

Published in final edited form as:

*J Cogn Neurosci.* 2011 January ; 23(1): 119–136. doi:10.1162/jocn.2010.21417.

## The Emergence of Perceived Position in the Visual System

Jason Fischer, Nicole Spotswood, and David Whitney

University of California, Davis

### Abstract

Representing object position is one of the most critical functions of the visual system, but this task is not as simple as reading off an object's retinal coordinates. A rich body of literature has demonstrated that the position in which we perceive an object depends not only on retinotopy but also on factors such as attention, eye movements, object and scene motion, and frames of reference, to name a few. Despite the distinction between perceived and retinal position, strikingly little is known about how or where perceived position is represented in the brain. In the present study, we dissociated retinal and perceived object position to test the relative precision of retina-centered versus percept-centered position coding in a number of independently defined visual areas. In an fMRI experiment, subjects performed a five-alternative forced-choice position discrimination task; our analysis focused on the trials in which subjects misperceived the positions of the stimuli. Using a multivariate pattern analysis to track the coupling of the BOLD response with incremental changes in physical and perceived position, we found that activity in higher level areas—middle temporal complex, fusiform face area, parahippocampal place area, lateral occipital cortex, and posterior fusiform gyrus—more precisely reflected the reported positions than the physical positions of the stimuli. In early visual areas, this preferential coding of perceived position was absent or reversed. Our results demonstrate a new kind of spatial topography present in higher level visual areas in which an object's position is encoded according to its perceived rather than retinal location.

### INTRODUCTION

Although retinotopy has been extensively studied throughout the visual cortex and is considered one of the fundamental principles by which visual areas are organized, there are many circumstances in which the perceived position of an object differs markedly from its retinal position. Object motion (Whitney, 2002; De Valois & De Valois, 1991; Ramachandran & Anstis, 1990), eye movements (Ross, Morrone, Goldberg, & Burr, 2001; Cai, Pouget, Schlag-Rey, & Schlag, 1997; Ross, Morrone, & Burr, 1997), attention shifts (Kerzel, 2003; Suzuki & Cavanagh, 1997), changes in frame of reference (Bridgeman, Peery, & Anand, 1997; Roelofs, 1935), and adaptation (Whitaker, McGraw, & Levi, 1997) are among the many factors that can lead to disparate physical and perceived position information. These examples speak to the fact that reading off an object's retinal coordinates is only a single step in the complex task of object localization; perceived and physical object position are often dissociated. Given this, we might well expect that some visual areas encode position in percept-based rather than retina-based coordinates.

Where might percept-centered position coding exist in the visual system? Retinotopy is well characterized in striate and early extrastriate visual cortex, but the nature of spatial coding in higher level object-, scene-, and motion-processing areas is still unclear. Activity in the

fusiform face area (FFA) and parahippocampal place area (PPA) exhibits relatively weak position selectivity (Schwarzlose, Swisher, Dang, & Kanwisher, 2008; Hemond, Kanwisher, & Op de Beeck, 2007; MacEvoy & Epstein, 2007), but these areas show biases in response amplitude for centrally versus peripherally presented stimuli (Levy, Hasson, Avidan, Hendler, & Malach, 2001). Retinotopic maps have been found in human lateral occipital (LO) cortex (Larsson & Heeger, 2006), but there is also evidence that the predominant organization in LO cortex is head or body centered rather than retina centered (McKyton & Zohary, 2007). Similarly, although position coding in the motion-selective middle temporal (MT) region has previously been reported as retinotopic (Huk, Dougherty, & Heeger, 2002), a recent study that manipulated gaze direction relative to a motion stimulus found that responses in MT were more consistent with a spatiotopic reference frame (d'Avossa et al., 2007; cf. Gardner, Merriam, Movshon, & Heeger, 2008). These conflicting results may be reconciled if the aforementioned areas are integrating retinal and extraretinal sources of information to construct a representation of perceived position—perceived position may or may not match retinal position on a given trial or for a given experimental paradigm.

In the current study, we tested the hypothesis that higher level visual areas preferentially represent perceived rather than physical object position. We measured the relative precision of perceived versus physical position coding in five functionally localized higher level visual areas: LO cortex, posterior fusiform gyrus (pFs), FFA, PPA, and MT+. Using subjects' mislocalizations to dissociate physical and perceived stimulus position, we found that changes in the patterns of activity in each of these higher level areas were more tightly coupled with changes in perceived position than in physical position. Our results reveal the existence of a percept-centered coordinate frame for position coding in higher level visual areas.

## METHODS

Eight subjects with normal or corrected-to-normal vision participated in this study (six participated exclusively in the main experiment, one participated in both the main experiment and the eye tracking control, and one participated exclusively in the eye tracking control). Each subject provided informed consent before participation, and all scanning procedures were approved by the University of California, Davis, institutional review board.

### Functional Localizers: Stimuli and Analysis

In separate functional runs for each subject, we localized visual areas MT+, FFA, PPA, LO cortex, and pFs as well as V1, V2, V3, V3a, VP, and V4. To localize and to demarcate areas V1 through V4, we used flickering wedge stimuli ("bow tie" patterns; Sereno et al., 1995). Bow ties consisted of counterphase flickering (7.5 Hz) radial sine wave patterns of  $11.79^\circ$  radius, subtending an arc of  $8.16^\circ$ . There were three conditions in these bow tie runs; in two conditions, the bow ties were centered on the vertical or horizontal meridians, and the third condition was a fixation baseline in which only the fixation point was present. Conditions were randomized in thirty-six 10-sec blocks. In all conditions, subjects were instructed to fixate while performing a counting task at the fixation point. Small ( $0.98^\circ$ ) radial and circular gratings appeared around the fixation point at random times during each 10-sec block, and one pattern was always presented more often than the other. At the end of each block, a white annulus ( $0.98^\circ$  diameter) appeared around the fixation point, prompting subjects to make a response indicating which pattern had occurred most often.

