

# Higher Level Visual Cortex Represents Retinotopic, Not Spatiotopic, Object Location

Julie D. Golomb and Nancy Kanwisher

McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Address correspondence to Julie D. Golomb, McGovern Institute for Brain Research, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Room 46-4141, Cambridge, MA 02139, USA. Email: jgolomb@mit.edu.

**The crux of vision is to identify objects and determine their locations in the environment. Although initial visual representations are necessarily retinotopic (eye centered), interaction with the real world requires spatiotopic (absolute) location information. We asked whether higher level human visual cortex—important for stable object recognition and action—contains information about retinotopic and/or spatiotopic object position. Using functional magnetic resonance imaging multivariate pattern analysis techniques, we found information about both object category and object location in each of the ventral, dorsal, and early visual regions tested, replicating previous reports. By manipulating fixation position and stimulus position, we then tested whether these location representations were retinotopic or spatiotopic. Crucially, all location information was purely retinotopic. This pattern persisted when location information was irrelevant to the task, and even when spatiotopic (not retinotopic) stimulus position was explicitly emphasized. We also conducted a “searchlight” analysis across our entire scanned volume to explore additional cortex but again found predominantly retinotopic representations. The lack of explicit spatiotopic representations suggests that spatiotopic object position may instead be computed indirectly and continually reconstructed with each eye movement. Thus, despite our subjective impression that visual information is spatiotopic, even in higher level visual cortex, object location continues to be represented in retinotopic coordinates.**

**Keywords:** eye position, multivoxel pattern analysis, object recognition, reference frame, ventral visual stream

## Introduction

How, and at what stage of visual processing, is information about an object’s location coded in the brain? It has been proposed that the visual system contains 2 separate visual processing pathways, a ventral “what” stream for object recognition and a dorsal “where” stream for spatial processing (Ungerleider and Mishkin 1982). However, recent evidence has challenged this classic dissociation, with reports of location representations in ventral areas (DiCarlo and Maunsell 2003; Hung et al. 2005; Sayres and Grill-Spector 2008; Schwarzlose et al. 2008; Arcaro et al. 2009; Kravitz et al. 2010; Carlson et al. 2011) and object representations in dorsal areas (Serenio and Maunsell 1998; Lehky and Sereno 2007; Janssen et al. 2008; Konen and Kastner 2008). Even if object location and object identity are represented in the same regions, they may still be processed independently, allowing for location-invariant information about object category and category-invariant information about object location (Hung et al. 2005; Schwarzlose et al. 2008).

The fact that location information is present throughout early, ventral, and dorsal visual areas begs the question of how

location is represented in these areas, and whether different visual areas code different types of location information. Specifically, is this location information coded relative to the eyes (retinotopic position) or relative to the world (spatiotopic position)? While the initial input into visual cortex is retinotopic, our experience of visual stability across eye movements suggests that this information must at some point be transformed into spatiotopic representations. Thus, it seems possible that while early visual cortex (EVC) codes location retinotopically, higher level visual cortex might accommodate spatiotopic information.

Topographic maps of spatial location have been reported throughout early (Serenio et al. 1995; Engel et al. 1997), ventral (Brewer et al. 2005; Larsson and Heeger 2006; Arcaro et al. 2009), and dorsal (Serenio et al. 2001; Silver et al. 2005; Swisher et al. 2007) visual regions; however, with the eyes held at fixation, retinotopic and spatiotopic positions are confounded. In the studies that have used changes in eye position to dissociate these 2 reference frames, early visual regions have been confirmed to be retinotopically organized (Gardner et al. 2008; Golomb et al. 2010; Crespi et al. 2011). However, much less is known about the organization of higher visual areas. Might any of these areas contain spatiotopic information? Parietal areas represent one putative source (Zipser and Andersen 1988; Galletti et al. 1993; Duhamel et al. 1997; Snyder et al. 1998; Pertzov et al. 2011), and the dorsal stream middle temporal visual area (MT) has also been debated (d’Avossa et al. 2007; Gardner et al. 2008; Crespi et al. 2011). The properties of ventral visual cortex also offer interesting potential for spatiotopic representations, with their relatively large receptive fields (Gross et al. 1972; MacEvoy and Epstein 2007) and tolerance across changes in position (Ito et al. 1995; Grill-Spector and Malach 2001). In particular, area lateral occipital complex (LOC) may be sensitive to relative position (Hayworth et al. 2011), perceived position (Fischer et al. 2011), and even spatiotopic position (McKyton and Zohary 2007; although see Gardner et al. 2008). Yet a systematic exploration of the coordinate systems of higher level visual cortex has never been conducted. Determining the type of location information represented in these ventral regions is particularly important given their role in object recognition, and the possibility they might be involved in combining information about “what” objects are with “where” they are in the world.

## Materials and Methods

We used multivariate pattern analysis (MVPA) techniques to test whether and where information about object category, retinotopic position, and spatiotopic position is present in the visual processing stream. Multivariate analyses allow for more sensitive and fine-scale probing of information content by asking if the spatial pattern of functional magnetic resonance imaging (fMRI) response can differentiate between different types of information (Haxby et al. 2001;

Kamitani and Tong 2005). In the present paper, we use this technique to ask whether given brain regions contain explicit information about object category, retinotopic, and spatiotopic position. By “explicit,” we specifically mean information that can be linearly decoded (here with MVPA). A distinction is often made between explicit representations (as defined above) and implicit representations, in which information about a stimulus property may be indirectly represented in a brain region but not in any way that can be read out with a linear decoder (deCharms and Zador 2000; Connor et al. 2007). Explicit (linear) representations have been argued to reflect biologically plausible neuronal processing; for example, while area V1 could be said to implicitly represent all of the necessary information to support object recognition, only at the level of inferior temporal cortex can this information be read out by a linear classifier, paralleling the response properties of the ventral processing stream (DiCarlo and Cox 2007).

### Subjects

Eight subjects (6 females; mean age 26.8, range 21–35) participated in at least one experiment. All 8 subjects participated in Experiment 1 and subsets of 4 subjects each (partially overlapping) participated in Experiments 2 and 3. All subjects were neurologically intact with normal or corrected-to-normal vision. Informed consent was obtained for all subjects, and the study protocols were approved by the Massachusetts Institute of Technology Committee On the Use of Humans as Experimental Subjects.

### fMRI Acquisition

MRI scanning was carried out with a Siemens Trio 3-T scanner using a 32-channel receiver array head coil. Functional data were acquired with a  $T_2^*$ -weighted gradient-echo sequence (repetition time = 2500 ms, echo time = 30 ms, flip angle = 90°, matrix = 192 × 192). Parallel imaging with an acceleration factor of 2 was used, and 36 axial slices were taken oriented roughly parallel to the calcarine sulcus, to maximize high-resolution (1.5 × 1.5 × 2 mm voxels, 0.4-mm interslice gap) coverage of occipital, parietal, and posterior temporal cortices.

All subjects were scanned for multiple sessions across different days: Every subject completed one session of Experiment 1 (10–12 runs) and functional localizers (3 runs) and one session of retinotopic mapping (4–8 standard retinotopic mapping runs and 6–8 delayed saccade mapping runs). For the subjects that participated in additional experiments, Experiment 2 was conducted in a single session (12 runs) and Experiment 3 across 2 sessions (14 runs each). The functional data from each session were first coregistered to a high-resolution 3D magnetization prepared rapid gradient echo anatomical scan taken in the same session; anatomical scans were then coregistered to the anatomical scan from the initial session, such that functional data were precisely aligned across sessions.

Stimuli were generated using the Psychtoolbox extension (Brainard 1997) for Matlab (The Mathworks, Inc., Natick, MA) and displayed with an LCD projector onto a screen mounted in the rear of the scanner bore, which subjects viewed from a distance of 120 cm via a mirror attached to the head coil (maximal field of view: 21°).

### Eye Tracking

Eye position was monitored using a modified ISCAN eye-tracking system (ISCAN, Inc., Burlington, MA), with the camera and infrared source placed directly in front of the bottom of the rear screen. Pupil and corneal reflection (CR) were recorded at 120 Hz and analyzed offline to ensure accurate fixation performance. The eye tracker was calibrated at the beginning of the session and repeated between runs if necessary and time permitted. Occasionally, the eye-tracker signal in the scanner was too noisy to achieve reliable calibration, and eye position was monitored via video trace. All subjects were expert subjects practiced in eye-tracking tasks outside the scanner, and fixation behavior was very reliable. We did not exclude individual trials for imperfect fixation behavior; however, we repeated the analyses including only subjects for whom calibration was accurate enough to ensure precise fixation (within 2°) on greater than 90% of the run for at least 10 runs and found the same pattern of results.

### Stimulus Arrangement

The stimulus arrangement consisted of 2 possible fixation locations (located 7° apart, centered on the screen) and 3 possible stimulus locations (left, middle, and right; Fig. 1). The middle location was positioned directly between the fixation locations and the 2 outer locations were positioned at equidistant eccentricity (3.5° from nearest fixation to center of the image). This arrangement allowed manipulation of both eye position and stimulus position to generate pairs of conditions in which stimuli appeared in different retinotopic positions but the same spatiotopic position, the same retinotopic position but different spatiotopic positions, the same in both retinotopic and spatiotopic positions, or different in both. The stimulus locations were marked by 3 faint gray placeholder squares (5.25° × 5.25°) that were continually present on the screen. The placeholders were included to create and reinforce a spatiotopic reference frame. On each trial a stimulus of the same size appeared replacing one of the placeholders; fixation crosses were presented at only one fixation location at a time.

### Experiment 1

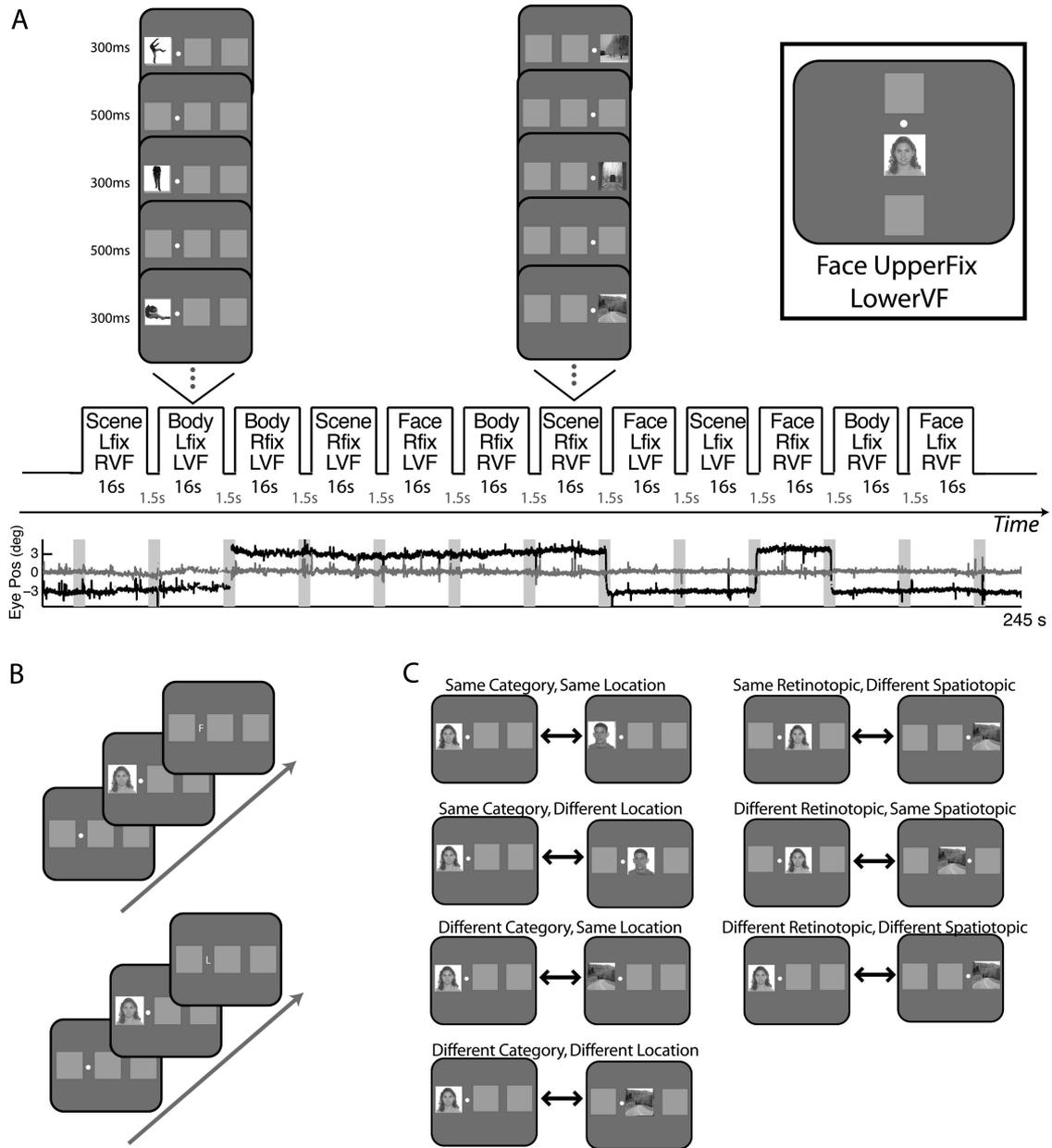
Subjects fixated on a fixation cross located to the left or right of the center of the screen while viewing blocks of faces, scenes, and bodies. The images were presented either to the left or to the right of fixation (in the immediately adjacent stimulus position), generating 4 different location combinations varying in eye position and stimulus position (Fig. 1A). Each of these 12 conditions (3 stimulus categories × 4 locations) was presented in a blocked fashion, one block per run, in pseudorandomized order, as in the Schwarzlose et al. (2008) task. Each block lasted 16 s and consisted of 20 images presented in succession (300-ms stimulus presentation, 500-ms interstimulus interval). A 1.5-s fixation interval separated each block, and the 245 s run started and ended with fixation-only blocks. Subjects completed 10–12 runs during the scanning session. Eye position was monitored to ensure successful fixation, and subjects performed a 1-back task in which they were instructed to press a button whenever an image was repeated consecutively in the stream. Stimuli were grayscale images drawn from pools of 40 faces, 40 scenes, and 40 headless bodies.

