Where Vision Meets Memory: Prefrontal–Posterior Networks for Visual Object Constancy during Categorization and Recognition

Objects seen from unusual relative to more canonical views require more time to categorize and recognize, and, according to object model verification theories, additionally recruit prefrontal processes for cognitive control that interact with parietal processes for mental rotation. To test this using fundamental functional magnetic resonance imaging, people categorized and recognized known objects from unusual and canonical views. Canonical views activated some components of a default network more on categorization than recognition. Activation to unusual views showed that both ventral and dorsal visual pathways, and prefrontal cortex, have key roles in visual object constancy. Unusual views activated object-sensitive and mental rotation (and not saccade) regions in ventrocaudal intraparietal, transverse occipital, and inferotemporal sulci, and ventral premotor cortex for verification processes of model testing on any task. A collateral–lingual sulci “place” area activated for mental rotation, working memory, and unusual views on correct recognition and categorization trials to accomplish detailed spatial matching. Ventrolateral prefrontal cortex and object-sensitive lateral occipital sulcus activated for mental rotation and unusual views on categorization more than recognition, supporting verification processes of model prediction. This visual knowledge framework integrates vision and memory theories to explain how distinct prefrontal–posterior networks enable meaningful interactions with objects in diverse situations.

Keywords: long-term memory, occipital cortex, parietal cortex, spatial cognition, temporal lobe

Introduction

People can quickly and accurately categorize objects (e.g., horse, plane, or pen) and recognize them as familiar, despite changing viewing conditions, an ability known as object constancy. Yet, people do worse when visual input is relatively impoverished, as when objects are occluded or viewed at unusual angles (Palmer et al. 1981; Tarr et al. 1998). Of the few vision frameworks explaining object cognition under impoverished input, perceptual hypothesis testing theories propose that top-down processes in a frontoparietal network drive iterative processes of object model verification that determine the stored structural model in posterior object-sensitive cortex that best accounts for the visual percept (Lowe 1985, 2000; Kosslyn et al. 1994; Grossberg 1999; Lee and Mumford 2003; Ganis et al. 2007). One class of such theories, multiple views plus transformation (MVPT) (Bülthoff et al. 1995; Ullman 1996), would predict that the role of the parietal part of the network is a transformation process, such as mental rotation (Tarr and Pinker 1989), that is required for unusual views of objects to bring the visual input and stored representation(s) into sufficient visuospatial correspondence to achieve a match. To test key predictions of model verification and MVPT, brain activity was compared between canonical (best) and unusual views of real objects during categorization and recognition using functional magnetic resonance imaging (fMRI). Although vision frameworks tend to posit generalization across tasks, memory frameworks predict task dissociations because distinct brain systems support perceptual and semantic, generic memory for categorization versus episodic memory for recognition (Hintzman 1978; Schacter and Tulving 1994).

Feedforward computational models explain rapid detection of a target category (i.e., animal or not), but model verification theories can better explain object cognition under impoverished viewing conditions (Serre et al. 2007). According to the latter, impoverished images of objects recruit model verification processes implemented in ventralprefrontal cortex (VLPFC) and parietal networks that are involved in top-down control of long-term memory retrieval, working memory, and covert attention processes and modulate activity in ventral object-sensitive regions, with greater engagement as image impoverishment increases (Ganis et al. 2007). Verification processes are recruited after the initial bottom-up pass through the visual system provides only weak candidate object models for the match with the more impoverished input. Similarly, MVPT and related psychophysical and computational accounts also include iterative model testing, but posit processes and representations that enable more specific hypotheses about the role of posterior cortex, despite being largely agnostic about neural correlates (Bülthoff et al. 1995; Peissig and Tarr 2006). Object constancy is posited to be achieved by storing multiple views of an object, consistent with findings of view-dependent performance on diverse tasks (Tarr et al. 1998). A familiar or canonical view can be matched directly to stored views. When an object is instead encountered from a noncanonical, unusual, or unfamiliar viewpoint, an extra process of mental rotation (or transformation, normalization, or alignment) is needed to compensate for any shape differences between the perceived input and stored views to enable them to be matched to each other (Palmer et al. 1981; Tarr and Pinker 1989; Ullman 1998).

Here, we focus on testing the MVPT variant that implicates the core process underlying the well-studied spatial ability of mental rotation on object cognition (Tarr and Pinker 1989). On mental rotation tasks ( Shepard and Cooper 1982), 2 objects appear at different orientations and are either identical or left–right mirror images of each other. Response times (RTs) and error rates to judge whether the 2 objects are the “same” or “different” increase linearly as the orientation disparity between the 2 objects increases. This linear relation has been taken as evidence that observers imagine one of the objects

Haline E. Schendan1,2,5 and Chantal E. Stern2,3,4

1Department of Psychology, Tufts University, 490 Boston Avenue, Medford, MA 02155, USA, 2Department of Psychology, Boston University, Boston, MA 02215, USA, 3Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Charlestown, MA 02129, USA and 4Center for Memory and Brain, Boston University, Boston, MA 02215, USA
moving through space along the same continuous trajectory as if being physically rotated to match the orientation of the other object (Shepard and Cooper 1982). Many tasks of object memory also show linear changes in performance with increasing rotations away from a target view, suggesting a common process is responsible for the curves in both mental rotation and object memory tasks, and motivating the MVPT account (Jolicoeur 1985; Tarr and Pinker 1989; Tarr 1995). A key finding is that performance on categorization and recognition is monotonically slower and less accurate as objects rotate farther from a canonical view (Palmer et al. 1981; Edelman and Bulthoff 1992; Srinivas 1993, 1995; Cutzu and Edelman 1994; Schendan and Kutas 2003).

MVPT variants are consistent with evidence that the dorsal pathway has a key role in both mental rotation and object cognition. Although more superior intraparietal sulcus (IPS) regions are also involved in saccades, the ventral-caudal IPS (vIPS) and adjacent transverse occipital sulcus (TOS) are not, and show a linear parametric increase as the degree of rotation increases, implicating these object-sensitive dorsal foci (Schendan and Stern 2007) specifically in the scaling of mental rotation task performance with angular disparity (Carpenter et al. 1999; Harris et al. 2000; Podzebenko et al. 2002; Ecker et al. 2006). Occipitoparietal damage impairs mental rotation and object categorization. Patients with the “perceptual categorization” or “unusual views deficit” type of visual apperceptive agnosia (Farah 1990) have right inferior parietal damage, including vIPS and TOS (Warrington 1982; Turnbull and McCarthy 1996; Turnbull 1997; Turnbull, Beschin, et al. 1997). They have problems with impoverished images, in general, and are impaired with unusual, but not canonical, views of objects. Although some neuropsychological work suggests dissociations between mental rotation and object cognition, the patients performed somewhat abnormally on both tasks (Farah and Hammond 1988; Turnbull and McCarthy 1996).

Few tests of visual object cognition theories have used more than one memory task in the same study, and none with neuroimaging. One behavioral study found that view dependence differs across 3 memory tests: categorization, recognition, and working memory (Srinivas 1995). Although another behavioral test of MVPT used multiple tasks and found generalization (Tarr et al. 1998), all were tests of episodic explicit memory for personally experienced events (Schacter and Tulving 1994). Of neuroimaging studies, 2 found occipitoparietal activation for unusual relative to canonical views during categorization (Kosslyn et al. 1994; Sugio et al. 1999). Another study found substantial overlap in vIPS and TOS during object categorization and mental rotation (and not saccades) (Schendan and Stern 2007). Of 2 other studies that also included a mental rotation task, one compared mental rotation of letters and real objects with object naming (Wilson and Farah 2006), whereas the other compared mental rotation with immediate recognition of cube objects rotated from 15° to 90° (Gauthier et al. 2002), and both found some occipitoparietal overlap. Still, some anatomical dissociations were found, suggesting a greater role for ventral posterior regions (e.g., lateral occipital sulcus [LOS]) in view dependence of object cognition and for right rostromedial IPS in mental rotation (Gauthier et al. 2002; Wilson and Farah 2006). However, these superior IPS regions are those implicated in saccades, working memory, and attention (Courtney 2004; Postle 2006; Schendan and Stern 2007), as opposed to the inferior IPS areas considered critical for scaling effects on mental rotation (Podzebenko et al. 2002). Also, interpreting task dissociations in one study (Gauthier et al. 2002) was complicated by stimulus (i.e., different objects for each task), repetition, and working memory factors that differed between tasks, which could explain the direction of some IPS task dissociations and dorsal-ventral differences.