To define the boundaries of areas V1 through V4, we traced the mirror reversals in the cortical representations of the horizontal and vertical meridians, as identified by the horizontal and vertical bow tie stimuli (Sereno et al., 1995). To do this, we constructed a general linear model (GLM) contrast between the horizontal and the vertical bow tie stimuli. This contrast yielded a striated map of activity across the early visual areas, revealing their retinotopic organization.

We separately overlaid each subject's contrast map on his or her inflated brain and traced the boundaries of areas V1 through V4 by following the horizontal and the vertical meridians.

To define ROIs for areas LO cortex and pFs, we conducted separate localizer scans using methods similar to those used by Grill-Spector et al. (1998). Stimuli in these scans consisted of intact and scrambled objects (Supplementary Figure 1a). Subjects performed a one-back matching task, indicating whether the current object matched the previously presented object, while maintaining fixation throughout each run. Intact objects were gray scaled and centered within  $4.1^\circ \times 5.7^\circ$  rectangles. Scrambled objects were created by dividing each of the object images into a grid of 875 squares and shuffling the locations of these squares within the rectangle. Each run consisted of ten 30-sec stimulation blocks (five with intact objects and five with scrambled objects), interleaved with five 20-sec fixation periods. Within each stimulation block, 40 stimuli were presented (1.33 Hz). Each subject participated in two runs, except for FF, who participated in one run.

To localize areas PPA and FFA, we used a one-back matching task and block design similar to those that we used to localize LO cortex and pFs (Epstein, Harris, Stanley, & Kanwisher, 1999; Kanwisher, McDermott, & Chun, 1997). Within each run, face and house stimuli (Supplementary Figure 1b and c) were presented in 10 alternating 30-sec blocks, interleaved with five 20-sec fixation periods; within each block, 40 stimuli were presented. As with intact objects, the face and the house stimuli were gray scaled and centered within  $4.1^\circ \times 5.7^\circ$  rectangles. Each subject participated in two runs.

ROIs for LO cortex, pFs, PPA, and FFA were defined by conducting GLM analyses on data collected from the localizer runs. For all subjects, each area was separately defined as the region with the strongest contrast in activations when the subject viewed intact objects versus scrambled objects (LO cortex and pFs), houses versus faces (PPA), or faces versus houses (FFA). To select ROIs for six of our subjects, we set the threshold at  $t > 4.9$ ,  $p < .05$ , Bonferroni corrected. One subject had a weak BOLD response, so the threshold was reduced to  $t > 3.6$ ,  $p < .05$ . The inclusion or exclusion of this subject's data did not significantly influence any results in this study. Several subregions of the LO cortex have been identified; however, these have not been clearly established and universally agreed upon (McKyton & Zohary, 2007; Larsson & Heeger, 2006; Grill-Spector et al., 1999). The Talairach coordinates we found corresponded most with studies that identified these subregions as LO and pFs; therefore, we chose to retain this nomenclature in our study (Altmann, Deubelius, & Kourtzi, 2004; Avidan, Hasson, Hendler, Zohary, & Malach, 2002; Grill-Spector et al., 1999).

Each subject also participated in separate runs to functionally localize area hMT+ (the human homologue of monkey areas MT and MST, commonly referred to as MT+). The MT+ localizer runs consisted of three conditions: Gabors with inward motion (drifting toward fixation), Gabors with outward motion (drifting away from fixation), and a fixation baseline condition in which only the fixation point was present. The Gabors were situated as in the main experiment: One Gabor was presented in each visual quadrant at an eccentricity of  $9^\circ$  from fixation. The Gabors had a spatial frequency of  $0.38^\circ$  and drifted inward or outward at 2.5 Hz for the duration of each block. The three stimulus conditions were randomized in eighteen 10-sec blocks, and during these blocks, subjects performed a task at the fixation point identical to the one described for the bow tie localizer above. Subjects participated in a minimum of five MT+ localizer runs.

MT+ ROIs were functionally defined for each subject by contrasting moving versus baseline responses in a GLM applied to the localizer runs. The threshold for inclusion in the ROI was  $t > \pm 6.3$ ,  $p < .001$ , Bonferroni corrected. Two subjects had weak responses to all stimuli, so the threshold for inclusion was dropped to  $t > \pm 3$ ,  $p < .003$ . The inclusion or exclusion of these

subjects did not significantly change the results. The Talairach coordinates for the functionally defined MT+ ROIs were consistent with those estimated in previous studies (Dukelow et al., 2001; Dumoulin et al., 2000; Kourtzi & Kanwisher, 2000; Tootell et al., 1995; Watson et al., 1993). See Supplementary Table 1 for the averaged Talairach coordinates of LO, pFs, PPA, FFA, and MT+ across all seven subjects (Talairach & Tournoux, 1988).

## Main Experiment Stimuli

Stimuli consisted of four flickering Gabor patterns (sinusoidal luminance modulations within Gaussian contrast envelopes; Figure 1A). The Gaussian contrast envelope of the Gabors was

defined as  $L(x, y) = A \times e^{\frac{-r^2}{(\sigma M)^2}}$ , where  $A$  is the peak contrast amplitude,  $r$  is the distance of  $(x, y)$  from the center of the Gaussian,  $\sigma$  is the standard deviation, and  $M$  is the maximum radius. Gabors had a spatial frequency of 0.38 cycles/deg and flickered in counterphase at 7.5 Hz; the phase of each Gabor was independently randomized on each trial. One Gabor was positioned in each visual quadrant, with the peak contrast (87% Michelson) always located 9.04° from a central fixation point. We used Gabors rather than the optimal stimulus for each ROI to avoid generating activity specific to one particular area. In addition, the physical characteristics of the Gabor stimuli (e.g., position, contrast, and spatial frequency) are easily quantified and can be independently manipulated; the features of objects, faces, or houses (e.g., surface properties, contours, and asymmetries) are difficult to control.

