., Visscher, K. M., Palmer, E. D., Miezin, F. M., Wenger, K. K., Kang, H. C., et al. (2006). A core system for the implementation of task sets. *Neuron, 50,* 799–812.
Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour. *Trends in Cognitive Sciences, 14,* 172–179.
Efron, B., & Stein, C. (1981). The jackknife estimate of variance. *Annals of Statistics, 9,* 586–596.
Fischl, B. (2004). Automatically parcellating the human cerebral cortex. *Cerebral Cortex, 14,* 11–22.
Fitzgerald, J. K., Freedman, D. J., & Assad, J. A. (2011). Generalized associative representations in parietal cortex. *Nature Neuroscience, 14,* 1075–1079.
Freedman, D. J., & Assad, J. A. (2006). Experience-dependent representation of visual categories in parietal cortex. *Nature, 443,* 85–88.
Gardner, H. E., Lambon Ralph, M. A., Dodds, N., Jones, T., Ehsan, S., & Jefferies, E. (2012). The differential contributions of pFC and temporo-parietal cortex to multimodal semantic control: Exploring refractory effects in semantic aphasia. *Journal of Cognitive Neuroscience, 24,* 778–793.
Hanke, M., Halchenko, Y. O., Sederberg, P. B., Hanson, S. J., Haxby, J. V., & Pollmann, S. (2009). PyMVPA: A Python toolbox for multivariate pattern analysis of fMRI data. *Neuroinformatics, 7,* 37–53.
Harel, A., Kravitz, D. J., & Baker, C. I. (2014). Task context impacts visual object processing differentially across the cortex. *Proceedings of the National Academy of Sciences, U.S.A., 111,* E962–E971.
Hindy, N. C., Solomon, S. H., Altmann, G. T., & Thompson-Schill, S. L. (2015). A cortical network for the encoding of object change. *Cerebral Cortex, 25,* 884–894.
Holmes, A. P., Blair, R. C., Watson, J. D., & Ford, I. (1996). Nonparametric analysis of statistic images from functional mapping experiments. *Journal of Cerebral Blood Flow & Metabolism, 16,* 7–22.
Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage, 17,* 825–841.
Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). FSL. *Neuroimage, 62,* 782–790.
Kan, I. P., & Thompson-Schill, S. L. (2004). Selection from perceptual and conceptual representations. *Cognitive, Affective & Behavioral Neuroscience, 4,* 466–482.
Kanwisher, N., & Wojciulik, E. (2000). Visual attention: Insights from brain imaging. *Nature Reviews Neuroscience, 1,* 91–100.
Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences, U.S.A., 103,* 3863–3868.
Liu, T. (2016). Neural representation of object-specific attentional priority. *Journal of Neuroscience, 31,* 4484–4495.
Liu, T., Hospadaruk, L., Zhu, D. C., & Gardner, J. L. (2011). Feature-specific attentional priority signals in human cortex. *Neuroimage, 129,* 15–24.
Mahmoudi, A., Takerkart, S., Regragui, F., Boussaoud, D., & Brovelli, A. (2012). Multivoxel pattern analysis for fMRI data: A review. *Computational and Mathematical Methods in Medicine, 2012,* 1–14.
Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology, 58,* 25–45.
Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24,* 167–202.
Misaki, M., Kim, Y., Bandettini, P. A., & Kriegeskorte, N. (2010). Comparison of multivariate classifiers and response normalizations for pattern-information fMRI. *Neuroimage, 53,* 103–118.
Mur, M., Bandettini, P. A., & Kriegeskorte, N. (2008). Revealing representational content with pattern-information fMRI—An introductory guide. *Social Cognitive and Affective Neuroscience, 4,* 101–109.
Nagel, I. E., Schumacher, E. H., Goebel, R., & D'Esposito, M. (2008). Functional MRI investigation of verbal selection mechanisms in lateral prefrontal cortex. *Neuroimage, 43,* 801–807.
Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping, 15,* 1–25.
Olivetti, E., Greiner, S., & Avesani, P. (2012). Testing multiclass pattern discrimination. In *International Workshop on Pattern Recognition in Neuroimaging (PRNI)* (pp. 57–60). Washington, DC: IEEE Computer Society.
Olivetti, E., Veeramachaneni, S., & Nowakowska, E. (2012). Bayesian hypothesis testing for pattern discrimination in brain decoding. *Pattern Recognition, 45,* 2075–2084.
Owen, A. M., Schneider, W. X., & Duncan, J. (2000). Executive control and the frontal lobe: Current issues. *Experimental Brain Research, 133,* 1–2.
Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience, 8,* 976–987.
Pereira, F., Mitchell, T., & Botvinick, M. (2009). Machine learning classifiers and fMRI: A tutorial overview. *Neuroimage, 45,* S199–S209.
Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage, 59,* 2142–2154.
Ramachandran, P., & Varoquaux, G. (2011). Mayavi: 3D visualization of scientific data. *Computing in Science & Engineering, 13,* 40–51.
Rissman, J., & Wagner, A. D. (2012). Distributed representations in memory: Insights from functional brain imaging. *Annual Review of Psychology, 63,* 101–128.
Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science, 20,* 207–214.
Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage, 23,* S208–S219.
Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences, U.S.A., 94,* 14792–14797.
Waskom, M. L., Kumaran, D., Gordon, A. M., Rissman, J., & Wagner, A. D. (2014). Frontoparietal representations of task context support the flexible control of goal-directed cognition. *Journal of Neuroscience, 34,* 10743–10755.

Whitney, C., Grossman, M., & Kircher, T. T. J. (2009). The influence of multiple primes on bottom–up and top–down regulation during meaning retrieval: Evidence for 2 distinct neural networks. *Cerebral Cortex, 19,* 2548–2560.
Whitney, C., Kirk, M., O'Sullivan, J., Lambon Ralph, M. A., & Jefferies, E. (2011). The neural organization of semantic control: TMS evidence for a distributed network in left inferior frontal and posterior middle temporal gyrus. *Cerebral Cortex, 21,* 1066–1075.
Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., & Nichols, T. E. (2014). Permutation inference for the general linear model. *Neuroimage, 92,* 381–397.
Woolgar, A., Hampshire, A., Thompson, R., & Duncan, J. (2011). Adaptive coding of task-relevant information in human frontoparietal cortex. *Journal of Neuroscience, 31,* 14592–14599.
Woolgar, A., Thompson, R., Bor, D., & Duncan, J. (2011). Multi-voxel coding of stimuli, rules, and responses in human frontoparietal cortex. *Neuroimage, 56,* 744–752.
Woolgar, A., Williams, M. A., & Rich, A. N. (2015). Attention enhances multi-voxel representation of novel objects in frontal, parietal and visual cortices. *Neuroimage, 109,* 429–437.
Woolrich, M. W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., et al. (2009). Bayesian analysis of neuroimaging data in FSL. *Neuroimage, 45,* S173–S186.
Yee, E., Chrysikou, E. G., & Thompson-Schill, S. L. (2014). Semantic memory. In K. Ochsner & S. Kosslyn (Eds.), *The Oxford Handbook of Cognitive Neuroscience* (pp. 353–374). Oxford University Press.
Yee, E., & Thompson-Schill, S. L. (2016). Putting concepts into concepts into context. *Psychonomic Bulletin and Review, 23,* 1015–1027.