Throughout the present study, participants saw objects from unusual or canonical views. First, each object was categorized as belonging indoors or outdoors. Next, participants discriminated new objects from old objects recognized as seen before during categorization. Object model verification theories predict greater activation for unusual than canonical views in VLPFC and ventral posterior cortex (Ganis et al. 2007). Both model verification and MVPT also predict such view atypicality effects in the vIPS and TOS, with MVPT predicting these regions are involved also in mental rotation. Although view atypicality and image impoverishment effects in these cortical areas have been found on categorization (Kosslyn et al. 1994; Sugio et al. 1999; Ganis et al. 2007), we tested the predictions in regions of interest (ROIs) defined independently as object-sensitive and/or mental rotation- (and not saccade-) related in the same subjects (Schendan and Stern 2007) and assessed whether they held also for recognition.

### Materials and Methods

#### Materials

Ten lists of 35 grayscale objects were matched for name frequency (Kucera and Francis 1967) and category (20 tools [57.1%] and 15 were animals [12.9%], furniture [12.3%], musical instruments [3.1%], vehicles [8.0%], food/plants [5.1%], buildings [2.2%]). Canonical and unusual views of 85 objects were taken from our prior work (Verfaillie and Boutsen 1995; Schendan and Kutas 2003). Views of 265 other objects were chosen based on pilot results for computer models from copyright free sources or real objects on a turntable photographed with a digital camera on a tripod. Views were taken at 30° and 45° increments from 0° (most symmetric), and the depth rotation axis was chosen to maximize view differences; for object database models, top and bottom, but no 45°, views existed (Tarr and Yu 1996). In pilot work, participants performed 2 tasks on all views of each object (N = 24 for digital photographs and object database; N = 7 for other models). First, up to 16 views of each object were shown all at once arranged in a 4 column x 4 row grid (−6° x 4° visual angle) in random order on a 17” cathode ray tube monitor (Samsung, South Korea; 1024 x 768 resolution). Subjects typed the object’s name into the computer. Next, they ranked ordered the goodness of each view from best (rank 1) to worst, as elsewhere (Palmer et al. 1981; Verfaillie and Boutsen 1995; Schendan and Kutas 2003). Second, each view was shown sequentially at the size used in the experiment centered on the monitor. Subjects rated how hard each view was to identify on a 7-point Likert scale (1 = definitely identify easily). Results were used to choose canonical views as those with the lowest mean goodness rank (i.e., best view) and subjects were used to choose canonical views as those with the lowest mean goodness rank (i.e., better) for canonical (M = 2.9, SE = 0.06) than unusual views (M = 10.6, SE = 0.14), $F_{1,264} = 34.28, P < 0.0001$.

#### Design and Procedure

We used a 2 view (canonical, unusual) x 2 task (categorization, recognition) repeated measures, fast event-related design (Fig. 1). Object lists were assigned to conditions in a Latin square design counterbalanced across participants. Conditions were randomly intermixed. Object order avoided semantic associates. Participants performed each task by pressing 1 of 3 keys (right hand) fast and accurately,
receiving instructions and practice shortly before doing each. In runs 1-3, they categorized objects from canonical or unusual views as indoor or outdoor objects, or unidentified. In runs 4-6, participants saw objects from canonical or unusual views, and recognized each object as remembered from the categorization task with high or low confidence or as a new object. Because recognition requires old and new objects, the categorization runs included old objects categorized before fMRI, counterbalanced across conditions; for old objects in both tasks, half were shown in the same view and half in the other view (i.e., canonical then unusual, or vice versa), but view was task-irrelevant, similar to related work on mental rotation and recognition (Gauthier et al. 2002); only new objects from the categorization task were later shown as old objects on recognition. Finally, 3 localizer tasks used standard methods and blocked designs (for details see Schendan and Stern 2007). The first 2 localizers defined object-sensitive areas and regions activated during mental rotation, with task order randomized; note, this design ensured that all subjects were naı¨ve about mental rotation until after the localizer, and 12 the saccade localizer; for details see Schendan and Stern (2007). All but one were right-hand dominant (Oldfield 1971).

Magnetic Resonance Imaging
At 3 T (Siemens Allegra, Munich, Germany), we acquired high-resolution T1-weighted scans (magnetization prepared rapid gradient echo [MP-RAGE]; 192 x 256), a T1-weighted echo-planar image (EPI) image, and T2*-weighted functional blood oxygen level-dependent (BOLD) scans using a gradient-echo EPI pulse sequence (21 anterior-posterior commissure line slices, 5 mm thick, 1 mm skip, time repetition [TR] = 2 s, time echo [TE] = 30 ms; flip angle = 90°, 64 x 64). If time allowed, a second MP-RAGE scan was acquired during practice for the mental rotation localizer. Using SPM99 software (Wellcome Department of Cognitive Neurology), anatomical and functional scans were coregistered using the T1-EPI image. Coregistered anatomical images were then normalized to MNI305 (Montreal Neurological Institute [MNI]) stereotactic space (interpolating to 3-mm3 voxels; neuroanatomical convention). Preprocessing of fMRI data included slice timing correction (for event-related data), motion correction, normalization (using the anatomical parameters), and spatial smoothing (8-mm3 Gaussian kernel). Statistical analyses used the general linear model. High-pass filtering was applied, but global signal scaling was not to avoid spurious deactivations. Each event in the design matrices was modeled using 2 gamma functions (positive, negative) shifted 2 s apart convolved with a canonical hemodynamic response function. Activation was assessed in linear contrasts between unusual and canonical views in each task; new objects were analyzed, because old objects confound view atypicality and repetition factors. Contrast images for each subject were used in second-level analyses treating subjects as a random effect. One-sample t tests assessed contrasts for each task. Paired t tests assessed contrasts between tasks. Extent threshold was 5 voxels. Anatomical structures were labeled using the human brain atlas of Duvernay et al. (1999).

Participants
Data from the categorization task were available from 18 young, neurologically normal people with normal or corrected vision (age $M = 22$ years; education $M = 15$ years; 8 males). Some of these completed other tasks during the allotted 2-h session: 14 completed recognition (age $M = 22$ years; education $M = 15$ years; 7 males), but 4 others did not due to technical problems. The localizer tasks were run last: 16 stayed to complete the object-sensitive localizer, 13 the mental rotation localizer, and 12 the saccade localizer; for details see Schendan and Stern (2007). All but one were right-hand dominant (Oldfield 1971).
Results

Performance
To assess view effects, RT was the main measure (Bülthoff et al. 1995), and the study was designed for high accuracy. Anticipatory RTs under 300 ms and RTs over ±2.5 SD of the mean per subject and condition were not analyzed; RT was not used to exclude trials for fMRI analyses. Categorization and recognition accuracy was high, and RTs and errors for new objects were higher for unusual than canonical views on both tasks, confirming the classic behavioral finding (Palmer et al. 1981; Cutzu and Edelman 1994). A 2 × 2 omnibus repeated measures ANOVA (SPSS 14.0, SPSS Inc., Chicago, IL) showed significant main effects of view (RT: F₁,1₃ = 33.72, P < 0.001; accuracy: F₁,1₃ = 12.26, P = 0.004), and task (RT: F = 36.23, P < 0.001; accuracy: F = 23.1, P < 0.001), that did not interact (RT: F < 0.8, P > 0.4; accuracy: F < 0.2, P > 0.6). Separate ANOVAs for each task showed view effects on both tasks (Fig. 2).