In each of five experimental conditions, the centroids of all four Gabors were set to one of five possible eccentricities. Gabor centroids were manipulated by applying a skew to the Gaussian contrast envelopes (Whitaker, McGraw, Pacey, & Barrett, 1996). A central condition had no skew applied to the contrast envelope; its size (standard deviation) was 1.66° and its centroid was at 9.039° eccentricity. In two more foveal conditions, the contrast envelopes were skewed by 0.19° and 0.38° toward fixation, and in two more eccentric conditions the contrast envelopes were skewed by 0.19° and 0.38° away from fixation. The Gabors skewed in this way had centroids at 8.430°, 8.735°, 9.039°, 9.343°, and 9.647° from fixation on the basis of the equation  $(\sigma_2 - \sigma_1) \sqrt{\frac{2}{\pi}}$  (Whitaker & McGraw, 1998; Whitaker et al., 1996). Eccentricities were manipulated in this way rather than by shifting the peak contrast because skewing the contrast envelope is a better method of isolating the perceptual mechanisms that perform centroid analysis (Whitaker et al., 1996), and skewing the Gabor envelope dissociates peak contrast from centroid information. In a previous study, we showed that skewing a Gabor's Gaussian envelope versus shifting its peak contrast are both equally valid means of altering the retinotopic representation of the pattern (Whitney & Bressler, 2007). A sixth condition consisted of a fixation baseline in which only the fixation point, a 0.39° diameter bulls eye, was present.

In a control analysis, we presented faces instead of Gabors (Supplementary Figure 5). The stimuli consisted of four faces, situated symmetrically about the fixation point in the four visual quadrants, just as with the Gabor stimuli. The faces were drawn from the PICS database (University of Stirling Psychology Department; <http://pics.psych.stir.ac.uk/>). Each face was enveloped in a Gaussian contrast window to define its centroid and standardize its size. The Gaussian contrast profiles had a standard deviation of 1.66°; to yield five position conditions, the envelope was skewed by -0.38°, -0.19°, 0°, 0.19°, or 0.38°, and applied to the faces, yielding centroids at 8.430°, 8.735°, 9.039°, 9.343°, and 9.647° from fixation, just as with the Gabor stimuli. The identity of the four faces updated at 7.5 Hz, and on each update, the next identity was randomly drawn from a pool of 20 possible identities. In all other respects, the experimental parameters for the face position discrimination experiment were the same as those for the main experiment.

## Experimental Design and Task

In each functional imaging run, the six stimulus conditions were randomly interleaved in thirty-six 10-sec blocks (each condition was presented six times; runs were 360 sec in length). A blocked design was used to maximize signal to noise ratio in the resulting maps of BOLD response. Each subject participated in five functional runs.

Subjects maintained fixation at a central point throughout the entire experiment. On each trial, subjects attended to the locations of the surrounding Gabor stimuli and judged the position (centroid eccentricity) of the Gabors. This task was a five-alternative forced-choice (5AFC) classification task (MacMillan & Creelman, 2004), with one Gabor position assigned to each of five buttons on a button box (method of single stimuli; McKee, Silverman, & Nakayama, 1986; Volkman, 1932). At the end of each 10-sec trial, a white annulus (0.98° diameter) appeared around the fixation point, prompting subjects to respond by pressing the button associated with the position in which the Gabors appeared.

To prevent subjects from making the position judgments on the stimuli quickly and then attending elsewhere for the remainder of each trial, subjects performed a secondary sustained attention task at the locations of the Gabors. During the first 8 sec of each 10-sec block, at a randomly chosen time, a patterned circle was briefly superimposed on one of the four Gabors (chosen at random) for 500 msec. The patterned circle (either a circular or a radial pattern) was always presented at an eccentricity of 9.04°. A second patterned circle was presented again in the same manner during the last 2 sec of each 10-sec block, and subjects responded to indicate whether the first patterned circle presented matched the second circle (patterns matched with a probability of 50%). Figure 1D depicts an example trial. In this way, we ensured that subjects maintained attention at the locations of the surrounding stimuli for the entire 10 sec of each trial. In addition to the functional runs described above, additional runs were interleaved in which subjects performed a task at the fixation point and did not perform the 5AFC position discrimination task. These runs were part of a separate analysis and are not included in the present results.

We also ran a subset of the subjects from the main experiment in an event-related experiment in which stimulation intervals were only 2 sec long. We found very similar results to those from the main experiment, but the signal strength in the BOLD response was substantially reduced. Because the power of our pattern analysis depends on signal-to-noise ratio in the BOLD response (Fischer & Whitney, 2009a), our main experiment's blocked design was aimed at achieving an optimal compromise between the psychological and the analytical demands of the study.

## fMRI Data Acquisition and Preprocessing

Imaging was conducted at the UC Davis Imaging Research Center on a 3-T Siemens TRIO scanner. Each subject's head was placed in a Siemens eight-channel phased-array head coil, and padding was placed on the side and forehead of the subject to restrict movement. Using a Digital Projection Mercury 5000HD projector, stimuli were back projected at 75 Hz onto a semitransparent screen from outside the bore. Subjects were able to see the screen and the stimuli via a mirror angled at 45°, located 9.5 cm directly above their eyes. Functional images were collected with a gradient-recalled echo EPI sequence. Whole-brain anatomical images were acquired with a high resolution (1 mm<sup>3</sup>) Turbo Spin Echo scan. The acquisition parameters were repetition time = 2000 msec, echo time = 26 msec, flip angle = 90°, field of view = 22 × 22 cm<sup>2</sup>, voxel size = 1.528 × 1.528 × 2.5 mm<sup>3</sup>, 20 slices per volume. The imaging volume was centered on the calcarine sulcus, covering the occipital lobe.