### Experiment 2

Experiment 2, which tested whether spatiotopic information might be more evident if it did not cross hemispheres, was identical to Experiment 1, except the fixation and stimulus locations were arranged vertically instead of horizontally on the screen (Fig. 1A, inset). To fit within the viewable screen dimensions, fixation positions were positioned above or below center, 5.2° apart, and stimuli were sized 3.9° and positioned at a 2.6° eccentricity above or below fixation.

### Experiment 3

Experiment 3, which tested whether the prevalence of category, retinotopic, and spatiotopic information is affected by task, used the same stimulus arrangement as Experiment 1 but employed a fast event-related design (Fig. 1B). Subjects performed 2 different tasks, one per scanning session, with order counterbalanced across subjects. As each stimulus appeared, subjects made a button press response based on the category of the stimulus (Category Task: face or scene) or the spatiotopic location of the stimulus (Location Task: left, middle, or right position). However, because direct report of the category or location would confound manual response with the stimulus property being measured (see Supplementary Material), we implemented an indirect report technique. Stimuli were presented for 300 ms, and subjects were instructed to attend to the relevant property as each stimulus appeared. At stimulus offset, the fixation dot was replaced by a small letter (F or S in the Category Task representing “face” or “scene”; L, M, or R in the Location Task representing “left,” “middle,” or “right”). If the letter correctly matched the relevant stimulus property, subjects pressed a button with their index finger; if the letter did not match they pressed a different button with their middle finger. For example, in the Category Task, if the letter “F” followed presentation of a face, the answer was true; in the Location Task, if an “R” followed a stimulus in the center position, the answer was false. Performance on



**Figure 1.** Stimuli and design. (A) Experiments 1 and 2. Stimuli were faces, scenes, and bodies presented to the left or right of fixation, at 1 of 2 fixation locations. Stimuli were presented in a block design, with one block per each of the 12 conditions (scrambled order) per run. Each block lasted 16 s and consisted of a series of 20 images from that condition. Subjects performed a one-back repetition task and pressed a button whenever the same exact stimulus exemplar was presented twice in a row. A sample eye trace from an entire run of a single subject is shown beneath the schematic showing block order for that run; black trace is horizontal eye position, gray is vertical eye position, and light gray bars denote the beginning of each block. An example stimulus from Experiment 2 is shown in the inset at top. The design was the same as Experiment 1, using vertically arranged stimulus configurations. (B) Experiment 3. Stimuli were faces and scenes presented in the left, middle, or right positions at both fixation locations. Stimuli were presented one at a time in a fast event-related design. Each stimulus appeared for 300 ms, followed by a probe. In the Category Task, subjects reported the category of each stimulus by making a true/false button press response indicating whether the probe (F for face, S for scene) matched the stimulus category. In the Location Task, subjects reported the spatiotopic location of each stimulus by making a true/false button press response indicating whether the probe (L for left, M for middle, R for right) matched the spatiotopic stimulus location. (C) Examples of pairs of conditions used for “same” and “different” comparisons. Left column: category and location comparisons for pairs sharing the same eye position. Note that location is the combined retinotopic and spatiotopic locations because there is no difference in eye position. Right column: retinotopic and spatiotopic comparisons across eye positions. Only different-category comparisons are depicted here.

both tasks was near ceiling (Category Task: mean 98.6%, standard deviation [SD] 1.1%; Location Task: mean 98.1%, SD 2.1%). Subjects had 2 s to respond, after which the letter was removed, and the fixation cross was positioned for the upcoming trial. Stimulus onset asynchrony (SOA) was 2.5, 5, or 7.5 s. Jittered SOA and pseudorandomized trial order were optimized to reduce serial correlations between conditions. To minimize frequent back-and-forth eye movements, fixation position was blocked and alternated every 10 trials (8 alternations per run).

For the Location Task, it was important that stimuli could appear in all 3 locations for both fixation positions. We thus included a small portion of trials in Experiment 3 in which stimuli appeared at the more remote position (e.g., left fixation, far right stimulus and right fixation, and far left stimulus). Because these stimuli appeared at a greater visual eccentricity and did not have an equivalent retinotopic position at the other fixation location, we did not include these trials in the MVPA analyses. To compensate for the increased number of location conditions, we reduced the number of category conditions and only

included faces and scenes in this experiment. The conditions and number of trials per run were as follows: Face\_Lfix\_Lvisualfield (8), Face\_Lfix\_Rvisualfield (8), Face\_Lfix\_farRvisualfield (4), Face\_Rfix\_Lvisualfield (8), Face\_Rfix\_Rvisualfield (8), Face\_Rfix\_farLvisualfield (4), Scene\_Lfix\_Lvisualfield (8), Scene\_Lfix\_Rvisualfield (8), Scene\_Lfix\_farRvisualfield (4), Scene\_Rfix\_Lvisualfield (8), Scene\_Rfix\_Rvisualfield (8), and Scene\_Rfix\_farLvisualfield (4).

### **Functional Localizers and Retinotopic Mapping**

Category-selective ventral stream areas were identified in each subject individually using a localizer task. Blocks of faces, scenes, headless bodies, everyday objects, and scrambled versions of the objects were presented with the same timing and one-back task as Experiment 1. Fixation was always in a central position, and images appeared simultaneously on both sides of fixation at the same size and eccentricity as Experiment 1. Images on the left and right were always identical; both visual fields were stimulated to replicate the stimulus positions in the main task.

Category-selective regions were identified separately for each subject using the following contrasts: fusiform face area (FFA: Kanwisher et al. 1997): faces > objects, extrastriate body area (EBA: Downing et al. 2001): bodies > objects, parahippocampal place area (PPA: Epstein and Kanwisher 1998): scenes > faces. Object-selective region LOC (Kourtzi and Kanwisher 2000) was identified using an object > scrambled contrast. Regions of interest (ROIs) were defined as clusters of at least 25 contiguous voxels exceeding an uncorrected  $P < 0.001$  statistical threshold. All ROIs were identified in at least one hemisphere of every subject.

An additional localizer task was conducted to identify motion-sensitive area MT/V5 (Tootell et al. 1995). Subjects fixated at the center of the screen while viewing blocks of either stationary or moving random dot displays. The stimuli were full screen dot patterns (Huk et al. 2002); moving patterns alternated between concentric motion toward and away from fixation at 7.5 Hz. Area MT+ was defined with a moving > stationary contrast.

Early visual and parietal regions were identified (again on an individual subject basis) using 2 complementary techniques. First, an all conditions > fixation contrast was carried out on the localizer data to select bilateral ROIs for EVC (covering approximately V1–V3) and intraparietal sulcus (IPS). For more detailed analysis of these regions, retinotopic mapping was carried out to separate individual retinotopic regions.

Retinotopic mapping was conducted on all subjects using standard checkerboard rotating wedge and expanding/contracting ring stimuli to map polar angle and eccentricity of early visual regions (Engel et al. 1994; Sereno et al. 1995; DeYoe et al. 1996). High-contrast radial checkerboard patterns were presented as 60° wedges or rings and flickered at 4 Hz. Maximal eccentricity was 12°, and the central 1.2° foveal region was not stimulated. Each run rotated clockwise or counter clockwise or expanded or contracted through 7 cycles with a period of 24 s/cycle. Subjects fixated at the center of the display and pressed a button every time the black fixation dot dimmed to gray.

A memory-guided saccade task was also used to map topographic areas of parietal cortex (Sereno et al. 2001; Schluppeck et al. 2005; Swisher et al. 2007; Konen and Kastner 2008). Subjects fixated at the center fixation dot, and a cue briefly appeared (500 ms) in the periphery (6.65° ± 1° eccentricity) indicating the location of the upcoming saccade. Subjects remained fixated as a ring of distractors flickered at 4 Hz for 3 s. When the distractors disappeared, subjects had 1.5 s to saccade to the remembered location and return their eyes to the central fixation dot before the next cue appeared. Saccades were directed sequentially to 8 polar angle locations for each 40 s cycle. Each run consisted of 8 cycles of a clockwise or counter-clockwise sequence.

Data from both techniques were analyzed using standard phase-encoded analysis methods: the best-fitting phase and correlation coefficient was obtained for each voxel using Fourier analysis and averaged across clockwise and counter-clockwise runs to compensate for hemodynamic lag. Phase angle maps, thresholded based on correlation coefficient, were displayed on the flattened cortex. Visual field boundaries were defined following standard phase-reversal criteria (Sereno et al. 1995; DeYoe et al. 1996; Engel et al. 1997; Larsson and

Heeger 2006; Swisher et al. 2007) and bilateral ROIs were created for each of the following regions: V1, V2, V3, ventral V4, V3A, and V7 from the checkerboard mapping and IPS1–4 from the combination of delayed saccade mapping and checkerboard mapping. ROIs for early visual areas were restricted to the approximate visual field location/eccentricity of the stimuli in the main task.

The following regions were identified in at least one hemisphere of each subject: LOC (right hemisphere: 8 subjects, mean Talairach coordinates [40,–72,–5]; left hemisphere: 8 subjects, [–36,–75,0]), PPA (right: 8 subjects, [25,–48,–11]; left: 8 subjects, [–27,–49,–10]), FFA (right: 8 subjects, [36,–46,–18]; left: 5 subjects, [–39,–46,–16]), EBA (right: 8 subjects, [44,–70,2]; left: 6 subjects, [–43,–78,7]), MT+ (right: 8 subjects, [42,–68,1]; left: 8 subjects, [–43,–72,1]), IPS (right: 8 subjects, [26,–56,43]; left: 8 subjects, [–25,–60,46]), and EVC (right: 8 subjects, [12,–84,–3]; left: 8 subjects, [–16,–84,–4]). Retinotopic regions V1–V7 and IPS1–4 were also identified in every subject. Significant differences were not found across hemispheres, so data were combined across both hemispheres of a given region for subjects with bilateral ROIs. Results were highly similar among the individual early visual regions (retinotopic V1, V2, V3, V4, and localizer EVC) and the parietal regions (retinotopic IPS1, IPS2, IPS3, IPS4, and localizer IPS), so data for the individual retinotopically defined regions are only shown for Experiment 1 MVPA; remaining analyses focus on the localizer-defined regions, including representative regions EVC and IPS.

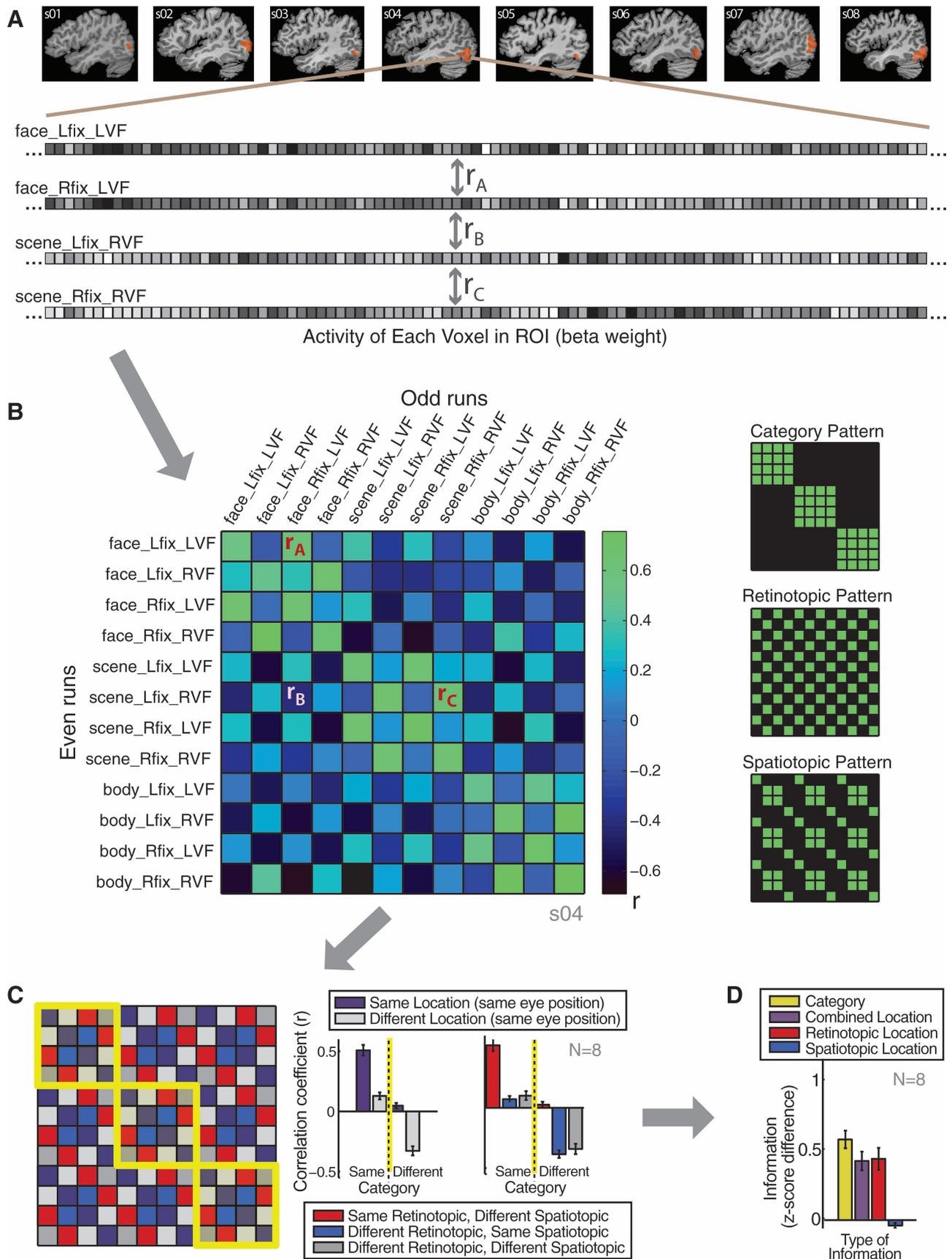
### **fMRI Preprocessing and Analysis**

Preprocessing of the data was done using Brain Voyager QX (Brain Innovation). All data were corrected for slice acquisition time and head motion, temporally high-pass filtered with a 128-s period cutoff, normalized into Talairach space (Talairach and Tournoux 1988) and interpolated into 2 mm isotropic voxels. The linear Talairach normalization was conducted to warp individual brains into a common space, which allowed us to report and compare ROI coordinates as well as average across subjects for the group searchlight analysis. With the exception of the group searchlight analysis, all analyses reported here were conducted entirely on an individual subject level based on ROIs that were functionally localized for each individual, regardless of Talairach location. Localizer task data were spatially smoothed with a 4-mm full-width at half-maximum kernel; data from the main experiments were left unsmoothed. For retinotopic mapping analyses, the high-resolution 3D anatomical images were used to create flattened representations of the cortical surface for each hemisphere, after segmenting the gray and white matter, inflating the cortical sheet, and cutting and unfolding the inflated brain along 5 segments, including the calcarine sulcus.