Discrimination of old and new objects was high, as shown by the discrimination index (d’) of signal detection theory with logistic distributions using a corrected matrix (Snodgrass and Corwin 1988), which was 3.69 for canonical, and 3.62 for unusual views, and showed no view effect, F₁,1₃ = 0.06, P = 0.81; correct recognition of old objects (“hits” for old objects from the same view collapsed across confidence) was 89.6% for canonical and 91.5% for unusual, with no view effect, F = 2.31, P = 0.15. Note, for recognition RTs on new objects, data from one participant were unavailable due to a broken key. Localizer results were reported elsewhere (Schendan and Stern 2007).

Functional MRI
Analyses of ROIs Using Small-Volume Correction
Analyses focused on a priori ROIs. Analysis methods and results for the saccade, object-sensitive, and mental rotation localizers, and methods for using them to define the ROIs and volume masks (ROI Analysis Toolbox v. 1.7, SPM99), and the resultant ROIs, have been described elsewhere (Schendan and Stern 2007). For the most focused ROI analysis, the Gaussian-field small-volume correction (SVC) for multiple voxel-wise comparisons was applied in ROI volume masks defined from mental rotation and object-sensitive localizer results, separately, with saccades for fronto and posterior eye fields exclusively masked out (Supplemental Fig. 1). In addition, 2 VLPFC ROIs were based on published coordinates for model verification and memory effects (Kirschhoff et al. 2000; Gainis et al. 2007); anterior VLPFC (aVLPFC; Brodmann area [BA] 45/47/12 ± 37 26 5) and posterior VLPFC (pVLPFC; BA 44/6 ± 48 9 34). For these ROIs, SVC was applied to a sphere of 16-mm radius around the coordinate (Worsley et al. 1996); if results were nonsignificant but some voxels were suprathreshold (uncorrected P value [P₀] < 0.05), the nearest cluster SVC algorithm was applied, and significant clusters in the ROI reported (Friston 1997). To apply SVC, each categorization or recognition SPM was threshold at P₀ < 0.05 to exclude nonsignificant single voxels. SVC results are reported if significant at corrected P value (Pᵢ) ≤ 0.05 based on the lower of random field theory with Bonferroni family-wise error (FWE) or false discovery rate (FDR) correction (Genovese et al. 2002).

Saccade ROIs. Saccade task results (saccade > fixation) were used to create a bilateral volume mask of frontal eye fields and posterior eye fields of the IPS in anterior (aIPS) and posterior (pIPS) parts in order to exclude all saccade-related voxels (P₀ < 0.05) from other task results using SPM99 exclusive masking.

Object-sensitive ROIs. Six ROIs based on object-decision task results (intact > scrambled) were combined into 4 volume masks: 1) of 2 ventral occipitotemporal (VOT) volumes, one mask was in posterior fusiform gyrus (pFG), also known as “LOa” (Malach et al. 2002), “mid-FG” (Denys et al. 2004), or “vTO” (James et al. 2002). 2) Another VOT mask was in anterior collateral sulcus (CoS) (Hasson et al. 2003). 3) For a dorsal occipitotemporal (DOT) area mask, one part of the ROI encompassed occipitotemporal and inferotemporal sulci (ITS), also known as “LO” (Malach et al. 2002) or “post-ITG” (Denys et al. 2004), and another part of the ROI was around LOS in the most posterior part of the lateral occipital complex (Hasson et al. 2003), with the more ventral part perhaps corresponding to retinotopic area L01 (Larsson and Heeger 2006). 4) For a dorsal foci mask (Grill-Spector et al. 2000), one ROI was around the ‘‘TOS” (TOs = “DF1”) in a building-selective area (Hasson et al. 2003; Epstein et al. 2005), and another ROI was in the most vlIPS region (vclIPS = “DF2”) (Grill-Spector et al. 2000), also known as ventral IPS, which is also motion-sensitive (Denys et al. 2004). (Other dorsal shape-sensitive areas, Denys et al. 2004, located superior to vlIPS were almost entirely masked out of our object-sensitive localizer results by the saccade localizer. Supplemental Fig. 2 shows these areas.)

Mental rotation ROIs. Mental rotation task results (rotation > no rotation control) were used to create 4 ROI masks. 1) The main ROI volume covered occipitoparietal areas that were not saccade-related but overlapped object-sensitive TOS and vlIPS (Schendan and Stern 2007). 2) An ROI volume for a caudal IPS region (cIPS) just superior to vlIPS corresponds to retinotopic area IPS2 (Silver et al. 2005) or shape-selective parieto-occipital IPS (POIPS) (Denys et al. 2004) and was tested for completeness, and showed effects similar to vlIPS, but we do not focus on

Figure 2. Performance results. View atypicality effects on performance with novel real objects were found on both tasks and were comparable between tasks. (A) On both tasks, mean RTs were significantly slower for unusual than canonical views. The view effect was significant for both categorization, F₁,1₃ = 20.116, and recognition, F₁,1₃ = 10.93. (B) On both tasks, mean error rates were higher for unusual than canonical views. The view effect was significant for categorization, F₁,1₃ = 19.536, and showed a trend for recognition, F₁,1₃ = 4.07; for recognition, errors were false alarms (i.e., responding “old” to the new objects). **P < 0.0001, *P = 0.006, +P = 0.065.
cIPS effects because saccade-related activity was found there (Schendan and Stern 2007); note, more rostrodorsal parietal regions have not shown parametric modulation with mental rotation and so were not of interest, but we observed these areas to show no view effects. 3) One ROI volume covered parts of object-sensitive left ITS and left LOS. This more dorsal LOS region may correspond to retinotopic area LO2, which lies just posterior to motion area MT+ (Larsson and Heeger 2006). 4) Another ROI in the right hemisphere covered an area centered at the junction where the lingual sulcus (LS) intersects the object-sensitive CoS area between fusiform and posterior parahippocampal gyri (x 30, y –42, z –9. Fig. 5C), sometimes referred to as a CoS “building,” lingual “landmark” (Aguirre et al. 1998; Hasson et al. 2003), or parahippocampal “place” area (Epstein et al. 1999); we use the anatomical label of CoS/LS.

Analyses of ROIs Using SPM Overlays

Another ROI approach was used because the volume masks for SVC analyses are large and may include structures outside the cortex of interest. To examine 4 bilateral ROIs of VOT (pFG, CoS, CoS/LS, DOT (ITS, LOS), dorsal foci (vcIPS, TOs), and vLPFC, the contrast SPMs of categorization and recognition results were superimposed on the mental rotation and object-sensitive localizer SPM contrasts at P_{z} < 0.05, excluding saccade voxels; only left pVLPFC was active in the localizers, but the corresponding right pVLPFC was assessed because picture novelty effects are bilateral (Kirchhoff et al. 2000). For categorization, the threshold for the SPMs was P_{u} < 0.005, and significant voxels were defined as those located in a ROI. For recognition and task comparisons, which involved fewer subjects, the threshold for the SPMs was P_{u} < 0.01 (0.05 ÷ 4 ROIs).

View Atypicality (Unusual > Canonical) Results

Data from all trials and correct trials were analyzed, separately, to facilitate interpretation of task effects; because the 2 tasks depend upon distinct memory processes, different object trials would be correct for each task so results overall are considered more valid (Tulving 1972; Curran et al. 2002). To analyze objects eliciting correct responses, the SPH9 design matrix further sorted trials for each condition according to subject response: for categorization, correctly categorized new objects and effects of no interest (i.e., correctly categorized old objects, unidentified, or no responses); for recognition, correct rejections (new objects correctly reported as new), and effects of no interest (i.e., hits [old objects correctly remembered], false alarms [new incorrectly reported as remembered], misses [old incorrectly reported as new, and no responses]). Participants were included in analyses of correct recognition trials if correct rejection rates were over 50%, which excluded 1 person (n = 13). Too few incorrect trials were available to assess these. Because most trials were correct, results should be similar for both analyses, as was found.