All preprocessing and GLM analyses were conducted using Brain Voyager QX (Brain Innovation B.V., Maastricht, The Netherlands). Preprocessing included linear trend removal and 3D motion correction on a run-by-run basis. Before all GLM analyses, corrections for serial correlations (removal of first-order autocorrelations) were applied. The images from each functional run were individually aligned to the subject's respective high-resolution anatomical image, reducing the effects of head movement between runs. The anatomical images were then transformed into Talairach space.

## Data Analysis

In the main analysis, we discarded the “hit” trials and analyzed only the “miss” trials to focus on the subset of the data in which there was a difference between the physical and the perceived locations of the stimuli. To measure position selectivity within each ROI, we first generated a map of the BOLD response corresponding to each stimulus eccentricity condition. We did this separately for each subject's five functional runs in a GLM analysis, using the six stimulus conditions (five Gabor eccentricities, plus a baseline condition) as predictors. We separately contrasted each of the five Gabor eccentricities against the baseline condition to produce five statistical maps ( $t$  values, unthresholded) of the BOLD activity unique to the five stimulus eccentricities. The subsequent correlation analysis measured position discrimination by tracking systematic changes in the patterns of activity in these maps.

To quantify the precision of position discrimination, we generated a position discrimination plot for each ROI. To do this, we first computed the correlations between all possible pairings of activity maps within each ROI. Given any two of a subject's five maps of BOLD response (corresponding to the five stimulus eccentricities), we computed the correlation between the maps by pairing the  $t$  values from the two on a voxel-by-voxel basis and computing a Pearson  $r$  for the resulting set of pairs. Figure 2 illustrates this process for two of the 10 correlations: correlating the activity from adjacent stimuli generally yielded large  $r$  values (Figure 2A), whereas correlating the activity from more distant stimuli yielded smaller  $r$  values (Figure 2B). We converted the 10 resulting  $r$  values to Fisher  $z$  scores so that they could be linearly compared with each other. This process yielded 10  $z$  scores for each ROI; to create a position discrimination plot, we plotted each of the  $z$  scores against the spatial separation (in degrees visual angle) of the two stimulus eccentricities from which it was produced (Figure 2C). Adjacent conditions had centroids separated by  $0.304^\circ$ ; more distant conditions were separated by multiples of this value ( $0.609^\circ$ ,  $0.912^\circ$ , and  $1.216^\circ$ ). To avoid a loss of information due to imperfect registration between runs, we performed this process separately for each run; thus, each subject's individual position discrimination plot for a given ROI contained a total of 50 points (Figure 2C depicts data for a single run). For group analyses, we fit a regression to all subjects' data taken together. As we wished to combine data across several subjects to make inferences about the population at large, we used a random effects analysis to account for between-subject variability (Holmes & Friston, 1998). Our regression model took the form  $z_{ijkl} = \beta_0 + \tau_i + \beta_1 x_{jk} + \varepsilon_{ijkl}$ , where  $i$  indexed the subjects, the pair  $(j, k)$  indexed the 10 stimulus pairings, and  $l$  indexed the run number;  $\tau_i$  accounted for baseline differences between subjects. The goodness of fit of the group regression ( $r$ ) for each ROI provided an index of the precision of position coding there—the stronger the inverse relationship between the BOLD response correlations and their corresponding stimulus separations, the more precise the position discrimination (Fischer & Whitney, 2009a, 2009b; Bressler, Spotswood, & Whitney, 2007). We expressed the precision of position coding for each plot as the Fisher  $z$ -transformed  $r$  value for that plot. Although the slope of a position discrimination plot also reflects the precision of position coding in the corresponding ROI,  $r$  serves as a better indicator of coding precision because it captures the *significance* of slope. That is,  $r$  reflects the steepness of the linear fit and the compactness of points around the regression line, both of which are important indicators of the precision of position coding. Although the addition of an outlier to a plot might increase

the slope of the linear fit, the corresponding  $r$  value will generally decrease, reflecting a greater scatter in the points around the regression line. The latter is desirable, as in our case, greater scatter means less systematic encoding of stimulus position. We took the Fisher  $z$ -transform of each  $r$  value to linearize the scale of the precision estimates so that they could be directly compared. Table 1 gives the position discrimination fit estimates for each ROI.

We measured the precision of *perceived* position coding in the same manner, substituting BOLD response maps corresponding to subjects' responses for the original maps that reflected physical stimulus position (see Supplementary Figure 2). We created these perceived position maps by using the subjective reports provided by each subject as predictors in the GLM. For example, if on a given trial a subject indicated that the Gabors were at the most eccentric position, then that trial was coded as condition 5 in the GLM, regardless of what was actually presented during that trial. We used only "missed" trials in the analysis of both physical and perceived position coding, so each trial was coded differently in the two analyses. We performed the correlation analysis a second time, using each subject's five perceived position maps (corresponding to the five possible responses) to create perceived position discrimination plots for each ROI. Figure 3A shows the perceived and physical position discrimination plots for area LO. The two plots differ with respect to the goodness of their linear fits, hence with respect to the precision of coding that they reveal for physical and perceived position. Table 1 gives the precision of perceived position discrimination for each ROI (group data). To compare the precision of physical and perceived position coding within each ROI, we performed a  $Z$  test:

$$Z_{\text{diff}} = \frac{Z_{\text{perceived}} - Z_{\text{physical}}}{\sqrt{\frac{1}{N_1 - 3} + \frac{1}{N_2 - 3}}},$$

where  $N_1$  and  $N_2$  are the numbers of points in the perceived and physical position discrimination plots, respectively. Note that in Figure 3A, there are fewer total points in the plot for perceived position than in the plot for physical position. It was occasionally not possible to create a map for a particular perceived position when no corresponding trials were available (e.g., on a given run, a subject may never have made an erroneous "5" response). The  $Z$  test for significance accommodates this difference in the number of data points (Snedecor & Cochran, 1980). In subsequent control analyses, we also conducted the position discrimination analysis using all trials instead of only the missed trials. In this case, the plots for physical and perceived position had the same number of points, and the results showed the same effects as in the main experiment (see Figure 5 and Supplementary Figure 4).