Multiple regression analyses and phase-encoded analyses were performed separately for each subject to obtain subject-specific ROIs from the localizer and retinotopic mapping runs. Then, for each main experiment, a whole-brain random-effects general linear model (GLM), using a canonical hemodynamic response function, was used to extract beta weights for each voxel, for each condition and subject. Separate GLMs were run for odd runs and even runs to allow split-half comparison of voxelwise patterns in each ROI (Haxby et al. 2001). Data were exported to Matlab (Mathworks) using Brain Voyager's BVQXtools Matlab toolbox, and all subsequent analyses were done in Matlab.

### **Multivoxel Correlation Analysis**

Multivoxel pattern analysis was performed separately for each subject and ROI following the method of Haxby et al. (2001). Figure 2 illustrates the entire workflow. First, the data from each ROI were split into an odd runs data set and an even runs data set. For each data set separately, the mean response across all conditions was subtracted from the responses to individual conditions, normalizing each voxel's response. Next, the voxelwise response patterns for each of the 12 conditions in the even runs were correlated with each of the 12 conditions in the odd runs, generating 144 correlations. These correlations were then classified as same or different in 1) category, 2) "combined location," 3) retinotopic location, and 4) spatiotopic location (Fig. 1C). For example, a same category, same retinotopic



**Figure 2.** Analysis workflow. Sample data are shown for each stage of the MVPA analysis for a single region (LOC). (A) ROIs were chosen separately for each individual based on independent functional localizers; each subject's LOC is shown overlaid in orange. Below, examples of actual responses (real data) to 4 different conditions for an array of 100 voxels from one subject's LOC are shown. The data were normalized to remove mean response. The voxelwise patterns evoked by each of the 12 conditions in odd and even runs

correlation could be Face\_Lfix\_LVF(even) versus Face\_Rfix\_LVF(odd); a different category, same retinotopic correlation could be Face\_Lfix\_LVF(even) versus Scene\_Rfix\_LVF(odd).

Finally, the amount of category and location information contained within an ROI was quantified by converting the correlations to  $z$ -scores and subtracting the “same” minus “different” correlations for that type of information. Category information was quantified as “same Category” – “different Category.” Location information was quantified using 2 sets of analyses:

1. Analyses of Location Information for Same Fixations: “Combined Location.” For comparisons in which eye position was the same, location could be considered a combination of spatiotopic and retinotopic components. Combined location information was calculated as the difference between both retinotopic and spatiotopic staying the same and both retinotopic and spatiotopic changing:

$$\text{Combined location} = \begin{pmatrix} \text{same Retinotopic,} \\ \text{same Spatiotopic,} \\ \text{same FixLoc} \end{pmatrix} - \begin{pmatrix} \text{different Retinotopic,} \\ \text{different Spatiotopic,} \\ \text{same FixLoc} \end{pmatrix}$$

2. Analyses of Location Information for Different Fixations: Retinotopic and Spatiotopic Information. Retinotopic and spatiotopic information could only be measured for comparisons involving a change in eye position. To control for any eye position effects, they were compared with a baseline involving a similar change in eye position but with both retinotopic and spatiotopic varying:

$$\text{Retinotopic} = \begin{pmatrix} \text{same Retinotopic,} \\ \text{different Spatiotopic,} \\ \text{different FixLoc} \end{pmatrix} - \begin{pmatrix} \text{different Retinotopic,} \\ \text{different Spatiotopic,} \\ \text{different FixLoc} \end{pmatrix}$$

$$\text{Spatiotopic} = \begin{pmatrix} \text{different Retinotopic,} \\ \text{same Spatiotopic,} \\ \text{different FixLoc} \end{pmatrix} - \begin{pmatrix} \text{different Retinotopic,} \\ \text{different Spatiotopic,} \\ \text{different FixLoc} \end{pmatrix}$$

The correlation coefficients were  $z$ -transformed, and random-effects statistical analyses were conducted to assess whether the amounts of Category, Combined Location, Retinotopic, and Spatiotopic information were significant. For the same-fix data, 2 (category)  $\times$  2 (location) analyses of variance (ANOVAs) were run for each region comparing same and different category and combined location information. For the different-fix data, an omnibus 7 (ROI)  $\times$  2 (spatiotopic vs. retinotopic) ANOVA was run on the information values comparing retinotopic and spatiotopic information across the 7 regions. Post hoc 2-tailed one-sample  $t$ -tests were then run comparing each information value to zero (indicating no difference in correlation strength for “same” vs. “different” correlations). Significant values for an ROI indicate that the voxels in that region evoke more similar patterns of response when 2 stimuli share a given property than when they differ in that property, implying that the region carries information about that property.

### Searchlight Analysis

To see whether information about category, retinotopic, or spatiotopic location was present outside the a priori ROIs, we conducted a “searchlight” analysis (Kriegeskorte et al. 2006) across our entire slice coverage. For each subject, we iteratively searched through the brain conducting MVPA within a “moving” ROI defined as a sphere of radius 3 mm (ca. 100 voxels). On each iteration, the ROI was chosen as a sphere centered on a new voxel, and multivoxel correlation analyses were performed exactly as described above. The magnitudes of

category, retinotopic, and spatiotopic information (as defined by the  $z$ -transformed “same” – “different” correlation differences) were then plotted for each voxel, creating a  $z$ -map for each type of information for each subject. Statistical thresholds for the individual subject “information maps” were determined using permutation tests, as described below.

For each subject, condition labels were shuffled (e.g., by assigning run 1 the labels for run 7, keeping the shuffling in the space of real possibilities). The searchlight analysis was rerun on the shuffled data, and each voxel was assigned a correlation value ( $z$ -scored) reflecting retinotopic, spatiotopic, and category information. The permutation analysis was repeated 10 times to create a “chance” distribution for each type of information. Because each voxel is a sample in this distribution (Gardner et al. 2008), 10 permutations result in a distribution of about 2 600 000 samples (10 permutations  $\times$  260 000 voxels). The  $P < 0.01$  statistical threshold was calculated from the permuted distributions, separately for each subject and type of information. These thresholds were then applied to the individual subjects’ searchlight maps. Clusters above this threshold reflect retinotopic, spatiotopic, or category information significantly greater than expected by chance, without making any assumptions about the underlying distribution.

Group-averaged maps were also calculated for retinotopic, spatiotopic, and category information. Searchlight maps were combined across subjects using one-sample  $t$ -tests to identify clusters containing significant information about a given stimulus property. The resulting  $t$ -maps were statistically thresholded at  $P < 0.01$ , cluster thresholded at 125 contiguous voxels, and projected back onto inflated brains for visualization.

### Spatiotopicity Index Analysis

We also analyzed our data with a univariate reference frame approach described in prior fMRI studies (d’Avossa et al. 2007; Gardner et al. 2008; Crespi et al. 2011). The approach is based on a similar idea to MVPA: according to a retinotopic prediction, 2 stimuli in the same retinotopic position should evoke more similar responses, but according to a spatiotopic prediction, 2 stimuli in the same spatiotopic position should evoke more similar responses. Stimuli were aligned according to retinotopic and spatiotopic predictions, and the summed squared difference in mean response amplitude was calculated for stimuli sharing the same retinotopic position ( $\text{resid}_R$ ) and same spatiotopic position ( $\text{resid}_S$ ), where smaller residuals indicate that the BOLD response more closely fits a given prediction. The spatiotopicity index was then calculated as

$$\text{SI} = \frac{\text{resid}_S - \text{resid}_R}{\text{resid}_S + \text{resid}_R}$$

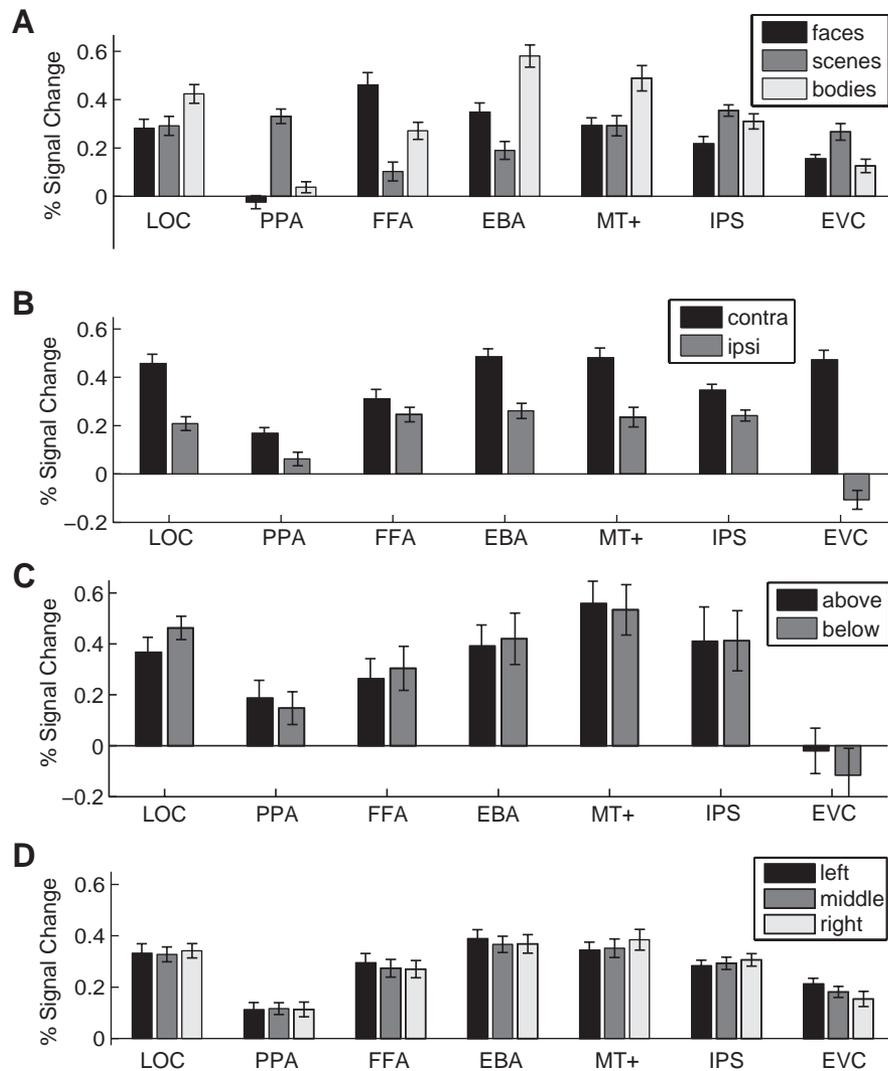
where positive values indicate retinotopic behavior and negative values indicate spatiotopic behavior.

## Results

### Mean Response Magnitudes

For each ROI, the average response magnitudes across the entire ROI were calculated for each condition using traditional univariate methods. Collapsing across location, each of the scene-, face-, and body-selective regions exhibited the expected category selectivity (Fig. 3A). Responses to the preferred category were significantly greater than to the second highest category (PPA: scenes > bodies,  $t_7 = 15.51$ ,  $P < 0.001$ ; FFA: faces >

were then correlated with each other ( $B$ ). The  $12 \times 12$  correlation matrix for a single subject’s left LOC is illustrated—these data were calculated directly from the voxel data displayed in  $A$ : cells marked  $r_A$ ,  $r_B$ , and  $r_C$  represent the 3 correlations indicated. Next to the actual correlation matrix, inset illustrates hypothetical correlation matrices expected from pure category information, retinotopic information, and spatiotopic information. ( $C$ ) Each cell in the correlation matrix was coded according to whether it represented comparisons between conditions that had the same or different category, combined location, retinotopic location, and spatiotopic location. Colors in the key at left correspond to the different conditions in the figure and legends at right. This analysis was done on individual subjects, then averaged across subjects to produce the data shown. ( $D$ ). Correlation coefficients were  $z$ -scored, and the amount of category and location information was quantified by taking the difference of [same – different] for each type. The data in  $C$  and  $D$  are the actual data for area LOC shown in Figure 5 and submitted to group statistics.