Categorization. Overlying results from all categorization trials on the localizers showed that activation was greater for unusual than canonical views of new instances of common objects (Fig. 3B,C,E) in both mental rotation ROIs and object-sensitive ROIs of bilateral vlIPS (and right cIPS), TOs, ITS, and LOS. Activation was also found in ROIs that were only object sensitive in right TOs caudomedially and a mid-ITS region bilaterally (in-between mental rotation parts of ITS and LOS), extending ventromedially a little into pFG, and in the mental rotation defined ROIs of pVLPFC and aVLPFC bilaterally (Fig. 4; Supplemental Fig. 3); note, the aVLPFC region is located in the caudal part of BA 47/12 in the horizontal ramus of the lateral fissure, which is just anterior to insular cortex (Kostopoulos et al. 2007). Overlaying results for correct trials on the localizers showed a similar regional pattern (Supplemental Fig. 4B,C,E). SVC results for all and correct trials supported this pattern, except an effect in the mental rotation ROI of right CoS/LS was revealed (Table 1). Assessing results also without the saccade mask uncovered view effects for all and correct trials in saccade areas of right aIPS, a small part of left aIPS, and right dorsal shape-selective areas of POIPS and medial dorsal IPS (Denys et al. 2004).

Recognition. Overlying results for all recognition trials on the localizers showed greater activation for unusual than canonical views (Fig. 3B,D,E) in both mental rotation ROIs and object-sensitive ROIs of bilateral vlIPS (and right cIPS), right TOs, and bilateral ITS. Activation was found in ROIs that were only object sensitive in right TOs anterolaterally, and the mental rotation ROI of right CoS/LS, right pVLPFC, and a small part of left pVLPFC (z 23; inferior precen sulcus; Supplemental Fig. 3). SVC results further supported the regional pattern (Table 2). For correct trials, a similar regional pattern was found, except for no effect in right CoS/LS (Supplemental Fig. 4B,D,E). SVC results were significant in right pVLPFC (nearest cluster Z = 3.06, P_{u} = 0.038; x 60, y 18, z 33). Assessing results without the saccade mask, revealed the same additional regions over all trials as found for categorization but, for correct trials, only a few voxels near right pIPS.

Task dependence: categorization versus recognition. For categorization relative to recognition, view atypicality effects for all trials were greater in the mental rotation ROI and object-sensitive ROI of left LOS and the corresponding object-sensitive ROI of right LOS, a few voxels bordering the object-sensitive ROI of right ITS, and the mental rotation ROIs of bilateral aVLPFC (Fig. 3A). SVC results confirmed significant effects in the mental rotation ROI of left LOS (Z = 3.11, P_{u} = 0.043; x –30, y –90, z –6), the object-sensitive ROI of right LOS (Z = 3.77, P_{u} = 0.031; x 30, y –87, z –6), and the nearest cluster in left vlIPS (Z = 2.82, P_{u} = 0.049; x –33, y 30, z –6). For recognition relative to categorization, task differences in view atypicality effects across all trials were found in the mental rotation ROI of right CoS/LS (Fig. 5A, Supplemental Fig. 6A), which was also significant in SVC results (Z = 2.42, P_{u} = 0.027; x 30, y –45, z –9). For correct trials, we were concerned that, because subject’s criteria for making categorization versus recognition decisions must differ, trials sorted by response would have uncontrolled stimulus differences between tasks (i.e., unmatched subsets of object trials for each task), which may reduce task effects by increasing noise or induce spurious stimulus-induced effects. These concerns did not apply to results over all trials because stimuli were counterbalanced between tasks. Results over all trials are thus the primary basis for our conclusions (for a similar approach, see Curran et al. 2002), and used to further constrain analyses of correct trials. For view atypicality effects on categorization relative to recognition, results over all trials showed task differences in 2 mental rotation ROIs (LOS, vlIPS), so we assessed whether, for correct trials, the overlay results would show task effects there at P_{u} < 0.025 (0.05 ÷ 2 ROIs), which were found (Supplemental Fig. 4A); SVC results showed nearest clusters.
were significant in left aVLPFC (Z = 2.43, P = 0.05) and right aVLPFC (Z = 1.98, P = 0.049). In contrast, for view atypicality effects on recognition relative to categorization, results over all trials suggested task differences in the ROI of right CoS/LS, so we assessed whether, for correct trials, the overlay results would show a task effect there at P < 0.05, which was not found (Supplemental Fig. 5D); SVC results showed no effect. In sum, view atypicality effects in LOS and aVLPFC were generally larger on categorization than recognition. Also, view atypicality effects in right CoS/LS were larger on recognition than categorization over all trials but indistinguishable between tasks on correct trials.

Figure 4. aVLPFC showed view atypicality effects (unusual > canonical) for all new objects on the categorization task. The aVLPFC region is located in the caudal part of BA 47/12 in the horizontal ramus of the lateral fissure, which is just anterior to insular cortex (Kostopoulos et al. 2007). The view atypicality effect on categorization was located bilaterally in the same aVLPFC ROI as found during mental rotation of objects (Schendan and Stern 2007). Note: Activation in SPM is shown with saccade voxels masked out exclusively and superimposed on an individual canonical brain (MN1, SPM99) in coronal slices at the location of the maximum activated voxel for aVLPFC in left (−30 24 −3) and right (39 30 −3) hemispheres. MNI y coordinate labeled in center. The ROIs, which were defined based on the mental rotation and object-sensitive localizer results (corrected PFDR < 0.05 with saccade voxels at P < 0.05 masked out exclusively), are superimposed on each slice in green and blue outlines, respectively. HR = horizontal ramus of lateral fissure; AR = ascending ramus of lateral fissure; IFS = inferior frontal sulcus; SFS = superior frontal sulcus.

Additional analyses of CoS/LS: The anatomical landmarks, coordinates, shape, and spatial extent of the right CoS/LS activations fit those of a "building," "landmark," or "place" area (Aguirre et al. 1998; Hasson et al. 2003; Epstein et al. 2005). To further explore this important region for vision and memory theories, first, we compared the categorization, recognition, and between task results for view atypicality effects with results from our prior studies using the same threshold of P < 0.01. We found activation bilaterally in the same CoS/LS region not only during mental rotation and recognition, and more so for recognition than categorization, but also during the delay period of a working memory task (Fig. 5A–D) (Schon et al. 2004). In contrast, right CoS/LS showed minimal effects for our indoor/outdoor categorization task, another categorization task, that is, for object-sensitivity during an object-decision task from our localizer data, and processes related to long-term memory encoding (Schon et al. 2004) (Fig. 5E–G). Overall, the CoS/LS showed its known right laterality (Epstein et al. 2007). Even so, left CoS/LS also showed effects of mental rotation and working memory delay (Fig. 5C,D) (Schon et al. 2004; Schendan and Stern 2007), and marginal view atypicality effects during recognition and a marginal task difference in view atypicality effects on recognition relative to categorization (Supplemental Fig. 6A,B). Second, we asked whether view atypicality effects in right CoS/LS were merely below the SPM threshold over all categorization trials and for correct trials on both tasks by assessing results in the single ROI of right CoS/LS at P < 0.05, and found significant effects. For all trials, effects were significant on categorization in an anterovenal part of the right CoS/LS region (Supplemental Fig. 6D), which had also
shown effects on recognition (Fig. 5B, Supplemental Fig. 6B), and did not show a larger view atypicality effect on recognition than categorization, in contrast to a posterior part of right CoS/LS, which did show this task effect (Fig. 5A; Supplemental Fig. 6A). For correct trials, effects were also significant in the right CoS/LS on both categorization and recognition, but this ROI did not show a task effect (Supplemental Fig. 5). In sum, right CoS/LS showed view atypicality effects for all stimulus–response combinations, except incorrect categorizations; over all trials, view effects for correct rejections plus false alarms on recognition were greater than view effects for correct plus incorrect categorizations, and correct trial results showed comparable view effects for correct rejections on recognition and correct categorizations; thus, the task effect over all trials reflected smaller view effects for incorrect categorizations than false alarms on recognition.