To test for significant heterogeneity in the nature of position coding across the visual areas in the dorsal and ventral streams, we performed a chi-square test on the  $(Z_{\text{perceived}} - Z_{\text{physical}})$  scores for each of the two collections of visual areas. The test is given by  $\chi^2 = \sum ((N_i - 3)(z_i - \bar{z})^2)$ , where each  $z_i$  is a  $(Z_{\text{perceived}} - Z_{\text{physical}})$  score,  $\bar{z}$  is the weighted mean of all such scores, and  $N_i$  is the number of points on each position discrimination plot. Subsequently, we tested for a trend in the nature of position coding across visual areas by computing a Spearman rank correlation coefficient between the areas' ordering in the visual processing hierarchy and their bias for physical or perceived position coding, given by  $(Z_{\text{perceived}} - Z_{\text{physical}})$ . We ordered the early visual areas according to the order in which they are encountered when moving anteriorly across the cortex from V1 (allowing for duplicate ranks). As a guide, we used Figure 1 from Tootell, Tsao, and Vanduffel (2003). Thus, the early visual areas were ranked as follows: V1 – 1, V2 – 2, V3 – 3, VP – 3, V3a – 4, and V4 – 4. All of the higher level visual areas were assigned the same rank of 5. We computed  $\rho$  and its corresponding  $p$  value in SPSS 17.0.

## Eye Tracking Data Collection and Analysis

For two control subjects, we monitored eye position during scanning. Eye position was monitored at 60 Hz using an ASL Eye-Trac 6 series long-range eye tracker. Eye position data were collected using EyeTrac6000 software and subsequently analyzed in Matlab 7.1. To determine whether the subjects' eye movements were correlated with either the stimuli or their responses, we aligned the eye position data with the stimulus presentation epochs for each run (Figure 6B). We then populated a separate list of sampled eye positions for each of the five physical positions and each of the five perceived conditions. All of the eye positions in the physical condition lists were also present in the perceived condition lists, but many were assigned to a different condition number because the subject misperceived the stimulus position. For each list, we computed the mean eye position in the  $x$  direction (purple data points in Figure 6C) and in the  $y$  direction (green data points) as well as the mean squared distance from the fixation point as a measure of variability. For each subject, separately for the physical and perceived stimulus positions, we performed three one-way ANOVAs, testing for significant differences in  $x$  position,  $y$  position, and variability.

## RESULTS

In separate imaging runs, we functionally localized visual areas LO, pFs, FFA, PPA, and MT + using standard techniques, including the presentation of objects, buildings, faces, and moving stimuli (Supplementary Figure 1; see Methods) (Epstein et al., 1999; Grill-Spector et al., 1998; Kanwisher et al., 1997; Sereno et al., 1995). The Talairach coordinates of our ROIs were in good agreement with those reported in previous studies (Supplementary Table 1; Yi, Kelley, Marois, & Chun, 2006; Altmann et al., 2004; Grill-Spector, Knouf, & Kanwisher, 2004; Epstein, Graham, & Downing, 2003; Avidan et al., 2002; Dumoulin et al., 2000; Kourtzi & Kanwisher, 2000; Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999; Grill-Spector et al., 1999; Watson et al., 1993). In the main experiment, our goal was to measure the precision of stimulus position information in the pattern of BOLD response in each ROI. During scanning, we presented flickering Gabor patches at five possible eccentricities, ranging from 8.43° to 9.65° from fixation (Figure 1A). In each 10-sec block, the stimuli were presented at one of the five eccentricities, and subjects reported their apparent position in a 5AFC response. Seven subjects participated in the main experiment.

Figure 1B shows behavioral performance: Subjects' ability to discriminate between two different conditions (discrimination sensitivity,  $d'$ ) is plotted as a function of the separation between the two eccentricities (for details on calculating pairwise sensitivity, see MacMillan & Creelman, 2004). A one-way ANOVA across the five conditions at the smallest (0.30°) separation showed no performance bias for any particular stimulus eccentricity,  $F(4, 30) = 0.75$ ;  $p = .57$ . The positive trend in  $d'$  for increasing stimulus separation indicates that when subjects made errors, they were most likely to mistake the presented condition for an adjacent one. This tendency is also evident in Figure 1C, which shows a histogram of subjects' errors. Trials are binned according to the difference between the subject's response and the correct response; the trials in bin 0 are correct trials, and trials in positive bins are those in which the subject perceived the stimulus as more eccentrically positioned than it was presented. Subjects performed well at determining the positions of the Gabors (~58% correct; chance is 20%), but the task was sufficiently difficult to elicit a substantial number of errors. In the subsequent data analysis, we focused exclusively on these "missed" trials, in which subjects mislocalized the positions of the stimuli to dissociate physical and perceived position.