**Figure 3.** Mean response magnitudes. Mean responses across the entire ROI (traditional univariate approach) are shown for each region. (A,B) Experiment 1. (A) Responses to each category, collapsing across location. The scene-, face-, and body-selective regions exhibited the expected preferred responses. (B) Responses to contralateral and ipsilateral stimuli, collapsing across category. All regions show some degree of contralateral preference. (C) Experiment 2, responses to stimuli above and below fixation, with little difference between conditions. (D) Responses as a function of spatiotopic location (Experiment 1). Collapsing across fixation location and category, there was no difference in response for stimuli appearing in the left, middle, or right positions on the screen. Error bars are standard error of the mean.

bodies,  $t_7 = 4.80$ ,  $P = 0.002$ ; EBA: bodies > faces,  $t_7 = 10.91$ ,  $P < 0.001$ ). Each region also exhibited greater responses to stimuli (both preferred and nonpreferred) presented in the contralateral versus ipsilateral hemifield (Fig. 3B; LOC:  $t_7 = 7.83$ ,  $P < 0.001$ ; PPA:  $t_7 = 6.58$ ,  $P < 0.001$ ; FFA:  $t_7 = 4.59$ ,  $P = 0.003$ ; EBA:  $t_7 = 11.57$ ,  $P < 0.001$ ; MT+:  $t_7 = 6.25$ ,  $P < 0.001$ ; IPS:  $t_7 = 6.74$ ,  $P < 0.001$ ; EVC:  $t_7 = 9.02$ ,  $P < 0.001$ ). In Experiment 2, when stimuli were arranged above and below fixation instead of to the left and right, mean responses showed no systematic location bias (Fig. 3C; as expected given that ROIs averaged across upper and lower visual fields within each hemisphere), and category-selective responses were preserved (data not shown). Mean response magnitude did not differ as a function of spatiotopic location (Fig. 3D), which is expected given that standard univariate analyses are not likely to be particularly well suited for comparisons of spatiotopic reference frames. Multivariate pattern analyses, on the other hand, allow us to test for more subtle differences and more directly compare whether spatiotopic stimulus location is represented independently of eye position.

### Correlation Patterns

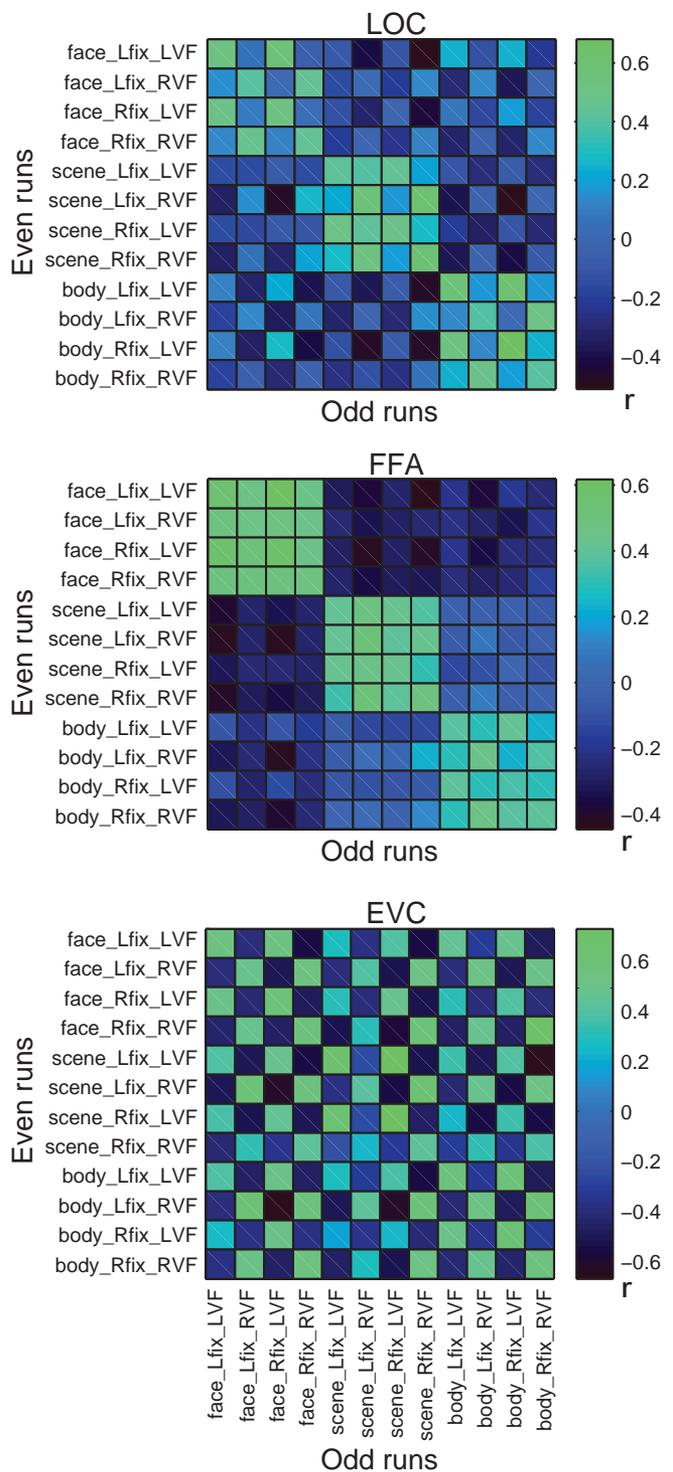
The MVPA analysis was conducted in 3 stages: creation of a correlation matrix, pooling across cells of the matrix to calculate same and different category and same and different location correlations, and finally, quantifying the amount of each type of information by calculating the difference of these “same” and “different” correlations (Haxby et al. 2001). Each of these analysis stages was done separately for each subject and region. Group data and within-subjects random-effects statistics from the final 2 stages are presented for each region and experiment below. In addition, to illustrate the analysis process and give a sense of the raw data, the full workflow of data from a sample subject’s left LOC is illustrated in Figure 2, and results from the correlation stage are shown for a few sample regions in Figure 4. The  $12 \times 12$  correlation matrix illustrates correlation strengths between every possible pair of conditions (comparing each condition in the odd data set to each condition in the even data set).

LOC (Fig. 4, top) was chosen as an example region because of its intermediate location in the visual stream, and the fact

that it responds comparably to various types of object categories. Clear clustering is seen by category, correlations are higher between faces and other blocks of faces than between faces and scenes or bodies. This indicates that while LOC responds to all categories (as evidenced by mean response magnitude), the voxelwise pattern of activation can successfully differentiate between responses to faces, houses, and bodies. In addition to object category information, the correlation matrix also reveals location information. Because of the order of conditions in the matrix, image on the left (left fixation), image on the right (left fixation), image on the left (right fixation), image on the right (right fixation), retinotopic position alternates with each cell. Thus, the checkerboard-like correlation pattern is indicative of retinotopic location information. While the LOC appears to exhibit a mixture of category and retinotopic location information, the FFA (Fig. 4, middle) seems more biased toward category information, particularly when discriminating between faces and scenes, whereas EVC (Fig. 4, bottom) appears to contain primarily retinotopic location information. These patterns are investigated more quantitatively in the sections below.

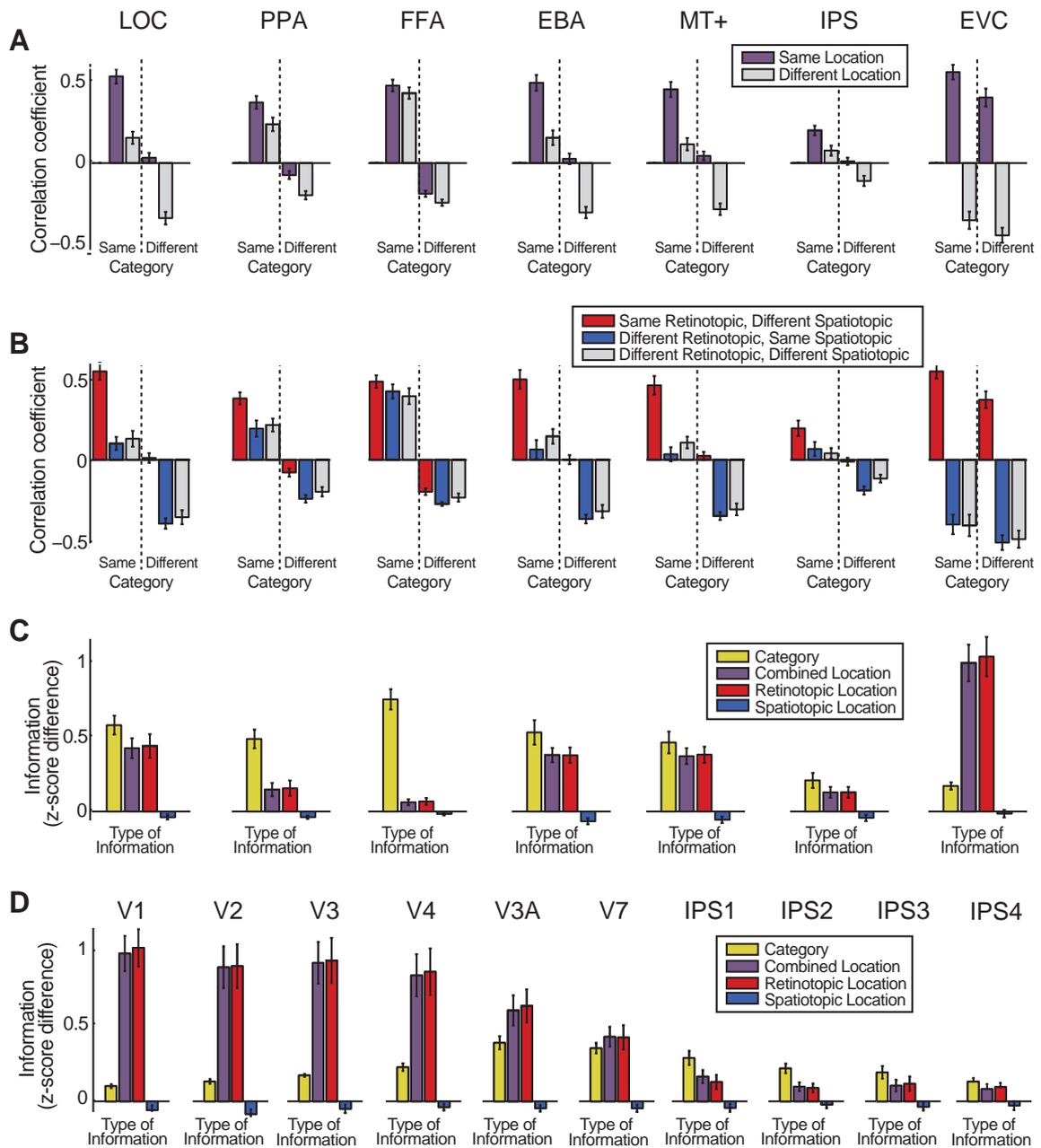
#### All Regions Contain Information about Object Category and Location

To determine whether information about object category and location is represented in ventral, dorsal, and early visual regions, correlations were pooled across cells according to which comparisons were same category, different category, same location, and different location. Because we measured the correlations between every possible pair of conditions, some correlations represent comparisons between conditions sharing the same fixation location (same-fix) and others represent comparisons across different fixation locations (different-fix). The same-fix comparisons allow us to replicate the Schwarzlose et al. (2008) results, which measured category and (combined) location information without varying eye position, and to extend these results to other regions. The Schwarzlose et al. (2008) paper focused on ventral visual areas, such as the LOC, PPA, FFA, and EBA. As can be seen in Figure 5A, in all of these regions, the highest correlations were for the same category of image presented in the same combined location; correlations were weaker when the same category was presented in a different location or when location was preserved but category differed, and the weakest correlations were for comparisons differing in both category and location. In each of these regions, there were significant main effects of both same versus different category (LOC:  $F_{1,7} = 101.7$ ,  $P < 0.001$ ; PPA:  $F_{1,7} = 75.0$ ,  $P < 0.001$ ; FFA:  $F_{1,7} = 166.8$ ,  $P < 0.001$ ; EBA:  $F_{1,7} = 47.5$ ,  $P < 0.001$ ) and same versus different combined location (LOC:  $F_{1,7} = 43.4$ ,  $P < 0.001$ ; PPA:  $F_{1,7} = 11.5$ ,  $P = 0.012$ ; FFA:  $F_{1,7} = 8.47$ ,  $P = 0.023$ ; EBA:  $F_{1,7} = 64.9$ ,  $P < 0.001$ ), replicating the finding that these ventral visual areas contain both location-invariant category information and category-invariant location information (Schwarzlose et al. 2008). Additional analyses confirmed location and category invariance (tolerance) by finding significant category information even when location differed and significant location information across differences in category (data not shown). The same pattern held for the dorsal MT+ and IPS regions, although correlations overall were weaker in IPS, with significant main effects of category (MT+:  $F_{1,7} = 50.1$ ,  $P < 0.001$ ; IPS:  $F_{1,7} = 17.0$ ,  $P = 0.004$ ) and combined location (MT+:  $F_{1,7} = 55.1$ ,  $P < 0.001$ ; IPS:  $F_{1,7} = 13.1$ ,  $P = 0.008$ ).



**Figure 4.** Example correlation matrices. Correlations are shown for each of the 144 comparisons (even runs vs. odd runs for each of the 12 conditions) from Experiment 1 for a few sample regions. LOC exhibits a combination of category and retinotopic location information, FFA exhibits primarily category information and EVC primarily retinotopic location information.  $N = 8$ .

In contrast, EVC favored location information (combined location information:  $F_{1,7} = 80.2$ ,  $P < 0.001$ ). Nonetheless, this region also contained significant category information ( $F_{1,7} = 48.9$ ,  $P < 0.001$ ). These results confirm that all of the regions examined contain information about both object category and object location, when eye position is preserved.



**Figure 5.** Results, Experiment 1. (A,B) Same and different category and location correlations for each region. (A) Comparisons for conditions sharing the same eye position; location is the combined retinotopic and spatiotopic locations. All regions show main effects of category and location. (B) Comparisons for conditions differing in eye position, where retinotopic and spatiotopic locations can be dissociated. In addition to the category effect, all regions exhibit greater correlations when retinotopic location is preserved. (C) The amount of information about each type of information was calculated by subtracting “same” minus “different” correlations for that type of information. (D) Same as C, for regions defined using retinotopic mapping. In all regions, significant category and location information was present, though the magnitudes varied. All regions exhibited robust retinotopic information and no spatiotopic information. Error bars are standard error of the mean.  $N = 8$ .