### Table 1

<table>
<thead>
<tr>
<th>Region</th>
<th>Hemisphere</th>
<th>Z</th>
<th>P</th>
<th>MNI (mm)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>y</td>
<td>z</td>
<td></td>
</tr>
<tr>
<td>Object sensitive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vcIPS Right</td>
<td>3.68</td>
<td>0.0235</td>
<td>−33</td>
<td>−84</td>
<td>27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right</td>
<td>3.37</td>
<td>0.0033</td>
<td>−33</td>
<td>−72</td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOS Right</td>
<td>3.20</td>
<td>0.0308</td>
<td>−33</td>
<td>−81</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>4.72</td>
<td>0.0005</td>
<td>33</td>
<td>−81</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOS Right</td>
<td>3.30</td>
<td>0.0104</td>
<td>−36</td>
<td>−87</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>3.29</td>
<td>0.0090</td>
<td>33</td>
<td>−90</td>
<td>−6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITS Right</td>
<td>3.99</td>
<td>0.0104</td>
<td>−36</td>
<td>−69</td>
<td>−15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>3.81</td>
<td>0.0090</td>
<td>30</td>
<td>−81</td>
<td>−12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pFG Right</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>2.55</td>
<td>0.0377</td>
<td>36</td>
<td>−39</td>
<td>−21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental rotation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dIPS Right</td>
<td>3.17</td>
<td>0.0410</td>
<td>33</td>
<td>−63</td>
<td>48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>3.74</td>
<td>0.0160</td>
<td>−36</td>
<td>−84</td>
<td>24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOS Right</td>
<td>3.07</td>
<td>0.0060</td>
<td>33</td>
<td>−72</td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>3.29</td>
<td>0.0260</td>
<td>−33</td>
<td>−81</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOS Right</td>
<td>4.27</td>
<td>0.0020</td>
<td>36</td>
<td>−84</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>4.27</td>
<td>0.0020</td>
<td>36</td>
<td>−84</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITS Right</td>
<td>3.66</td>
<td>0.0040</td>
<td>−42</td>
<td>−63</td>
<td>−12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>2.55</td>
<td>0.0330</td>
<td>33</td>
<td>−39</td>
<td>−9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Object model verification</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aVLPC Right</td>
<td>4.10</td>
<td>0.0110</td>
<td>33</td>
<td>−27</td>
<td>−3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>3.80</td>
<td>0.0100</td>
<td>36</td>
<td>30</td>
<td>−3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pVLPC Right</td>
<td>2.78</td>
<td>0.0499</td>
<td>48</td>
<td>3</td>
<td>33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>3.51</td>
<td>0.0280</td>
<td>54</td>
<td>15</td>
<td>30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Reported P values reflect either a Bonferroni FWE or FDR correction for multiple voxel-wise comparisons (marginal effects in parentheses); #alcs are statistics from nearest cluster SVC in ventrolateral prefrontal cortex (VLPFC); aVLPC = anterior VLPC at +/−37 26 5 (BA 45/47/12); pVLPC = posterior VLPC at +/− 48 9 34 (BA 44/6); brain structures labeled based on Duvernoy et al. (1999).

### Table 2

<table>
<thead>
<tr>
<th>Region</th>
<th>Hemisphere</th>
<th>Z</th>
<th>P</th>
<th>MNI (mm)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>y</td>
<td>z</td>
<td></td>
</tr>
<tr>
<td>Object sensitive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vcIPS Right</td>
<td>3.64</td>
<td>0.0140</td>
<td>39</td>
<td>−81</td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOS Right</td>
<td>3.65</td>
<td>0.0140</td>
<td>42</td>
<td>−84</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental rotation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vcIPS Right</td>
<td>3.64</td>
<td>0.0110</td>
<td>39</td>
<td>−81</td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOS Right</td>
<td>3.65</td>
<td>0.0110</td>
<td>42</td>
<td>−84</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOT-ITS Right</td>
<td>2.80</td>
<td>0.0450</td>
<td>48</td>
<td>−90</td>
<td>−6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>2.69</td>
<td>0.0140</td>
<td>30</td>
<td>−45</td>
<td>−9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Same legend as in Table 1.

### View Canonicality (Canonical > Unusual) Results

Model verification theory makes no predictions for this contrast, but recent findings suggest that the "active task" network showing view atypicality effects is anticorrelated with activity in a "default resting state" network (e.g., Fox et al. 2005; Vincent et al. 2006) implicated in memory and mental imagery (Schacter et al. 2007), and we found that a no rotation control condition relative to mental rotation activates the default network (Schendan and Stern 2007). An MVPT variant of model verification theory is compatible with the idea that processing canonical views is similar to that in a no rotation condition, predicting canonical more than unusual views will activate the same default state network. To test this, in addition to the active task ROI analyses, which were not expected to show effects, we assessed a default state network by overlaying the categorization and recognition results at $P_{c} < 0.01 = 0.05 \div 5$ default network regions of medial posterior [i.e., retrosplenium, precuneus], posterior cingulate, medial prefrontal [including anterior cingulate], left lateral temporal, and left lateral parietal cortex) on the default network defined using the reverse contrast of mental rotation (no rotation control > rotation, $P_{c} < 0.05$).

As expected, analyses on the active task (mental rotation and object sensitive) ROIs showed no view effects for all and correct trials on categorization and recognition. In contrast, all regions in the default network ROIs were activated for canonical relative to unusual views over all categorization trials; precuneus, retrosplenial, posterior cingulate, anterior cingulate, medial superior frontal gyrus, and left lateral temporal and inferior parietal cortex (Fig. 6A, Table 3). Results for correct trials also showed these regions (Supplemental Fig. 7A; Table 3). For recognition, results revealed activation in left lateral parietal.
cortex over all \((x = -42, y = -45, z = 21; T = 3.98; \text{Fig. 6B})\) and on correct trials \((x = -57, y = -48, z = 18; T = 2.98; \text{Supplemental Fig. 7B})\).

**Task dependence: categorization versus recognition.** The contrast of canonical > unusual for categorization relative to recognition is mathematically equivalent to unusual > canonical for recognition relative to categorization (for the latter results in mental rotation and object-sensitive ROIs, see section above), and reveals both sets of task effects altogether. For all trials, assessing 5 default network ROIs, view Canonicality effects were larger for categorization than recognition in precuneus, retrosplenial, medial prefrontal, and left lateral temporal cortex (Fig. 6C); a task effect in left lateral parietal cortex was located posterior to the ROI. For correct trials, task effects were found in default state ROIs of precuneus, retrosplenial, and medial prefrontal cortex (Supplemental Fig. 7C). The contrast of canonical > unusual for recognition relative to categorization is mathematically equivalent to unusual > canonical for categorization relative to recognition (for the latter results of mental rotation and object-sensitive ROIs, see section above), and reveals both sets of task effects altogether. Assessing 5 default network ROIs, no task effect was found for all and correct trials. In sum, together with results for each task separately, the default network showed view canonicality effects, but more so for categorization than recognition, except in the lateral parietal region, which showed comparable effects on both tasks.