Subjects also performed a secondary task to ensure that they attended at the locations of the Gabors for the entirety of each trial. During the first 8 sec of each trial, a small pattern appeared at the centroid of one of the Gabors. At a randomly chosen time during the final 2 sec of each trial, a second pattern appeared in one of the Gabors; the second pattern matched the first on



half of the trials. At the end of each trial, in addition to reporting the positions of the Gabors, subjects indicated whether the two patterns matched (Figure 1D shows an example trial). Performance on this secondary task was 75.9% correct (chance was 50%). Subjects' responses on the secondary task were not correlated with their responses on the primary position discrimination task,  $F(4, 1040) = 0.53$ ;  $p = .72$ , nor were they correlated with the physical positions of the stimuli,  $F(4, 1040) = 0.25$ ;  $p = .91$ . Because responses on the primary and secondary tasks were not correlated, subjects did not use the patterns to judge the positions of the Gabor stimuli. The below ceiling performance on this task indicates that it was demanding enough to hold subjects' attention at the location of the Gabors for the duration of each trial.

Within each ROI, we measured the selectivity for stimulus position by tracking the change in the spatial pattern of the BOLD response corresponding to the incremental shifts in stimulus position (see Methods). We first discarded the "hit" trials and kept only the "miss" trials for the main analysis; by doing so, we focused on the subset of the data in which there was a difference between the physical and the perceived locations of the stimuli (for the results of the same analysis when all trials are included, see Figure 5 and Supplementary Figure 4). For each subject, we produced five maps of BOLD response corresponding to the five stimulus positions by separately contrasting each condition against a fixation baseline in a general linear model. We then tested for a trend in the spatial pattern of the BOLD response by measuring the similarity between pairs of maps. For any given pairing, we computed the correlation between the two maps within the current ROI (Figure 2A and B shows example correlations computed for conditions separated by  $0.30^\circ$  and  $1.22^\circ$ , respectively). Collecting the correlations from all 10 possible pairings of the five maps, we plotted each correlation against the spatial separation between the stimuli that produced it to construct a position discrimination plot (Figure 2C). An ROI that codes stimulus position will have a strong negative trend in its position discrimination plot because stimuli that are closer together in space produce more highly overlapping patterns of BOLD response. In an ROI that does not contain position information, on the other hand, there is no reason to expect any trend in the correlations on the position discrimination plot. Thus, we used  $R$ , the goodness of fit of a linear regression applied to the data on the position discrimination plot, as an index of the precision of position coding in each ROI. The goodness-of-fit measure captures how tightly clustered the correlations are at each separation as well as the slope of the data, which indicates how dramatically the pattern of BOLD changed with incremental changes in position (see Methods). We took the Fisher  $z$ -transform of each  $r$  value to linearize the scale of the precision estimates so that they could be directly compared. We have previously shown that this pattern analysis technique is able to measure position selectivity in the BOLD response on a submillimeter scale (Fischer & Whitney, 2009a, 2009b; Bressler et al., 2007; Whitney & Bressler, 2007).

To test the precision with which each ROI codes *perceived* rather than physical stimulus position, we constructed position discrimination plots in the same way, this time using maps of BOLD response corresponding to the five possible responses from the position discrimination task. These maps, corresponding to the positions in which subjects perceived the stimuli on a trial-by-trial basis, were produced using the exact same trials as in the analysis of physical position coding; the only difference was in how the trials were coded in the GLM analysis (see Supplementary Figure 2). The resulting maps reflected activity corresponding to perceiving the stimuli in the five different locations, regardless of variations in the actual locations of the stimuli. Because we used only missed trials, every trial was coded differently when coded by physical position versus the subject's response. In this way, we were able to dissociate physical and perceived position and to separately measure the amount of information about each in the BOLD response.

Figure 3 shows the results of the position discrimination analysis for area LO (Figure 3A) and the other four higher level areas we tested (Figure 3B). The position discrimination plots in

Figure 3A show all subjects' data plotted together; to measure a group-level effect, we fitted a regression to all subjects' data and included a random effect of subject in the regression model to account for between-subject variability (Holmes & Friston, 1998; see Methods). Precision of physical and perceived position coding is plotted as  $-z$ , reflecting the fact that a strongly negative linear fit to a position discrimination plot reflects precise position coding. Every area we tested showed significant discrimination of physical stimulus position ( $p < .0001$  for all areas). Strikingly, however, the coding of perceived position was significantly more precise than coding of physical position in every area (red bars in Figure 3; see Table 1 for statistics). Consistent with our hypothesis, comparing the relative precision of physical versus perceived position coding within each area revealed a preferential representation of perceived position in these higher level visual areas. Supplementary Figure 3 shows physical and perceived position discrimination in the five higher level areas from Figure 3, broken down by individual subject. With few exceptions, individual subjects showed effects in the same direction as in the group-level analysis. The fact that subjects showed a consistent pattern of results, with higher level visual areas more precisely reflecting perceived position than physical position, speaks to the existence of an underlying percept-centered coding scheme. Each subject made a unique set of errors, and any alternative coding scheme tied solely to the physical positions of the stimuli would have been washed out in our analysis, in which a given condition could be, at various times, reassigned to any of the other four conditions.

In the main analysis, we used only unthresholded BOLD response maps because nonsignificant voxels can still carry precise stimulus information when analyzed as a multivariate pattern (Norman, Polyn, Detre, & Haxby, 2006). We also tried the same analysis using thresholded BOLD maps and found very similar results to those in the main experiment (see Supplementary Figure 4).

The smallest and largest eccentricities define the endpoints of the continuum of possible responses, so we wondered whether a dramatic difference between those two conditions in the "percept" maps might be driving the higher precision of perceived position coding, relative to physical position coding. We performed the position discrimination analysis again after removing the correlations at the  $1.22^\circ$  separation, and we still found stronger discrimination of perceived position than physical position in every higher level visual area (the least significant difference was in the FFA:  $Z = 2.06$ ,  $p = .039$ ). We also measured the similarity of the physical and perceived BOLD response maps within each of the five position conditions to determine whether the most dramatic changes resulting from recoding were isolated to any particular condition(s). Correlating the physical and the percept maps with each other in each ROI, we found no difference across the five conditions for any ROI (most significant was FFA),  $F(4, 170) = 0.96$ ,  $p = .43$ . Thus, systematic differences between the patterns of BOLD corresponding to physical and perceived position are present across the whole continuum of stimulus positions.