### Is the Location Information Retinotopic, Spatiotopic, or Both?

The more novel question, of course, is how the pattern of information is represented across different eye positions, when retinotopic and spatiotopic positions are unconfounded. In the different-fix comparisons (Fig. 5B), higher correlations were again found between conditions sharing the same versus different category, as expected. However, the pattern of location information was quite striking. As hinted by the correlation matrix, correlations were highest when retinotopic location was

preserved, when stimuli shared the same retinotopic but different spatiotopic positions. Crucially, however, as Figure 5B illustrates, when stimuli shared the same spatiotopic but different retinotopic position, correlations were no higher than if the stimuli differed in both retinotopic and spatiotopic positions.

To quantify these effects, we calculated the amount of retinotopic and spatiotopic information (Fig. 5C; retinotopic information = “same retinotopic, different spatiotopic” minus “different retinotopic, different spatiotopic”; spatiotopic information = “different retinotopic, same spatiotopic” minus “different retinotopic, different spatiotopic”). A 7 (region: LOC,

PPA, FFA, EBA, MT+, IPS, and EVC)  $\times$  2 (type of information: retinotopic and spatiotopic) ANOVA revealed significant main effects of region ( $F_{6,42} = 33.8, P < 0.001$ ) and type of information ( $F_{1,7} = 99.1, P < 0.001$ ) as well as a region  $\times$  type of information interaction ( $F_{6,42} = 36.2, P < 0.001$ ). Post hoc  $t$ -tests revealed significant information about retinotopic location in every region tested (LOC:  $t_7 = 5.54, P = 0.001$ ; PPA:  $t_7 = 3.06, P = 0.018$ ; FFA:  $t_7 = 2.99, P = 0.020$ ; EBA:  $t_7 = 7.27, P < 0.001$ ; MT+:  $t_7 = 6.99, P < 0.001$ ; IPS:  $t_7 = 3.52, P = 0.010$ ; EVC:  $t_7 = 7.88, P < 0.001$ ), although the magnitude clearly varied with region. Critically, in none of the regions was spatiotopic information present; if anything, correlations were weaker when spatiotopic position was preserved than when both spatiotopic and retinotopic position differed (LOC:  $t_7 = -2.04, P = 0.081$ ; PPA:  $t_7 = -3.22, P = 0.015$ ; FFA:  $t_7 = -1.63, P = 0.147$ ; EBA:  $t_7 = -3.12, P = 0.017$ ; MT+:  $t_7 = -2.49, P = 0.042$ ; IPS:  $t_7 = -1.93, P = 0.094$ ; EVC:  $t_7 = -0.57, P = 0.585$ ). A direct comparison of retinotopic to spatiotopic information also revealed significantly more retinotopic than spatiotopic information in every region tested (LOC:  $t_7 = 7.35, P < 0.001$ ; PPA:  $t_7 = 4.02, P = 0.005$ ; FFA:  $t_7 = 3.92, P = 0.006$ ; EBA:  $t_7 = 9.33, P < 0.001$ ; MT+:  $t_7 = 9.81, P < 0.001$ ; IPS:  $t_7 = 3.55, P = 0.009$ ; EVC:  $t_7 = 8.43, P < 0.001$ ).

This pattern was also present across all of the topographically defined occipital and parietal regions (Fig. 5D). Going from early to higher visual areas, location information appears to progressively decrease while category information grows, as expected. The IPS regions all have lower magnitudes of information overall, as noted above, but clearly contain both category and location information. Crucially, in all of these regions, the location information is purely retinotopic (Supplementary Material). The fact that this pattern of retinotopic but not spatiotopic information was so robust across multiple regions (ventral and dorsal, defined with both functional localizers and retinotopic mapping) attests to the strength of this retinotopic advantage. Furthermore, the magnitude of retinotopic information in each of the regions was essentially equivalent to the magnitude of the combined location information in that region, suggesting that when retinotopic and spatiotopic information are confounded when eye position does not vary, this combined location information is driven entirely by the retinotopic contribution.

The lack of spatiotopic information in higher lever areas is somewhat surprising, given that spatiotopic effects are thought to increase for more complex stimuli and later brain areas (Andersen et al. 1997; Melcher and Colby 2008; Wurtz 2008). Spatiotopic or head-centered effects have been previously associated with dorsal areas IPS (Galletti et al. 1993; Duhamel et al. 1997; Snyder et al. 1998; Sereno and Huang 2006; Rawley and Constantinidis 2010) and area MT/V5 (Melcher 2005; d'Avossa et al. 2007; Ong et al. 2009; Crespi et al. 2011), although these effects are still actively disputed (Gardner et al. 2008; Knappen et al. 2009). In the ventral areas, studies that have investigated location information often report broader-scale location (Hemond et al. 2007), larger receptive fields (Gross et al. 1972; MacEvoy and Epstein 2007), and position tolerance, that is, the ability to represent object identity across changes in location (Ito et al. 1995; Grill-Spector and Malach 2001; Li and DiCarlo 2008; Schwarzlose et al. 2008; although see Kravitz et al. 2010; Carlson et al. 2011). Accordingly, we found evidence for position-invariant category information in these areas, as well as category-invariant location

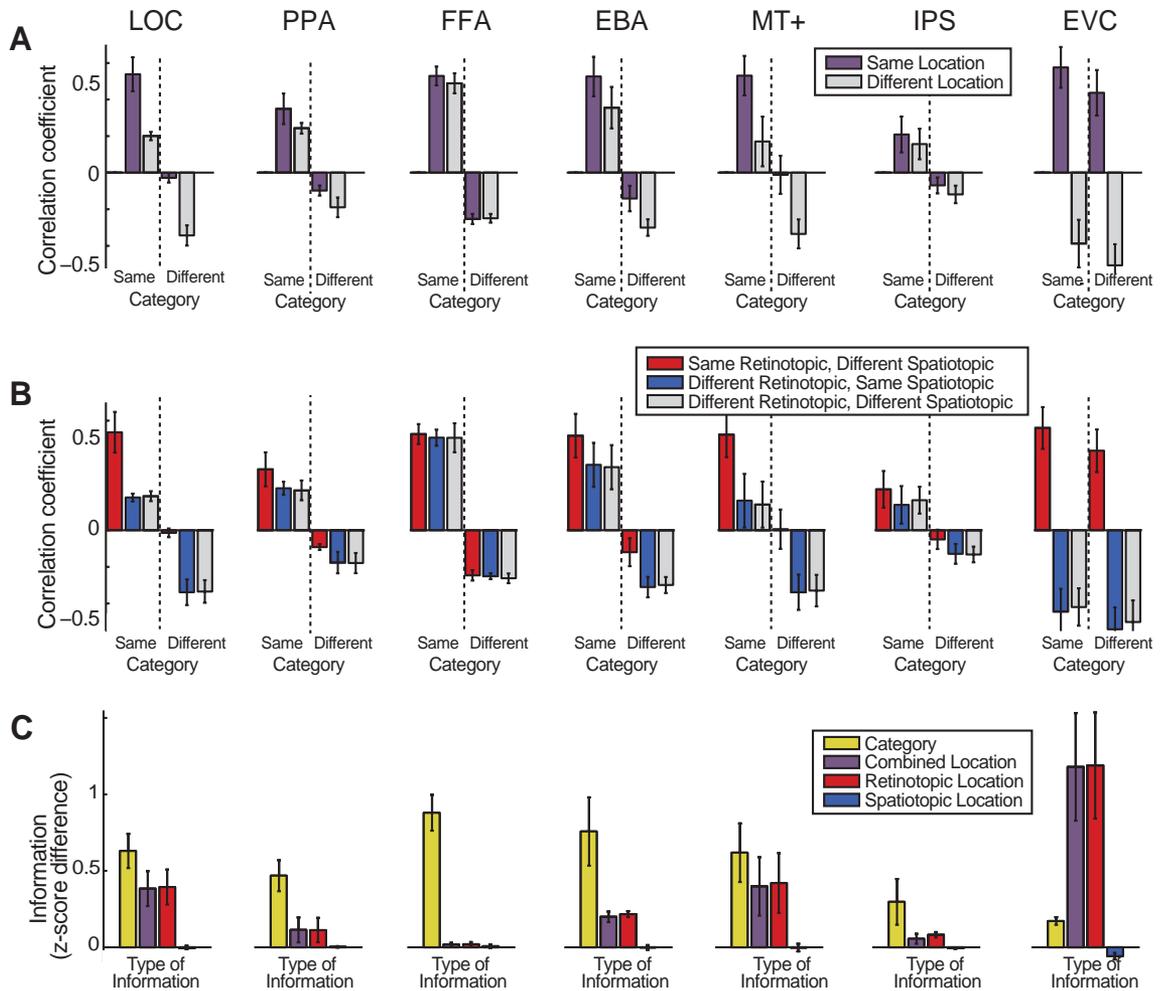
information. However, unlike a previous report (McKyton and Zohary 2007), we did not find any spatiotopic information in LOC nor in any other category-selective ventral regions. These results could not be attributed to a difference in the number of comparisons for spatiotopic and retinotopic conditions: Although twice as many correlation pairs shared the same retinotopic than spatiotopic position (36 vs. 18 of 144 possible pairs), we conducted an additional analysis matching for power by excluding half of the retinotopic comparisons and found the same pattern of results (data not shown). We also verified this pattern of results using an alternative MVPA technique (support vector machines: SVM; see Supplementary Material), since assumptions and sensitivity to different types of information can vary across MVPA methods (Pereira and Botvinick 2011).

### ***Is the Lack of Spatiotopic Information due to Interhemispheric Confounds?***

Experiment 1 demonstrated a lack of spatiotopic information across all the regions tested. However, due to the stimulus design, all spatiotopic comparisons involved comparing across hemispheres, whereas the retinotopic comparisons were all necessarily within hemisphere. To ensure that the lack of spatiotopic effects was not due to the crossing of hemispheres or vertical meridians, in Experiment 2, we conducted the same task but with the stimuli arranged vertically, such that all conditions involved equivalent stimulation in both hemispheres. The pattern of results (Fig. 6) was similar to Experiment 1. Category and combined location information was present in every region, with the exception of the FFA, which exhibited only negligible location information; the FFA had the weakest location information in Experiment 1 as well. Importantly, the location information was again exclusively retinotopic. As in Experiment 1, there were significant main effects of region ( $F_{6,18} = 7.59, P = 0.036$ ) and type of information ( $F_{1,3} = 13.81, P = 0.034$ ) and a region  $\times$  type of information interaction ( $F_{6,18} = 7.60, P = 0.033$ ). With the exception of the FFA, in each region 4 of 4 subjects exhibited both retinotopic information and greater retinotopic than spatiotopic information; in the FFA, 3 of 4 subjects exhibited this pattern.

### ***Does Location Information Become More Spatiotopic with a Spatiotopic Task?***

One could argue that spatiotopic information was not present in the first 2 experiments because the spatiotopic reference frame was not relevant for the task. The subjects' task was to detect repetitions in the identity of the stimulus, and stimulus location was always blocked. While retinotopic location was also irrelevant for the task, prior work has demonstrated that even when irrelevant to the task, information may be maintained in the native retinotopic coordinate system (Gardner et al. 2008; Golomb et al. 2008; Golomb et al. 2010), whereas spatiotopic representations are only formed when compelling spatiotopic task demands are present (Golomb et al. 2008). Furthermore, a recent study has reported spatiotopic information in several extrastriate areas (including MT, V4, and LO) that is revealed only in a passive condition when spatial attention is available to be allocated to the stimuli (Crespi et al. 2011). Thus, it is possible that the regions tested here may contain information about spatiotopic position but perhaps only when task relevant or attended. In Experiment 3,



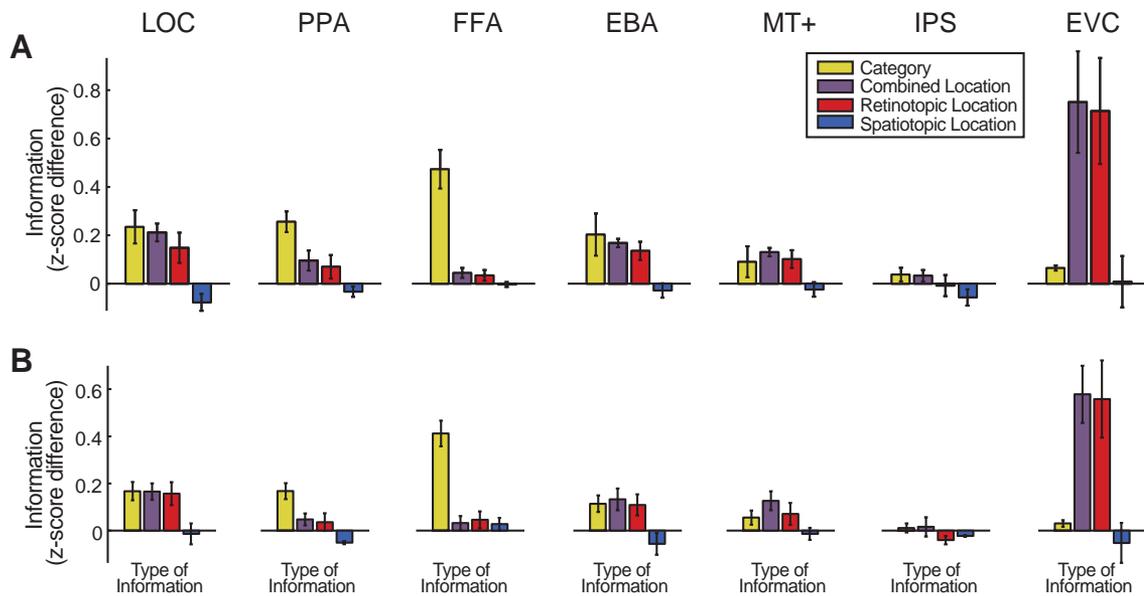
**Figure 6.** Results, Experiment 2, vertical stimulus arrangement. (A,B), Same and different category and location correlations for each region. (A) Comparisons for conditions sharing the same eye position; location is the combined retinotopic and spatiotopic locations. All regions show main effects of category and location. (B) Comparisons for conditions differing in eye position, where retinotopic and spatiotopic locations can be dissociated. In addition to the category effect, all regions exhibit greater correlations when retinotopic location is preserved. (C), The amount of information about each type of information was calculated by subtracting “same” minus “different” correlations for that type of information. As in Experiment 1, all location information was retinotopic. Error bars are standard error of the mean.  $N = 4$ .

we tested this by having subjects explicitly attend to the spatiotopic location of the stimuli. We switched to an event-related design where stimulus category and location varied pseudorandomly from trial to trial. Each subject completed 2 tasks in separate scanning sessions: a category task and a spatiotopic location task. In the category task, subjects reported the category of each stimulus (face or scene) as it appeared; in the location task, subjects reported the spatiotopic location of the stimulus (left, middle, or right position). Although we did not replicate the passive viewing conditions of the Crespi et al. (2011) task, our task is arguably an even stronger test of whether spatiotopic coding emerges only for attended stimuli.