**Discussion**

Brain activity related to categorization, recognition, mental rotation, object-sensitivity, and saccades were compared in order to examine the role of interactions between prefrontal-parietal and ventral posterior cortex in visual object cognition and memory. The fMRI results support key predictions of object model verification theory and MVPT. We found that the vclIPS and TOS are dorsal visual regions that unusual views of real objects recruit more strongly than canonical views during "nonspatial tasks" of object categorization and recognition. These areas are also object sensitive and involved in the cardinal spatial task of mental rotation (and not saccades) (Schendan and Stern 2007). (The same effects were found in a more anterodorsal region of cIPS, but most of this overlaps aIPS and pIPS saccade areas, Schendan and Stern 2007, so it will not be discussed in detail.) Although view atypicality effects in cortex (Supplemental Fig. 1). (A) View atypicality effects in CoS/LS are greater bilaterally during recognition than categorization over all trials (recognition > categorization for contrast of unusual > canonical). (B) View atypicality effects in CoS/LS bilaterally over all recognition trials (unusual > canonical). (C) Mental rotation activation (rotation > control) in CoS/LS bilaterally using data from Schendan and Stern (2007). (D) Delay activity during a delayed matching to sample type of working memory task with scenes was found in CoS/LS bilaterally and a region including mid-fusiform and adjacent parahippocampal gyr (delay period on working memory > control tasks; SPM2). Data were from Schon et al. (2004). (F) View atypicality effects during all categorization trials (unusual > canonical) were minimal in CoS/LS but significant in other areas, such as ITS. (F) Object-sensitive activation (intact > scrambled) was minimal or absent in CoS/LS but was high in other areas, including surrounding fusiform and parahippocampal gyr, using data from Schendan and Stern (2007); axial slice shown at \(z = -6\) for clarity. (G) Subsequent memory modulation during the delay period of a delayed matching to sample task with scenes (recognition confidence of 5 (maximum) \(> 1\); SPM2) was minimal in CoS/LS but significant in surrounding mid-FG and parahippocampal gyr. Data were from Schon et al. (2004).
prefrontal–posterior networks have been reported for categorization (Kosslyn et al. 1994; Sugio et al. 1999), we find that activity in this same network, involving the pVLPFC (BA 44/6), and object-sensitive areas in vcIPS, TOS, and ITS, shows view sensitivity not only for mental rotation and categorization but also for recognition, comprising a task-general prefrontal–posterior network (Schendan and Stern 2007). A right CoS/LS “place” area is also view-sensitive on diverse tasks, showing view sensitivity on categorization, recognition, and mental rotation (Schendan and Stern 2007), but less for uncategorized objects, and is activated during the delay period of a matching to sample task, regardless of subsequent memory confidence (Schon et al. 2004). In contrast, the aVLPFC (caudal part of BA 47/12) and part of object-sensitive LOS is view-sensitive on mental rotation (Schendan and Stern 2007) and more so on categorization than recognition, comprising a task-dependent prefrontal–posterior network. In addition, the reverse contrast shows view canonicality effects in a “default” state network, which has been shown to be anticorrelated with the “active” task networks of lateral prefrontal–IPS regions (Fox et al. 2005; Vincent et al. 2006). Canonical relative to unusual views activates medial parts of prefrontal and posterior cortex, and lateral temporal cortex on categorization more than

Table 3
Default State Network shows view canonicality effects (canonical > unusual) during categorization: Z scores and coordinates of most significantly activated volume cluster (>8 mm apart) within each anatomical structure at uncorrected P < 0.001.

<table>
<thead>
<tr>
<th>Region</th>
<th>All trials MNI (mm)</th>
<th>Correct trials MNI (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Z  x  y  z</td>
<td>Z  x  y  z</td>
</tr>
<tr>
<td>Retrosplenial cortex</td>
<td>4.32  0  -51  33</td>
<td>4.05  0  -51  33</td>
</tr>
<tr>
<td>Posterior cingulate</td>
<td>3.94  -3  -18  57</td>
<td>3.47  12  -24  42</td>
</tr>
<tr>
<td>Precuneus</td>
<td>3.99  9  -18  51</td>
<td>3.36  9  -12  36</td>
</tr>
<tr>
<td>Cuneus</td>
<td>4.04  0  -63  39</td>
<td>4.65  0  -78  36</td>
</tr>
<tr>
<td>STS/lateral parietal</td>
<td>4.18  -48  -48  21</td>
<td>3.70  -45  -48  21</td>
</tr>
<tr>
<td>Middle temporal gyrus</td>
<td>3.73  -60  -54  15</td>
<td>3.70  -45  -48  21</td>
</tr>
<tr>
<td>Medial prefrontal</td>
<td>3.63  -60  -57  21</td>
<td>4.05  60  -57  21</td>
</tr>
</tbody>
</table>

Note: STS = superior temporal sulcus (ascending or horizontal posterior segment), which corresponds to the lateral inferior parietal part of the network around the angular gyrus. Same conventions as in other tables.
recognition, and lateral inferior parietal parts on both tasks, and these regions lie within the same "default" network activated for no rotation (control) relative to a mental rotation task (Schendan and Stern 2007), and for small than large rotations on object cognition tasks (Gauthier et al. 2002). We aimed to test object model verification theory and MVPT, which makes the strongest predictions about view atypicality effects in active task VLPFC-posterior networks and so, next, we introduce an emerging account of the brain basis of visual object knowledge, and then detail the implications of the active task network findings for this account and the role of each network component in achieving object constancy.

Two State Interactive Account of Visual Object Knowledge

We propose a Two-State Interactive (2SI) account of the cortical dynamics of visual object knowledge activation (Schendan and Kutas 2007) that integrates model verification and MVPT accounts with each other and with memory theory (Fig. 7). The 2SI account combines timing ideas from event-related potential (ERP) studies on object constancy (Schendan and Kutas 2002, 2003, 2007) and emerging ideas that complex cognition, attention, feature binding, and fine discriminations depend upon later recurrent, feedback, and interareal contributions to visual information processing (Ullman 1995; Grossberg 1999; Lamme and Roelfsema 2000; David et al. 2005; Treisman 2006). The central idea of the 2SI account is that ventral object-sensitive cortex is activated in 2 functionally distinct states. First, a fast feedforward pass through ventral object-sensitive cortex takes place during an ‘‘initial classification’’ or structural encoding state (Bruce and Young 1986) between 125 and 200 ms, sufficient for perceptual categorization and global discriminations (e.g., face or not) (Schendan et al. 1998). We propose that more precise categorization and memory discriminations (i.e., determining a specific object class) are related to later neural computations supporting the critical matching processes of Object Model Selection (Schendan and Kutas 2002, 2003, 2007): between 200 and 500 ms, a second more sustained state of iterative activation of ventral object-sensitive cortex involves both feedforward and feedback interactions among these areas and with VLPFC and/or the mediotemporal lobe, with engagement of these other regions varying with memory task requirements. The purpose of the second interactive state is to disambiguate the input by integrating bottom-up and top-down information. At this later time, Object Model Verification processes in prefrontal–parietal networks modulate ventral posterior cortex (Ganis et al. 2007), facilitating Object Model Selection, to achieve object constancy despite impoverished viewing conditions. Accordingly, ERPs differ between unusual and canonical views after 200 ms (Schendan and Kutas 2003), and computational models suggest bottom-up processing dominates ERPs before 200 ms, whereas feedback connections along the hierarchical visual pathways dominate later ERPs between 200 and 500 ms (David et al. 2005). Our emphasis on later processes differs from other top-down accounts proposing earlier rapid feedback modulation in which extrastriate cortex projects via the magnocellular pathway into VLPFC and into orbitofrontal cortex within 130 ms, which then uses information about low spatial frequency candidate object models to send top-down feedback into ventral object-sensitive cortex by as early as 180 ms (Bar 2003; Bar et al. 2006).

Two VLPFC–Posterior Networks for Visual Object Constancy

The aVLPFC (BA 45/47/12) has been proposed to control network interactions during model verification in order to determine the object’s category, modulating more ventral posterior areas from which visual knowledge is retrieved (Ganis et al. 2007). We propose that the aVLPFC–posterior network is primarily important for 2 prediction processes of model verification: 1) object model selection to evaluate whether the visual percept matches candidate models of known object categories in memory and select the best match (Schendan and Kutas 2002, 2003, 2007); and 2) retrieval and working memory for salient visual features of the selected candidate model. Accordingly, the caudal part of BA 47/12 in aVLPFC, where we found view atypicality effects, has been implicated in actively selecting and comparing information in memory to disambiguate the interpretation of perceived inputs (Kostopoulos et al. 2007). The BA 45 part of left aVLPFC has been proposed to select from among competing alternatives stored in left ITS (Kan and Thompson-Schill 2004; Kan et al. 2006). Our aVLPFC effects are bilateral, consistent with prior studies of categorization and recognition (James et al. 2002) and view-specific knowledge in right posterior cortex (Burgund and Marsolek 2000). Our findings suggest an aVLPFC (BA 47/12)–LOS network for model prediction is more important for categorization than recognition. After all, categorization entails a many-instance-to-one-class mapping, enabling a pool of previously learned, instances of a category (i.e., object models) to be available for the verification processes of prediction that can then be tested for their similarity with the visual input (Rosch et al. 1976; Sigala et al. 2002).