In a follow-up analysis, we measured physical and perceived position discrimination in early visual areas V1, V2, V3, V3a, VP, and V4, which we functionally defined using a standard retinotopic localizer (Supplementary Figure 1; see Methods; Sereno et al., 1995). The collected data for all visual areas, divided into the dorsal and ventral visual streams, are presented in Figure 4. The within-area difference scores  $-(Z_{\text{percept}} - Z_{\text{physical}})$ , plotted below the raw discrimination data, are the most informative indices of the nature of position coding in each visual area because they reveal the direction and the strength of the bias within an area for encoding physical versus perceived position discrimination. It is important to note that due to potential inhomogeneities in signal quality across the brain, a comparison of position discrimination scores is only informative within, and not between, visual areas (see Discussion).

The data in Figure 4 reveal a trend in the nature of position representation: Although the higher level visual areas we tested showed a strong preference for coding perceived stimulus position, that effect is diminished or reversed in earlier areas. A chi-square test revealed significant heterogeneity in coding preference across both the dorsal and the ventral stream areas ( $\chi^2_{\text{dorsal}} = 20.92, p = .0003$ ;  $\chi^2_{\text{ventral}} = 35.29, p < .0001$ ). To test for a systematic progression in the nature of position coding across areas, we computed a nonparametric correlation (Spearman rho) between the rank-ordered visual areas and their coding preference  $-(Z_{\text{percept}} - Z_{\text{physical}})$  scores. We assigned all higher level areas the same rank (5), and we ranked the earlier visual areas according to the order they are encountered moving anteriorly from V1 (for reference, we used the inflated human cortex image from Figure 1 in Tootell et al., 2003): V1 – 1; V2 – 2; V3 and VP – 3; V3a and V4 – 4. The correlation was highly significant ( $\rho = .80$ ;  $p = .003$ ). To evaluate this a priori ranking relative to all possible rankings of the visual areas, we computed a 25,000-sample bootstrapped distribution of rank correlations, with a randomly drawn ranking of areas for each sample. The correlation of .80 obtained with the a priori ranking was larger in absolute value than 99.5% of the bootstrapped samples, indicating that our ranking based on functional anatomy is a good match to the independently measured position discrimination estimates for each area. The strong correlation between an area's location in the visual processing hierarchy and its position coding bias reinforces the idea that the nature of position coding evolves as information progresses through the visual processing hierarchy, becoming relatively more strongly tied to perception in higher level areas. This is not to say that early visual areas contain no information about perceived position but that percept-centered information does not yet dominate position representations at the earliest stages of visual processing. Our results provide a clear answer to our underlying question regarding the nature of position coding in higher level areas: The location in which an object is perceived matters more than the location in which it was presented.

If subjects' errors had been random or not directly related to perceived position, could recoding the condition labels according to those (uninformative) responses ever yield a spurious improvement in the precision of position coding measured by the correlation analysis? To evaluate how uniquely predictive subjects' actual responses were relative to other possible ways in which the trials could have been relabeled, we performed a bootstrapping analysis in which we repeated the correlation analysis for each subject and each ROI 1000 times. On each iteration, we added "errors" to the physical trials labels; the errors were randomly drawn from the distribution of subjects' actual errors, such that over all iterations, the distribution of simulated errors matched the distribution of actual errors that subjects made (Figure 1C). For each ROI, we ran the position discrimination analysis with the 1000 sets of simulated trial labels and obtained a distribution of discrimination values (Figure 5). Since the errors for each set of trials labels were randomly drawn, the bootstrapped discrimination values reflect the precision of position discrimination we would expect to measure if subjects' responses reflected random errors. In Figure 5, the bootstrapped discrimination values are shown in gray histograms, and the discrimination measured with the actual physical and percept trial labelings is indicated with blue and red bars, respectively (these values differ slightly from those measured in the main experiment; here, we used all trials in the analysis rather than only missed trials to allow for a direct comparison of all iterations). In each of the higher level visual areas, the discrimination measured with the actual percept labels falls in the extreme upper tail of the distribution of discrimination values for all possible trial relabelings (least significant was pFs;  $-Z = 3.41, p < .001$ ). The degree to which subjects' responses outperform other possible labelings is striking given that both the percept labelings and the bootstrapped labelings had, on average, 58% of their labels in common with the physical coding; the only difference is the precise arrangement of the errors. Thus, it is indeed critical to use the precise responses that subjects made to obtain the dramatic improvement in the precision of position coding that we measured in higher level areas.

The bootstrapping analysis also allowed us to further disentangle the measurements of percept- and retina-centered encoding. In an area dominated by retinotopic coding such as V1, one might ask whether there is also some degree of encoding of perceived position. In the main experiment, we found significant discrimination of the perceived stimulus positions in V1, but because subjects' responses were highly correlated with the physical stimulus positions (Figure 1C), that measurement could have been carried entirely by the underlying retinotopic coding. However, the histograms in Figure 5 provide a means of testing for a unique contribution of percept-related information after taking into account the correlation between the subjects' responses and the physical stimulus positions. If there was no unique percept information in V1, then the discrimination of subjects' responses (the red bar) should fall near the center of the bootstrapped distribution, indicating that labeling the trials according to subjects' responses is no better than using a random perturbation of the physical trial labels. In fact, while a percept-based labeling scheme is substantially worse than a physical labeling scheme at predicting the pattern of BOLD in V1, it still falls 2.38 standard deviations ( $p = .017$ ) above what would be expected by chance if there were no percept-specific information in V1. Thus, there is some information about perceived position encoded in V1, even after accounting for the correlation between the subjects' responses and the physical stimulus positions (the bootstrapped labels were, on average, as strongly correlated with the physical labels as subjects' responses were). Nonetheless, the relative precision of retinotopic coding versus percept-centered coding is stronger in V1 than in any other area we measured.