The pattern of data was remarkably similar across tasks (Fig. 7). The results were also quite similar to the first 2 experiments, with a few exceptions. Overall, the amount of both category and location information seemed somewhat decreased compared with the earlier experiments; this is probably due to the fact that stimulus exposure was shorter in the event-related compared with block design. This effect is most obvious in the IPS, which produced the smallest effects in

Experiments 1 and 2 and did not exhibit any location or category information in either event-related task in Experiment 3.

Despite these few differences across experiments, the critical question, whether location information is retinotopic, spatiotopic, or both, again revealed a purely retinotopic answer. Not only was this effect replicated with an unblocked event-related design but even when the task was to report spatiotopic location, none of the regions tested exhibited spatiotopic information. A 2 (task)  $\times$  7 (region)  $\times$  2 (type of information: retinotopic/spatiotopic) revealed a significant main effect of type of information ( $F_{1,3} = 13.07$ ,  $P = 0.036$ ) that was modulated by region ( $F_{6,18} = 16.26$ ,  $P = 0.023$ ) but not by task ( $F_{1,3} = 1.14$ ,  $P = 0.364$ ). Importantly, none of the interactions with task were significant (task  $\times$  region:  $F_{6,18} = 2.75$ ,  $P = 0.125$ ; task  $\times$  type of information:  $F_{1,3} = 1.14$ ,  $P = 0.364$ ; task  $\times$  region  $\times$  type of information:  $F < 1$ ) nor there was a significant main effect of task ( $F < 1$ ). Thus, when these regions contain information about stimulus location, that information reflects the retinotopic coordinate frame, even when location is either completely irrelevant to the task or when spatiotopic, not retinotopic, responses are required.

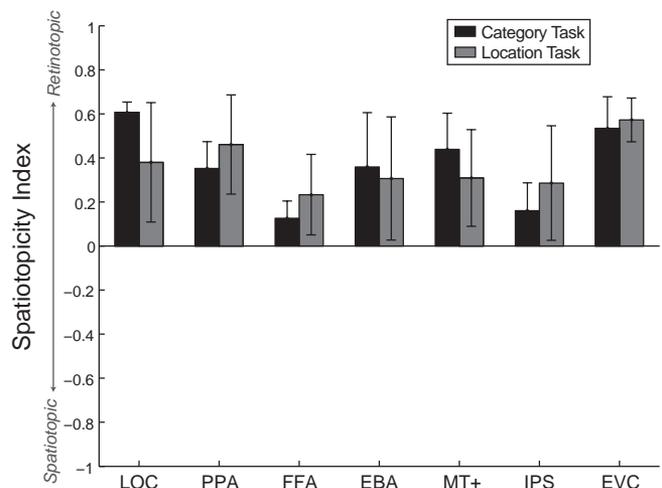


**Figure 7.** Results, Experiment 3, task manipulations. (A,B) The amount of information about each type of information was calculated by subtracting “same” minus “different” correlations for that type of information. (A) Category task. (B) Spatiotopic location task. In both tasks, location information was purely retinotopic. Error bars are standard error of the mean.  $N = 4$ .

As a further test, we analyzed our data using an additional univariate analysis method to allow for more direct comparison with prior fMRI reference frame studies (d’Avossa et al. 2007; Gardner et al. 2008; Crespi et al. 2011). We calculated a “spatiotopicity index” for each region, reflecting whether the BOLD responses better matched retinotopic or spatiotopic predictions (Fig. 8). Consistent with our MVPA results, each region behaved retinotopically, even when attention was explicitly allocated to the spatiotopic location of the stimuli. Thus, while it is possible that spatiotopic responses may be revealed in a passive viewing condition (Crespi et al. 2011), it is not likely that this difference is due to attention being allocated to the stimulus location. Our data are more consistent with the idea that visual areas (including higher level areas not previously examined) are fundamentally retinotopic in nature.

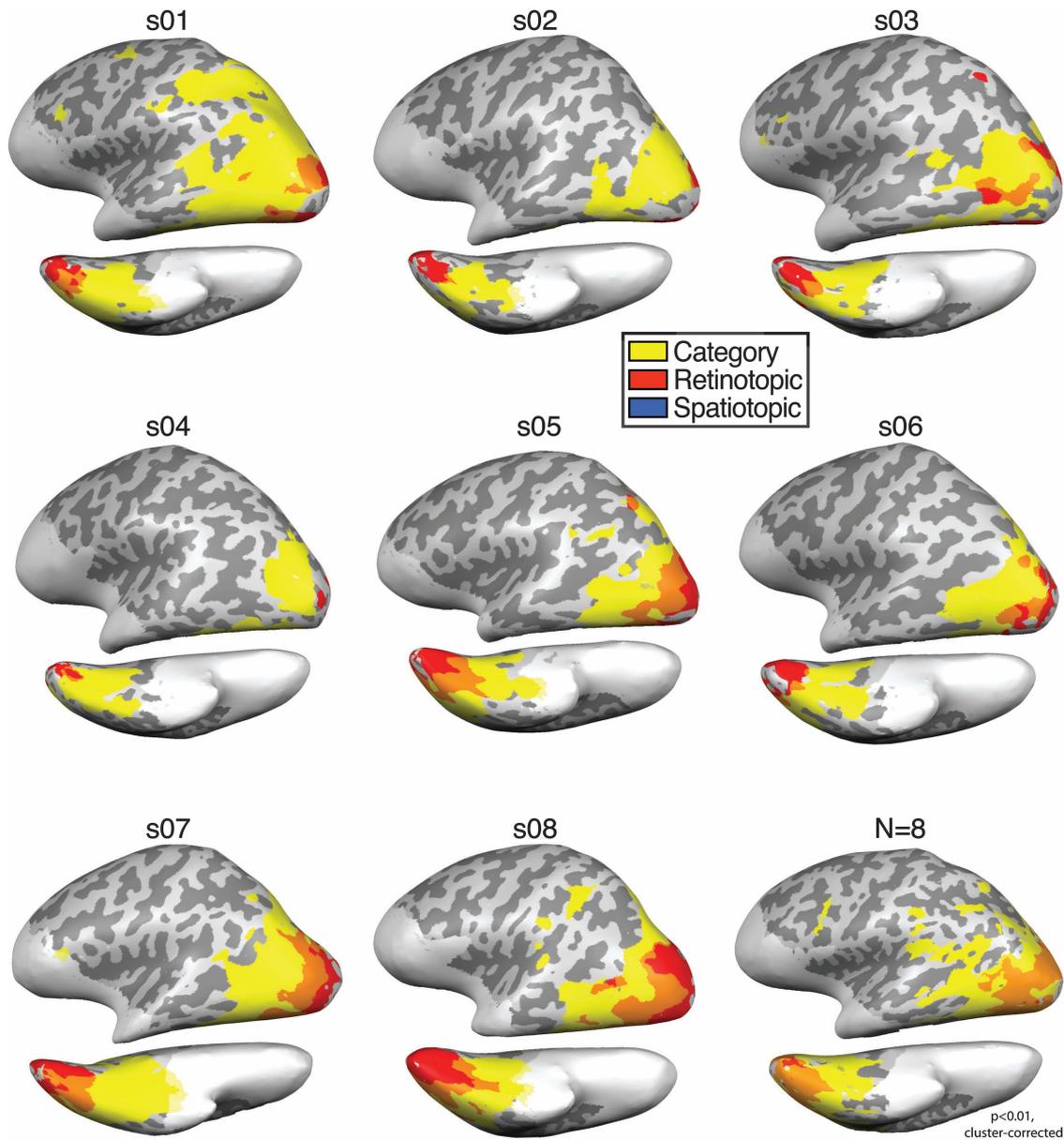
### Is There Spatiotopic Information Elsewhere in the Brain?

While none of the ventral, dorsal, or early visual regions tested exhibited any spatiotopic information; the ability to locate objects in spatiotopic coordinates is clearly a function we need to navigate our visual worlds. If explicit spatiotopic information is not present in these same regions that process visual objects, then where does the spatiotopic information come from? In the following sections, we explore 2 possibilities: that there is explicit spatiotopic information present elsewhere in the brain and/or that spatiotopic position is computed indirectly based on retinotopic position. To test whether spatiotopic information might be present outside the regions we tested, we conducted an MVPA “searchlight analysis” (Kriegeskorte et al. 2006). Instead of looking for voxelwise patterns within an a priori ROI, we used a moving “searchlight” ROI to systematically explore the brain and identify clusters with significant category, retinotopic, and spatiotopic location information. The searchlight was conducted across all voxels covered by our slice prescription, which included full coverage of occipital and parietal cortices and posterior coverage of temporal and frontal cortices but did not cover regions such as



**Figure 8.** Spatiotopicity index. A spatiotopicity (reference frame) index was calculated for each visual area for the category and location tasks of Experiment 3. The index is a univariate measure of whether the BOLD response is more consistent with retinotopic (positive index) or spatiotopic (negative index) predictions. All regions exhibited retinotopic indices, regardless of attentional condition. Compare with Crespi et al. (Figure 9) and Gardner et al. (Figure 5). Error bars are standard error of the mean.  $N = 4$ .

the hippocampus, anterior temporal, or prefrontal cortex. Results from Experiment 1 are plotted in Figure 9 for individual subjects, along with the group average. The searchlight patterns are quite consistent across subjects: Category information was found throughout posterior areas, including nearly all of occipital cortex and extending into posterior temporal and parietal regions, consistent with the ROI data. Also as expected, retinotopic location was represented robustly across visual cortex. At proper significance thresholds ( $P < 0.01$ , cluster corrected), no spatiotopic clusters were present in the group maps. Spatiotopic information was also notably absent in individual subjects’ maps, arguing against the



**Figure 9.** Searchlight analysis. Results from searchlight analysis projected onto inflated brains. Maps of significant category information (yellow), retinotopic information (red), and spatiotopic information (blue) overlaid for each individual subject, along with the group average. Anterior temporal and frontal areas marked in light gray indicate that these areas fell outside the scan coverage and thus could not be included in the searchlight. Statistical thresholds for the individual maps were calculated from permutation tests; the group average maps use standard parametric tests. All maps are thresholded at  $P < 0.01$ , cluster corrected. In all subjects, visual cortex is clearly dominated by category and retinotopic information; no spatiotopic clusters meet significance.

possibility that spatiotopic information might be present but in variable locations across subjects.

Although we did not find any statistically reliable spatiotopic clusters, in the interest of thoroughness, we conducted additional searchlight analyses on the group maps with more liberal statistical thresholds, as a purely exploratory analysis to look for any hints of smaller or weaker spatiotopic clusters. At more liberal thresholds ( $P < 0.05$ , uncorrected), a small cluster emerged bilaterally in an eccentric portion of EVC. This area probably reflects stimulation in the far periphery (perhaps the outline of the scanner bore) that did not change much across eye movements. When we further expanded the search to explore only same category location information, some small spatiotopic clusters appeared in left posterior parietal cortex.

Spatiotopic parietal involvement would be consistent with previous reports (Galletti et al. 1993; Duhamel et al. 1997; Snyder et al. 1998) and warrants further investigation. It should be emphasized, however, that these clusters were not statistically reliable in our study, and even if these areas do represent spatiotopic information, it is quite weak, whereas the retinotopic information represented in these same areas appears much more stable.

Interestingly, in a variation of Experiment 3 where subjects directly reported spatiotopic position with a left, middle, or right button press, the searchlight analysis revealed robust clusters of spatiotopic information centered on left motor cortex (Supplementary Material). However, this supplemental experiment confounded spatiotopic position with motor

response,, subjects responded with a different finger on their right hand for each spatiotopic position,, and thus, the searchlight was likely detecting a motor code. Nonetheless, the results suggest that this searchlight analysis should have the power to detect true spatiotopic information were it present.

## Discussion

Although we must be able to represent object location independent of eye position to function in the world, the current results suggest that explicit spatiotopic representations are not present in human visual cortex, even in higher level areas. We found information about both object category and object location in all ventral and dorsal regions. Crucially, however, all location information was retinotopic. Across a series of experiments, we demonstrated that the lack of spatiotopic information was not due to a specific stimulus configuration or task design. Retinotopic location information was preserved even when objects came from different categories, and it was present when location was irrelevant to the task. On the other hand, spatiotopic representations were never present, even when the task was to report the object's spatiotopic position. These results are consistent with previous work in early visual regions (Gardner et al. 2008; Golomb et al. 2010) but are somewhat surprising in light of the idea that spatiotopic effects are more common in later brain areas (Andersen et al. 1997; Melcher and Colby 2008; Wurtz 2008).