The pVLPFC–posterior network primarily contributes to the testing component of model verification for both categorization and recognition. Testing involves top-down control processes of 3) covert attention shifting and 4) attribute biasing to search for and switch attention between distinctive attributes in the input that are predicted by the candidate model in memory (Ganis et al. 2007), and 5) working memory for the relative spatial relations among parts of an object. The pVLPFC corresponds to the frontal part (ventral premotor area in precentral sulcus at the junction of BA 44/6; a human homolog to monkey F5) of a “canonical neuron” circuit preferentially recruited for perceiving objects, especially manipulable ones, executing a grasp toward an object, and motor knowledge retrieval (Grezes et al. 2003; Kan et al. 2006). Our findings suggest that, across tasks, the occipitoparietal region (vIPS, TOS) is most consistently coactivated with right pVLPFC (and ITS) for visual constancy with impoverished images of objects. This is consistent with evidence that the pVLPFC is connected reciprocally with the aIPS and pIPS saccade areas, and vIPS is an earlier process in an IPS cascade from inferior to superior parts for actions with objects (Chao and Martin 2000; Grezes et al. 2003; Kan et al. 2006), a role for pVLPFC, vIPS, and TOS (and cIPS) in modulating ventral posterior activity during object cognition (Wojciulik and Kanwisher 1999; Courtney 2004; Silver et al. 2005; Noppeney et al. 2006; Postle 2006; Xu and Chun 2006), and the bias competition model of selective attention in which VLPFC and ventral temporal cortex interact to generate object and feature templates (Desimone and Duncan 1995). In contrast to a common finding of left lateralization (Grezes et al. 2003;
We propose that the role of vcIPS and TOS is to compute the relative spatial relations between features and parts of an object, and, if necessary, transform the visuospatial structure to facilitate matching the perceived input with object models in memory, another perceived object, or a grasping action. Visual input has a multilevel hierarchical space. Spatial arrangements either of local parts define an object or of multiple objects define a scene at more global hierarchical levels (Cave and Kosslyn 1993; Robertson 1996; Newell et al. 2005). We suggest occipitoparietal cortex will be recruited as a function of stimulus and task requirements for additional analysis and/or retrieval of visuospatial structure, as opposed to abstract, semantic, or lexical knowledge, for which pVLPFC and ITS interactions are enough (Kan and Thompson-Schill 2004; Kan et al. 2006). The vcIPS and/or TOS have been implicated in perceiving and matching object structure in depth, stereopsis, symmetry, spatial cognition and coordinate transformation, grasping objects, and tasks demanding greater spatial analysis of object structure (Milner and Goodale 1993; Kanwisher and Wojciulik 2000; Backus et al. 2001; Faillenot et al. 2001; Beauchamp et al. 2002; Giliaie-Dotan et al. 2001; Assad 2003; Shikata et al. 2003; Tsao et al. 2003; Altmann et al. 2005; Arbib 2006). Schendan and Stern 2007), view atypicality effects in our pVLPFC--posterior network are more extensive in the right hemisphere. During model verification, the pVLPFC–vcIPS–TOS network modulates activity in ventral posterior cortex (LOS, ITS, CoS/LS) as a function of image impoverishment (Ganis et al. 2007), including view atypicality. The pVLPFC may interact with aVLPFC in a prediction-testing loop. Both are interconnected bidirectionally with each other and with the ventral stream (Ungerleider et al. 1989; Petrides and Pandya 2002; Petrides 2005), and have been implicated in visual working memory, in which parietal cortex is typically also involved (Passingham et al. 2000; Stern et al. 2000; Postle 2006). Considering our performance results, as well, the pVLPFC–posterior network may have a relatively greater role in performance speed, whereas the aVLPFC has a greater role in accuracy, though more for categorization than recognition.

**Dorsal Visual Pathway for Object Constancy**

We propose that the role of vcIPS and TOS is to compute the relative spatial relations between features and parts of an object, and, if necessary, transform the visuospatial structure to facilitate matching the perceived input with object models in memory, another perceived object, or a grasping action. Visual input has a multilevel hierarchical space. Spatial arrangements either of local parts define an object or of multiple objects define a scene at more global hierarchical levels (Cave and Kosslyn 1993; Robertson 1996; Newell et al. 2005). We suggest occipitoparietal cortex will be recruited as a function of stimulus and task requirements for additional analysis and/or retrieval of visuospatial structure, as opposed to abstract, semantic, or lexical knowledge, for which pVLPFC and ITS interactions are enough (Kan and Thompson-Schill 2004; Kan et al. 2006). The vcIPS and/or TOS have been implicated in perceiving and matching object structure in depth, stereopsis, symmetry, spatial cognition and coordinate transformation, grasping objects, and tasks demanding greater spatial analysis of object structure (Milner and Goodale 1993; Kanwisher and Wojciulik 2000; Backus et al. 2001; Faillenot et al. 2001; Beauchamp et al. 2002; Giliaie-Dotan et al. 2001; Assad 2003; Shikata et al. 2003; Tsao et al. 2003; Altmann et al. 2005; Arbib 2006). Schendan and Stern 2007), view atypicality effects in our pVLPFC–posterior network are more extensive in the right hemisphere. During model verification, the pVLPFC–vcIPS–TOS network modulates activity in ventral posterior cortex (LOS, ITS, CoS/LS) as a function of image impoverishment (Ganis et al. 2007), including view atypicality. The pVLPFC may interact with aVLPFC in a prediction-testing loop. Both are interconnected bidirectionally with each other and with the ventral stream (Ungerleider et al. 1989; Petrides and Pandya 2002; Petrides 2005), and have been implicated in visual working memory, in which parietal cortex is typically also involved (Passingham et al. 2000; Stern et al. 2000; Postle 2006). Considering our performance results, as well, the pVLPFC–posterior network may have a relatively greater role in performance speed, whereas the aVLPFC has a greater role in accuracy, though more for categorization than recognition.
The underlying function is not merely spatial but also object sensitive (Malach et al. 1995; Grill-Spector et al. 1998, 1999, 2001; Schendan and Stern 2007), and these dorsal regions are modulated routinely during object perception, categorization, and memory tasks in a way not directly attributable to saccade or attention processes (Grill-Spector et al. 2000, 2001; James et al. 2002; Vuilleumier et al. 2002; Hasson et al. 2003; Denys et al. 2004; Altmann et al. 2005; Ganis et al. 2007; Schendan and Stern 2007). Bilateral TOS has been implicated in Balint’s syndrome patients who perceive only one object or feature at a time without knowing where it is located (Friedman-Hill et al. 1995, 2003; Turnbull, Carey, et al. 1997). The occipitoparietal region has been implicated in the explicit dynamic binding into a coherent structural representation of the hierarchical spatial configuration of perceptual features, which can be bound together with information about object identity into a “temporary token (object file)” that enables appropriate actions and cognitive decisions about objects (Treisman 2006; Treisman and Kanwisher 1998). Similarly, monkey posterior parietal area 7A is involved in constructing the spatial configuration among the parts of objects (Chafee et al. 2005) and automatically provides a map of visual salience required to allocate attention (Constantinidis and Steinmetz 2005).

Mental rotation may not be a direct mental analog of physical rotation, as originally suggested (Shepard and Cooper 1982). Mental rotation performance curves vary with object complexity and dimensionality (two [2D] vs. three [3D]) (Jolicoeur and Landau 1984; Folk and Luce 1987; Bauer and Jolicoeur 1996), and 3D rotations that alter visuospatial structure widely yield more variable rotation curves and view canonicality effects than 2D rotations of known objects (Edelman and Bulthoff 1992; Cutzu and Edelman 1994; Blanz et al. 1999; Gauthier et al. 2002). Neuroimaging has shown violations from a simple linear relationship on cortical activity during object mental rotation and recognition (Gauthier et al. 2002), motivating multiple views accounts to dismiss mental rotation or any other process beyond the ventral stream as contributing to object constancy. Instead, generalization to novel or unusual views is achieved via time-consuming, evidence accumulation, or interpolation across multiple views within a neural population entirely in the ventral stream (Edelman and Bulthoff 1992; Bulthoff et al. 1995; Ullman 1996; Perrett et al. 1998; Tarr and Bulthoff 1998; Peissig and Tarr 2006). Our findings in VLPFC and occipitoparietal cortex, however, demonstrate view-sensitive neural activity outside the ventral stream, indicating additional processes do underlie scaling of performance on categorization (Kosslyn et al. 1994; Sugio et al. 1999; Ganis et al. 2007) and recognition with increasing rotations away from target, familiar, or canonical view(s), or greater image impoverishment.