Clearly, spatiotopic position must be represented at some level in order for us to function in the world, yet we found only retinotopic information. We explored a few possible explanations for this puzzle. The first is that spatiotopic information might exist elsewhere in the brain, outside our specific ROIs. However, our searchlight analysis did not convincingly reveal any additional spatiotopic regions, either in the individual subjects or in the group average. Second, it is possible that spatiotopic information is present at a finer or more distributed scale than can be detected with our techniques. Although MVPA can detect information on a finer grain than fMRI voxel size (Kamitani and Tong 2005), the extent of its sensitivity is still debated (Op de Beeck 2010; Freeman et al. 2011), and we cannot rule out the presence of spatiotopic information on a much finer scale. However, even if spatiotopic information were represented in such a manner, it is still notable that the grain of spatiotopic information would be so much smaller than that of retinotopic information. Finally, a third explanation, dealt with in more detail below, is that spatiotopic information need not be represented explicitly (i.e., in a way that can be linearly decoded) but could instead be represented implicitly and recalculated with each eye movement, based on the current retinotopic position of the object plus eye position.

### *The Native Retinotopic System*

The idea that spatiotopic representations may not be created as automatically or instantaneously as retinotopic ones has become an increasingly prevalent theme in human neuroscience (Gardner et al. 2008; Golomb et al. 2008; Cavanagh et al. 2010; Golomb et al. 2010; Mathot and Theeuwes 2010), supported by physiological and computational work (for review, see Cohen and Andersen 2002), as well as evidence

that infants initially behave according to retinotopic expectations and take several months to develop spatiotopic abilities (Gilmore and Johnson 1997). Humans are clearly capable of spatiotopic tasks, such as double-step saccades (Mays and Sparks 1980) and updating/remapping to maintain spatiotopic references (Duhamel et al. 1992; Medendorp et al. 2003; Merriam et al. 2003; Melcher 2007). However, spatiotopic abilities do not necessarily require explicit spatiotopic representations. Information need not actually reside in spatiotopic coordinates; instead, information could reside in natively retinotopic maps that are updated with each eye movement to reflect current spatiotopic position. Golomb et al. (2008) demonstrated that attention can be maintained at a spatiotopic location, but the native representations are retinotopic; spatiotopic representations are only created when task relevant, and the dynamic process leaves behind a "retinotopic attentional trace" with each saccade.

Similarly, the transformation from egocentric (body centered) to allocentric (world centered) processing is thought to involve a process of constant updating (Wang and Spelke 2000), and there is evidence of a noninstantaneous transition from retinal to object-centered representations as well (Crowe et al. 2008). In the current experiments, body/head position never varies nor do the positions of other stationary objects (the placeholders, projection screen, magnet, etc.). Thus, spatiotopic representations here could include contributions from any or all of these nonretinotopic coordinate frames,, yet the retinotopic representations still dominate. As noted above, this does not mean that these nonretinotopic reference frames are never used, just that they are not the native frame of spatial coding. In other words, nonretinotopic forms of representation,, body-centered, object-centered, world-centered,, could be calculated from a combination of retinotopic information and other input (Cohen and Andersen 2002) without being explicitly represented in visual areas. Perhaps a more accurate description of parietal involvement (long thought to be a source of extraretinotopic representations) would be not that it contains explicit maps of spatiotopic location per se but rather helps transform the retinotopic maps by conveying implicit spatiotopic information to downstream effector regions. This could explain why parietal damage results in disruptions of spatiotopic positional sense (i.e., neglect: Pisella and Mattingley 2004) and elimination of spatiotopic, but not retinotopic, inhibition of return (Sapir et al. 2004; van Koningsbruggen et al. 2009). Not having permanent explicit maps of all these different reference frames might reduce redundancy, maximize storage capacity, and allow for more flexible representations, in much the same way prefrontal cortex is thought to reuse the same neurons for different tasks (Cromer et al. 2010). However, it remains possible that explicit spatiotopic maps might be present downstream in areas such as premotor cortex (Graziano et al. 1994) or the hippocampus, a region which was outside our slice coverage and is not primarily visually driven, but where place cells have been shown to represent a given location independent of eye position (O'Keefe and Dostrovsky 1971; Ekstrom et al. 2003; Iglói et al. 2010).

### *A Spatiotopic Solution: The Role of Eye Position and Dynamic Eye Movements*

If the visual system does not support explicit spatiotopic representations, and location representations instead rely on

retinotopic input, how is this native retinotopic information transformed to accommodate spatiotopic behavior? A likely answer is that the brain uses information about eye position. Eye position gain fields have been reported throughout parietal (Andersen et al. 1985, 1993; Bremmer et al. 1997; Williams and Smith 2010) and earlier visual cortex (Galletti and Battaglini 1989; Weyand and Malpeli 1993; Guo and Li 1997; Trotter and Celebrini 1999; Bremmer 2000; DeSouza et al. 2002; Andersson et al. 2007) and are prominent in computational models of visual stability (Zipser and Andersen 1988; Andersen et al. 1990; Cassanello and Ferrera 2007). If both retinotopic and eye position information are present in the same cortical regions, these basic elements could be combined to implicitly represent spatiotopic position without the need for explicit spatiotopic visual maps. Indeed, representations of both retinotopic and eye position have been reported in human EVC (Merriam et al. 2008), and an exploratory analysis of our data suggests that eye position information may be present alongside the retinotopic information in many of the higher level areas as well (Supplementary Material).

If retinotopic and eye position information are present throughout visual cortex, what determines whether and when they are integrated to form a spatiotopic percept? Spatiotopic effects are present in some behavioral tasks (Davidson et al. 1973; McRae et al. 1987; Melcher and Morrone 2003; Burr et al. 2007; Pertzov et al. 2010) but weaker or absent in others (Irwin 1992; Abrams and Pratt 2000; Afraz and Cavanagh 2009; Knapen et al. 2009; McKyton et al. 2009), and some evidence suggests that the presence and magnitude of spatiotopic effects depend on their task relevance (Golomb et al. 2008; Rawley and Constantinidis 2010; Crespi et al. 2011). In the current task, we did not find an increase in spatiotopic information when spatiotopic location was task relevant; however, it is possible that spatiotopic behavior is only truly relevant when information must be maintained across an eye movement (our design did not involve eye movements during the trial), or that spatiotopic representations are more important for action than perception tasks (Pertzov et al. 2011) or for moving stimuli (e.g., Crespi et al. 2011). The eye movement difference could partially explain the discrepancy between our LOC findings and the spatiotopic adaptation findings of McKyton and Zohary (2007); object location might always be represented in retinotopic coordinates, but if the retinotopic representation is dynamically remapped with each saccade, this could manifest as spatiotopic adaptation. Similarly, the spatiotopic attentional facilitation reported in Golomb et al. (2010) does not mean that occipital cortex contains spatiotopic representations per se but rather that retinotopic representations can update following a saccade (Medendorp et al. 2003; Merriam et al. 2007). It has been suggested that an eye movement signal is required for successful remapping of position (Stevens et al. 1976; Wurtz 2008), that spatiotopic receptive field shifts and other dynamic changes occur specifically during the peri-saccadic time window (Hamker et al. 2008), and that scene-selective cortex adapts when the eyes move across a scene but not when the scene moves the same distance in the background with the eyes fixed (Golomb et al. 2011). These findings suggest that spatiotopic representations may be constructed from retinotopic position and eye position, but whether and how well this process is executed can be influenced by top-down task relevance and dynamic eye movement signals.

### **Combining Object Identity and Location Information**

For real-world vision, we often need to know not just about a spatial location but about the position of a particular object, requiring that information about object location be bound to its identity. At what point is this information combined? In particular, does the merging of identity and location information happen before or after the transformation from retinotopic to spatiotopic representation? The present results demonstrate that retinotopic, but not spatiotopic, position is represented explicitly in the same regions as object category, although our MVPA analysis alone cannot address whether this information is actually used (Williams et al. 2007). Intriguingly, a recent behavioral study from our group suggests that object identity is bound to retinotopic, not spatiotopic, representations (Tower-Richardi et al. 2011), further underscoring the idea that despite our subjective spatiotopic experience, most visual processing actually occurs in retinotopic coordinates. We suggest that properties such as object category, retinotopic location, and eye position are more basic units of perceptual processing, whereas spatiotopic position is an emergent property that must be continually recalculated.

### **Supplementary Material**

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

### **Funding**

National Institutes of Health (grants R01-EY13455 to N.K. and F32-EY020157 to J.D.G.); Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research, Massachusetts Institute of Technology.

### **Notes**

We thank C. Triantafyllou, S. Shannon, and S. Arnold for technical support and A. Afraz, J. DiCarlo, P. J. Hsieh, T. Konkle, A. Leber, and E. Vul for helpful discussion and suggestions. *Conflict of Interest*: None declared.

### **References**

- Abrams RA, Pratt J. 2000. Oculocentric coding of inhibited eye movements to recently attended locations. *J Exp Psychol Hum Percept Perform.* 26:776-788.
- Afraz A, Cavanagh P. 2009. The gender-specific face aftereffect is based in retinotopic not spatiotopic coordinates across several natural image transformations. *J Vis.* 9:10.1-10.17.
- Andersen RA, Bracewell RM, Barash S, Gnadt JW, Fogassi L. 1990. Eye position effects on visual, memory, and saccade-related activity in areas LIP and 7a of macaque. *J Neurosci.* 10:1176-1196.
- Andersen RA, Essick GK, Siegel RM. 1985. Encoding of spatial location by posterior parietal neurons. *Science.* 230:456-458.
- Andersson F, Joliot M, Percey G, Petit L. 2007. Eye position-dependent activity in the primary visual area as revealed by fMRI. *Hum Brain Mapp.* 28:673-680.
- Andersen RA, Snyder LH, Bradley DC, Xing J. 1997. Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annu Rev Neurosci.* 20:303-330.
- Andersen RA, Snyder LH, Li CS, Stricanne B. 1993. Coordinate transformations in the representation of spatial information. *Curr Opin Neurobiol.* 3:171-176.
- Arcaro MJ, McMains SA, Singer BD, Kastner S. 2009. Retinotopic organization of human ventral visual cortex. *J Neurosci.* 29:10638-10652.

- Brainard DH. 1997. The psychophysics toolbox. *Spat Vis.* 10:433-436.
- Bremmer F. 2000. Eye position effects in macaque area V4. *Neuroreport.* 11:1277-1283.
- Bremmer F, Distler C, Hoffmann KP. 1997. Eye position effects in monkey cortex. II. Pursuit- and fixation-related activity in posterior parietal areas LIP and 7A. *J Neurophysiol.* 77:962-977.
- Brewer AA, Liu J, Wade AR, Wandell BA. 2005. Visual field maps and stimulus selectivity in human ventral occipital cortex. *Nat Neurosci.* 8:1102-1109.
- Burr D, Tozzi A, Morrone MC. 2007. Neural mechanisms for timing visual events are spatially selective in real-world coordinates. *Nat Neurosci.* 10:423-425.
- Carlson T, Hogendoorn H, Fonteijn H, Verstraten FA. 2011. Spatial coding and invariance in object-selective cortex. *Cortex.* 47:14-22.
- Cassanello CR, Ferrera VP. 2007. Visual remapping by vector subtraction: analysis of multiplicative gain field models. *Neural Comput.* 19:2353-2386.
- Cavanagh P, Hunt AR, Afraz A, Rolfs M. 2010. Visual stability based on remapping of attention pointers. *Trends Cogn Sci.* 14:147-153.
- Cohen YE, Andersen RA. 2002. A common reference frame for movement plans in the posterior parietal cortex. *Nat Rev Neurosci.* 3:553-562.
- Connor CE, Brincat SL, Pasupathy A. 2007. Transformation of shape information in the ventral pathway. *Curr Opin Neurobiol.* 17:140-147.
- Crespi S, Biagi L, d'Avossa G, Burr DC, Tosetti M, Morrone MC. 2011. Spatiotopic coding of BOLD signal in human visual cortex depends on spatial attention. *PLoS One.* 6:e21661.
- Cromer JA, Roy JE, Miller EK. 2010. Representation of multiple, independent categories in the primate prefrontal cortex. *Neuron.* 66:796-807.
- Crowe DA, Averbach BB, Chafee MV. 2008. Neural ensemble decoding reveals a correlate of viewer- to object-centered spatial transformation in monkey parietal cortex. *J Neurosci.* 28:5218-5228.
- d'Avossa G, Tosetti M, Crespi S, Biagi L, Burr DC, Morrone MC. 2007. Spatiotopic selectivity of BOLD responses to visual motion in human area MT. *Nat Neurosci.* 10:249-255.
- Davidson ML, Fox MJ, Dick AO. 1973. Effect of eye movements on backward masking and perceived location. *Percept Psychophys.* 14:110-116.
- deCharms RC, Zador A. 2000. Neural representation and the cortical code. *Annu Rev Neurosci.* 23:613-647.
- DeSouza JF, Dukelow SP, Vilis T. 2002. Eye position signals modulate early dorsal and ventral visual areas. *Cereb Cortex.* 12:991-997.
- DeYoe EA, Carman GJ, Bandettini P, Glickman S, Wieser J, Cox R, Miller D, Neitz J. 1996. Mapping striate and extrastriate visual areas in human cerebral cortex. *Proc Natl Acad Sci U S A.* 93:2382-2386.
- DiCarlo JJ, Cox DD. 2007. Untangling invariant object recognition. *Trends Cogn Sci.* 11:333-341.
- DiCarlo JJ, Maunsell JH. 2003. Anterior inferotemporal neurons of monkeys engaged in object recognition can be highly sensitive to object retinal position. *J Neurophysiol.* 89:3264-3278.
- Downing PE, Jiang Y, Shuman M, Kanwisher N. 2001. A cortical area selective for visual processing of the human body. *Science.* 293:2470-2473.
- Duhamel JR, Bremmer F, BenHamed S, Graf W. 1997. Spatial invariance of visual receptive fields in parietal cortex neurons. *Nature.* 389:845-848.
- Duhamel JR, Colby CL, Goldberg ME. 1992. The updating of the representation of visual space in parietal cortex by intended eye movements. *Science.* 255:90-92.
- Ekstrom AD, Kahana MJ, Caplan JB, Fields TA, Isham EA, Newman EL, Fried I. 2003. Cellular networks underlying human spatial navigation. *Nature.* 425:184-188.
- Engel SA, Glover GH, Wandell BA. 1997. Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cereb Cortex.* 7:181-192.
- Engel SA, Rumelhart DE, Wandell BA, Lee AT, Glover GH, Chichilnisky EJ, Shadlen MN. 1994. fMRI of human visual cortex. *Nature.* 369:525.
- Epstein R, Kanwisher N. 1998. A cortical representation of the local visual environment. *Nature.* 392:598-601.
- Fischer J, Spotswood N, Whitney D. 2011. The emergence of perceived position in the visual system. *J Cogn Neurosci.* 23:119-136.
- Freeman J, Brouwer GJ, Heeger DJ, Merriam EP. 2011. Orientation decoding depends on maps, not columns. *J Neurosci.* 31:4792-4804.
- Galletti C, Battaglini PP. 1989. Gaze-dependent visual neurons in area V3A of monkey prestriate cortex. *J Neurosci.* 9:1112.
- Galletti C, Battaglini PP, Fattori P. 1993. Parietal neurons encoding spatial locations in craniotopic coordinates. *Exp Brain Res.* 96:221-229.
- Gardner JL, Merriam EP, Movshon JA, Heeger DJ. 2008. Maps of visual space in human occipital cortex are retinotopic, not spatiotopic. *J Neurosci.* 28:3988-3999.
- Gilmore RO, Johnson ME. 1997. Body-centered representations for visually-guided action emerge during early infancy. *Cognition.* 65:B1-B9.
- Golomb JD, Albrecht AR, Park S, Chun MM. 2011. Eye movements help link different views in scene-selective cortex. *Cereb Cortex.* 21:2094-2102.
- Golomb JD, Chun MM, Mazer JA. 2008. The native coordinate system of spatial attention is retinotopic. *J Neurosci.* 28:10654-10662.
- Golomb JD, Nguyen-Phuc AY, Mazer JA, McCarthy G, Chun MM. 2010. Attentional facilitation throughout human visual cortex lingers in retinotopic coordinates after eye movements. *J Neurosci.* 30:10493-10506.
- Graziano MS, Yap GS, Gross CG. 1994. Coding of visual space by premotor neurons. *Science.* 266:1054-1057.
- Grill-Spector K, Malach R. 2001. fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychol (Amst).* 107:293-321.
- Gross CG, Rocha-Miranda CE, Bender DB. 1972. Visual properties of neurons in inferotemporal cortex of the Macaque. *J Neurophysiol.* 35:96-111.
- Guo K, Li CY. 1997. Eye position-dependent activation of neurons in striate cortex of macaque. *Neuroreport.* 8:1405.
- Hamker FH, Zirnsak M, Calow D, Lappe M. 2008. The peri-saccadic perception of objects and space. *PLoS Comput Biol.* 4:e31.
- Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science.* 293:2425-2430.
- Hayworth KJ, Lescroart MD, Biederman I. 2011. Neural encoding of relative position. *J Exp Psychol Hum Percept Perform.* 37:1032-1050.
- Hemond CC, Kanwisher NG, Op de Beeck HP. 2007. A preference for contralateral stimuli in human object- and face-selective cortex. *PLoS One.* 2:e574.
- Huk AC, Dougherty RF, Heeger DJ. 2002. Retinotopy and functional subdivision of human areas MT and MST. *J Neurosci.* 22:7195-7205.
- Hung CP, Kreiman G, Poggio T, DiCarlo JJ. 2005. Fast readout of object identity from macaque inferior temporal cortex. *Science.* 310:863-866.
- Iglói K, Doeller CF, Berthoz A, Rondi-Reig L, Burgess N. 2010. Lateralized human hippocampal activity predicts navigation based on sequence or place memory. *Proc Natl Acad Sci U S A.* 107:14466-14471.
- Irwin DE. 1992. Perceiving an integrated visual world. In: Meyer DE, Kornblum S, editor. *Attention and performance XIV: synergies in experimental psychology, artificial intelligence, and cognitive neuroscience.* Cambridge (MA): The MIT Press. p. 121-142.
- Ito M, Tamura H, Fujita I, Tanaka K. 1995. Size and position invariance of neuronal responses in monkey inferotemporal cortex. *J Neurophysiol.* 73:218-226.
- Janssen P, Srivastava S, Ombelet S, Orban GA. 2008. Coding of shape and position in macaque lateral intraparietal area. *J Neurosci.* 28:6679-6690.
- Kamitani Y, Tong F. 2005. Decoding the visual and subjective contents of the human brain. *Nat Neurosci.* 8:679-685.
- Kanwisher N, McDermott J, Chun MM. 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci.* 17:4302-4311.
- Knapen T, Rolfs M, Cavanagh P. 2009. The reference frame of the motion aftereffect is retinotopic. *J Vis.* 9:16.1-16.7.
- Konen CS, Kastner S. 2008. Representation of eye movements and stimulus motion in topographically organized areas of human posterior parietal cortex. *J Neurosci.* 28:8361-8375.

- Kourtzi Z, Kanwisher N. 2000. Cortical regions involved in perceiving object shape. *J Neurosci.* 20:3310-3318.
- Kravitz DJ, Kriegeskorte N, Baker CI. 2010. High-level visual object representations are constrained by position. *Cereb Cortex.* 20:2916-2925.
- Kriegeskorte N, Goebel R, Bandettini P. 2006. Information-based functional brain mapping. *Proc Natl Acad Sci U S A.* 103:3863-3868.
- Larsson J, Heeger DJ. 2006. Two retinotopic visual areas in human lateral occipital cortex. *J Neurosci.* 26:13128-13142.
- Lehky SR, Sereno AB. 2007. Comparison of shape encoding in primate dorsal and ventral visual pathways. *J Neurophysiol.* 97:307-319.
- Li N, DiCarlo JJ. 2008. Unsupervised natural experience rapidly alters invariant object representation in visual cortex. *Science.* 321:1502-1507.
- MacEvoy SP, Epstein RA. 2007. Position selectivity in scene- and object-responsive occipitotemporal regions. *J Neurophysiol.* 98:2089-2098.
- Mathot S, Theeuwes J. 2010. Gradual remapping results in early retinotopic and late spatiotopic inhibition of return. *Psychol Sci.* 21:1793-1798.
- Mays LE, Sparks DL. 1980. Saccades are spatially, not retinocentrically, coded. *Science.* 208:1163-1165.
- McKyton A, Pertzov Y, Zohary E. 2009. Pattern matching is assessed in retinotopic coordinates. *J Vis.* 9:19.1-19.10.
- McKyton A, Zohary E. 2007. Beyond retinotopic mapping: the spatial representation of objects in the human lateral occipital complex. *Cereb Cortex.* 17:1164-1172.
- McRae K, Butler BE, Popiel SJ. 1987. Spatiotopic and retinotopic components of iconic memory. *Psychol Res.* 49:221-227.
- Medendorp WP, Goltz HC, Vilis T, Crawford JD. 2003. Gaze-centered updating of visual space in human parietal cortex. *J Neurosci.* 23:6209-6214.
- Melcher D. 2005. Spatiotopic transfer of visual-form adaptation across saccadic eye movements. *Curr Biol.* 15:1745-1748.
- Melcher D. 2007. Predictive remapping of visual features precedes saccadic eye movements. *Nat Neurosci.* 10:903-907.
- Melcher D, Colby CL. 2008. Trans-saccadic perception. *Trends Cogn Sci.* 12:466-473.
- Melcher D, Morrone MC. 2003. Spatiotopic temporal integration of visual motion across saccadic eye movements. *Nat Neurosci.* 6:877-881.
- Merriam EP, Gardner JI, Movshon JA, Heeger DJ. 2008. Modulation of human visual cortex responses by gaze direction. 2008 Neuroscience Meeting Planner. 2008 Nov 15-9; Washington. Washington (DC): Society for Neuroscience. Online.
- Merriam EP, Genovese CR, Colby CL. 2003. Spatial updating in human parietal cortex. *Neuron.* 39:361-373.
- Merriam EP, Genovese CR, Colby CL. 2007. Remapping in human visual cortex. *J Neurophysiol.* 97:1738-1755.
- O'Keefe J, Dostrovsky J. 1971. The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Res.* 34:171-175.
- Ong WS, Hooshvar N, Zhang M, Bisley JW. 2009. Psychophysical evidence for spatiotopic processing in area MT in a short-term memory for motion task. *J Neurophysiol.* 102:2435-2440.
- Op de Beeck HP. 2010. Probing the mysterious underpinnings of multi-voxel fMRI analyses. *Neuroimage.* 50:567-571.
- Pereira F, Botvinick M. 2011. Information mapping with pattern classifiers: a comparative study. *Neuroimage.* 56:476-496.
- Pertzov Y, Avidan G, Zohary E. 2011. Multiple reference frames for saccadic planning in the human parietal cortex. *J Neurosci.* 31:1059-1068.
- Pertzov Y, Zohary E, Avidan G. 2010. Rapid formation of spatiotopic representations as revealed by inhibition of return. *J Neurosci.* 30:8882-8887.
- Pisella L, Mattingley JB. 2004. The contribution of spatial remapping impairments to unilateral visual neglect. *Neurosci Biobehav Rev.* 28:181-200.
- Rawley JB, Constantinidis C. 2010. Effects of task and coordinate frame of attention in area 7a of the primate posterior parietal cortex. *J Vis.* 10:12.1-12.16.
- Sapir A, Hayes A, Henik A, Danziger S, Rafal R. 2004. Parietal lobe lesions disrupt saccadic remapping of inhibitory location tagging. *J Cogn Neurosci.* 16:503-509.
- Sayres R, Grill-Spector K. 2008. Relating retinotopic and object-selective responses in human lateral occipital cortex. *J Neurophysiol.* 100:249-267.
- Schluppeck D, Glimcher P, Heeger DJ. 2005. Topographic organization for delayed saccades in human posterior parietal cortex. *J Neurophysiol.* 94:1372-1384.
- Schwarzlose RF, Swisher JD, Dang S, Kanwisher N. 2008. The distribution of category and location information across object-selective regions in human visual cortex. *Proc Natl Acad Sci U S A.* 105:4447-4452.
- Sereno AB, Maunsell JH. 1998. Shape selectivity in primate lateral intraparietal cortex. *Nature.* 395:500-503.
- Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RB. 1995. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science.* 268:889-893.
- Sereno MI, Huang RS. 2006. A human parietal face area contains aligned head-centered visual and tactile maps. *Nat Neurosci.* 9:1337-1343.
- Sereno MI, Pitzalis S, Martinez A. 2001. Mapping of contralateral space in retinotopic coordinates by a parietal cortical area in humans. *Science.* 294:1350-1354.
- Silver MA, Ress D, Heeger DJ. 2005. Topographic maps of visual spatial attention in human parietal cortex. *J Neurophysiol.* 94:1358-1371.
- Snyder LH, Grieve KL, Brotchie P, Andersen RA. 1998. Separate body- and world-referenced representations of visual space in parietal cortex. *Nature.* 394:887-891.
- Stevens JK, Emerson RC, Gerstein GL, Kallos T, Neufeld GR, Nichols CW, Rosenquist AC. 1976. Paralysis of the awake human: visual perceptions. *Vision Res.* 16:93-98.
- Swisher JD, Halko MA, Merabet LB, McMains SA, Somers DC. 2007. Visual topography of human intraparietal sulcus. *J Neurosci.* 27:5326-5337.
- Talairach J, Tournoux P. 1988. Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system: an approach to cerebral imaging. New York (NY): Thieme.
- Tootell RB, Reppas JB, Kwong KK, Malach R, Born RT, Brady TJ, Rosen BR, Belliveau JW. 1995. Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *J Neurosci.* 15:3215-3230.
- Tower-Richardi S, Golomb J, Kanwisher N. 2011. Binding of location and color in retinotopic, not spatiotopic, coordinates. Naples (FL): Vision Science Society [abstract].
- Trotter Y, Celebrini S. 1999. Gaze direction controls response gain in primary visual-cortex neurons. *Nature.* 398:239-242.
- Ungerleider LG, Mishkin M. 1982. Two cortical visual systems. In: Ingle DJ, Goodale MA, Mansfield RJW, editors. *Analysis of visual behavior.* Cambridge, MA: MIT Press. p. 549-586.
- van Koningsbruggen MG, Gabay S, Sapir A, Henik A, Rafal RD. 2009. Hemispheric asymmetry in the remapping and maintenance of visual saliency maps: a TMS study. *J Cogn Neurosci.* 22(8):1730-1738.
- Wang RF, Spelke ES. 2000. Updating egocentric representations in human navigation. *Cognition.* 77:215-250.
- Weyand TG, Malpeli JG. 1993. Responses of neurons in primary visual cortex are modulated by eye position. *J Neurophysiol.* 69:2258.
- Williams AL, Smith AT. 2010. Representation of eye position in the human parietal cortex. *J Neurophysiol.* 104:2169-2177.
- Williams MA, Dang S, Kanwisher NG. 2007. Only some spatial patterns of fMRI response are read out in task performance. *Nat Neurosci.* 10:685-686.
- Wurtz RH. 2008. Neuronal mechanisms of visual stability. *Vision Res.* 48:2070-2089.
- Zipser D, Andersen RA. 1988. A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature.* 331:679-684.