Ventral Visual Pathway for Categorization, Memory, and View Representation

View atypicality effects reflect perceptual and memory factors (Palmer et al. 1981; Edelman and Bulthoff 1992; Perrett et al. 1992). The unique geometric properties of each view, such as the set of diagnostic 2D shapes visible, may differentially recruit perceptual and visuospatial processes, but this seems more important for non-sense; wire-frame objects with uniform parts visible from any angle (Cutzu and Edelman 1994), including stimuli used in seminal mental rotation studies (Shepard and Cooper 1982). For meaningful known objects, learning and memory are more important. Which views are stored depends upon the views learned, their exposure frequency, object structure, context, and the tasks performed on the object during encoding and retrieval (Edelman and Bulthoff 1992; Cutzu and Edelman 1994; Srinivas 1995; Blanz et al. 1999).

Dual visual stream frameworks posit that, whereas the dorsal visual pathway is responsible for action and spatial vision, the ventral pathway is responsible for perception and memory (Ungerleider and Mishkin 1982; Goodale and Milner 1992; Perrett et al. 1998), and object- and category-sensitive areas in human ventral posterior cortex have well-established roles in visual object cognition, learning, and memory (Kourtzi and Kanwisher 2000; Bar et al. 2001; Grill-Spector et al. 2001; Hasson et al. 2003; Courtney 2004; Kan and Thompson-Schill 2004; Price and Devlin 2003; van Turenout et al. 2003). Our findings of view sensitivity in LOS, ITS, and CoS/LS regions implicate these as sites from which view-specific object models and their features are retrieved and a source of a view atypicality signal driving VLPFC-posterior network interactions for model verification. Multiple memory systems theories posit generic explicit memory stores semantic knowledge of concepts and facts, such as a function or name, and nonsemantic knowledge (Tulving 1972; Hintzman 1978). The candidate object models evaluated during the object model selection process supporting categorization are stored as nonsemantic knowledge about visual structure (Warrington 1982; Humphreys et al. 1999). DOT cortex (LOS, ITS) has been implicated in computing spatial aspects of objects and their configuration of parts (e.g., stereoscopic structure, symmetry, depth), and relations between objects and space (e.g., position, motion, kinematics, scale) (Van Oostende et al. 1997; Grill-Spector et al. 1999; Gilaie-Dotan et al. 2001; Moore and Engel 2001; Kourtzi et al. 2002; Hasson et al. 2003; Sasaki et al. 2005; Behrmann et al. 2006). Monkey inferotemporal neurons code object parts in 2D but with some depth structure from qualitative information about disparity gradient, depth curvature, or horizontal disparity (Tanaka 2005). Multiple views accounts suggest that the 2D views have some depth or stereo information (Edelman and Bulthoff 1992). For categorization more than recognition, the LOS seems particularly important for the model prediction functions of the aVLPFC, as both regions show similar patterns of view and task effects. We propose that the ITS has the most task-general role in object cognition and in achieving visual constancy: ITS is modulated on categorization, recognition, mental rotation, and working memory tasks (Schon et al. 2004; Schendan and Stern 2007), and interacts with the right pVLPFC and occipitoparietal regions for model testing. Successful model selection can subsequently activate associated information in semantic memory, and semantic retrieval recruits the ITS (Tulving 1972; Kan and Thompson-Schill 2004), which may provide convergent information about both nonsemantic and semantic knowledge (Damasio et al. 1996).

For model testing, the CoS/LS area can be additionally recruited. We found evidence that right CoS/LS is viewsensitive, except for uncategorized objects, which by definition, have failed to match to any nonsemantic visuo-(structural) or semantic memory. The right CoS/LS has been found to be activated on N-back and recognition tasks for scenes, landmarks, and buildings relative to other images, including objects, and shows view-specificity (Aguirre et al. 1998; Epstein and Kanwisher 1998; Epstein et al. 2003, 2005, 2007; Hasson et al.
2003; Tsivilis et al. 2003; Wheeler and Buckner 2003). CoS/LS may be analogous to a part of higher-order visual area TE in monkeys, providing key input to medial temporal structures for explicit episodic memory, including parahippocampal gyrus, activation of which can be associated (Fig. 5) (Suzuki and Amaral 1994; Stern et al. 1996; Buffalo et al. 1999; Kirchhoff et al. 2006; Hasson et al. 2003; Rosenbaum et al. 2004; Yonelinas et al. 2005; Wais et al. 2006). CoS/LS effects often co-occur also with effects in TOS/vcIPS and retrosplenial cortex (Epstein et al. 2007), consistent with our results, and we find co-occurrence also with ITS on recognition, mental rotation (Schendan and Stern 2007), and working memory tasks (Schon et al. 2004), suggesting CoS/LS integrates input from the ITS, TOS, and vcIPS. We suggest that the CoS/LS area is recruited for perceiving and representing a more unitary spatial configuration of the features or parts of scenes, landmarks, or other objects as needed by the task to evaluate the detailed visuospatial match between an internal representation and a current percept or another internal representation(s). Detailed spatial information can help individuate between perceptually similar instances of a place or an object.

Conclusions
Visual object categorization and recognition tasks are where vision meets memory. The 2SI account is an initial attempt to develop a cognitive framework that integrates vision and memory theories. This account is compatible with psychophysical and computational models, including multiple views theories (perhaps facilitating processes of interrogation, evidence accumulation, or pooling across a neuronal population), and bottom-up architectures that could provide the initial object model(s) for the prediction-testing loops (Serre et al. 2007). By a 2SI account, after 200 ms, distinct VLPFC-posterior networks are engaged to facilitate object constancy. An aVLPFC-LOS network supports prediction verification processes on categorization more than recognition. A pVLPFC-vcIPS-POS-ITS network supports testing verification processes on diverse tasks. Right CoS/LS is also involved for not only view representation but also detailed spatial matching. Most important, the present findings demonstrate that not only the ventral stream but also object-sensitive occipitoparietal cortex (vcIPS, TOS), which is critical for mental rotation, also plays a key role in object constancy, computing the relative spatial relations among object parts, conscious spatial binding, and transforming hierarchical visuospatial structure. This framework may apply to other perceptual and memory situations in other sensory modalities when stimulus interpretation is ambiguous.

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

Funding
National Institute on Aging grant (F32 AG05914) and Tufts University Faculty start-up funds to H.E.S.; National Institute of Mental Health grant (P50 MH071702) to C.E.S. and H.E.S.; National Center for Research Resources grant (P41 RR14075); and the Mental Illness and Neuroscience Discovery (MIND) Institute.

Notes
Brain scanning was conducted at the MGH/MIT/HMS Athinoula A. Martinos Center for Biomedical Imaging. The Optseq software used to optimize fMRI signal discriminability was developed and made freely available at the Martinos Center. This study was conducted in the Vision and Memory Neuroimaging Lab. Department of Psychology at Tufts University and the Vision and Cognition Lab and Cognitive Neuroimaging Lab at the Center for Memory and Brain, Department of Psychology, Boston University. We are grateful to Courtney Horwitz, Mimi Castelo, Sule Tinzaz, MD, PhD, Michael J. Tavantzis, and Reyyan A. Bilge for assistance with data acquisition and/or analyses, and Dr Alice Cronin-Golomb for computer hardware support. Many thanks to Dr. Karin Schon for examining results from Schon et al. (2004) for delay and subsequent memory effects in the subregion of CoS/LS, and providing the results included in Figure 5 and Supplemental Figure 6 (panels D and G in both). We thank Dr Giorgio Ganis for helpful discussion. Conflict of Interest None declared.

Address correspondence to Haline E. Schendan, PhD, Department of Psychology, Tufts University. The Psychology Building, 490 Boston Avenue, Medford, MA 02155, USA: Email: Haline_E.Schendan@tufts.edu.

References