A potential concern is that subjects' eye movements might have been correlated with either the physical or the perceived stimulus positions, displacing the stimuli on the retina in a systematic fashion. In fact, the main experiment was designed specifically to factor out the impact of eye movements: Because the same trials were used to compare the relative precision of physical and perceived position coding in every visual area, any effects of eye movements would have had the same impact in every ROI and could not have produced effects in opposite directions as we found in higher level visual areas versus early areas (Figure 4). Nonetheless, to test for a possible correlation between eye movements and subjects' responses, we conducted a control experiment in which we tracked two subjects' eye positions during scanning (Figure 6). The position discrimination data from the control subjects were consistent with the results of the main experiment (Figure 6A; see caption for stats). Figure 6B shows a sample eye trace (sampled at 60 Hz) for one run from Subject 1a, with the stimulus conditions for that run indicated in shades of blue. The correlation between the eye position and the stimulus conditions for this run was  $r = .031$ ,  $p = .86$ . The largest correlation for any run was  $r = .048$ ,  $p = .78$ . Figure 6C shows the mean  $x$  position (purple) and  $y$  position (green) of gaze during each condition for both subjects. We performed one-way ANOVAs for each of  $x$  position,  $y$  position, and variability in gaze position, separately for the eye measurements grouped by physical position and perceived position. In neither subject was gaze position or variability related to the physical or perceived stimulus positions (Subject 1a: most significant test was  $F_{y\_pos \times percept} = 1.00$ ,  $p = .41$ ; Subject 2a: most significant test was  $F_{var \times pgghysical} = 1.57$ ,  $p = .18$ ), indicating that subjects did not make eye movements of different magnitudes or directions in the different conditions. Mean position and variability are easily visualized in Figure 6D, which show scatterplots of the recorded gaze positions for each condition for Subject 1a. Together, these data show that eye movements cannot account for our results.

Perhaps the faithful representation of retinal position within a visual area depends on the presentation of its preferred stimulus. To explore whether the preferential coding of perceived position in higher level visual areas is affected by stimulus type, we conducted an additional control experiment in which we presented faces instead of Gabors (Supplementary Figure 5a). The faces were enveloped in a Gaussian contrast profiles to define their centroids, and they were updated in identity at 7.5 Hz. In all other respects, the stimuli and the task were identical to those in the main experiment (see Methods). Two subjects from the main experiment

participated in this control experiment so that we had data for both Gabor stimuli and face stimuli for those subjects. We compared the precision of physical and perceived position coding within the FFA, for which faces are an optimal stimulus (Kanwisher et al., 1997). Supplementary Figure 5b shows the position discrimination scores measured independently using Gabor stimuli and face stimuli. Of critical interest is whether the perceived versus physical bias in the FFA, given by  $-(Z_{\text{percept}} - Z_{\text{physical}})$  (the green bars in Supplementary Figure 5b), differed depending on stimulus type. A Z test on the two bias estimates revealed no significant difference ( $Z = 0.10$ ,  $p = .92$ ); in fact, the bias estimates using the two different stimulus types were remarkably similar. Notably, the overall selectivity for both perceived and physical position was higher in the FFA for faces than Gabors (this difference was borderline significant:  $t = 11.90$ ,  $p = .053$ ). This suggests that while optimizing the stimulus for a given visual area might improve the overall precision of position discrimination there, changes in stimulus type do not change the underlying nature of position coding.

A final concern is that higher level visual areas may possess a representation of the digits used to make the 5AFC position discrimination response. Because subjects made two independent responses on each trial (a 5AFC position discrimination on the right hand and a same/different texture discrimination on the left hand), we were able to test for evidence of digit coding or response planning in each ROI, independently of the main position discrimination task. We applied the same correlation analysis as in the main experiment after recoding the design matrix according to subjects' responses on the secondary (pattern matching) task (Supplementary Figure 6). No ROI was able to discriminate the digit used to make the secondary response (Supplementary Figure 6a; most significant was MT+;  $-Z = 0.19$ ,  $p = .24$ ). More importantly, across visual areas, variation in the precision of digit coding was not correlated with the degree of bias for encoding physical or perceived position (Supplementary Figure 6b;  $r = .13$ ,  $p = .68$ ). Thus, we can be confident that the percept-centered position coding that we measured in higher level visual areas is not simply due to encoding of response planning or tactile input.

## DISCUSSION

Our results reveal a precise representation of perceived object position in every higher level visual area we tested. Although these areas do carry some information about a stimulus's retinotopic position, each carries significantly more information about the location in which the stimulus is perceived. The remarkable precision of perceived position coding that we found in areas pFs, FFA, and PPA, in which position information has previously been regarded as coarse at best (Hemond et al., 2007; MacEvoy & Epstein, 2007; Grill-Spector & Malach, 2001), suggests that it is critical to take into account a subject's perceptual experience—not just the stimulus properties—when measuring stimulus coding in higher level visual areas.

It is important to have a concrete interpretation of what it means to reassign the GLM predictors according to subjects' responses, as we did in the current analysis. Because subjects reported the apparent positions of the stimuli, there were effectively two distinct stimulus dimensions that we could define as predictors in the GLM (see Supplementary Figure 2). The first was physical object position (given by the five possible stimulus eccentricities), and the second was perceived object position (given by the five possible responses). The resulting two sets of BOLD maps revealed the variations in neural activity due to changes in physical stimulus position and the variations in activity due to changes in perceived position, respectively. Because we included only the missed trials in our primary analysis, each trial had a different value on the perceived and physical dimensions, which provided the maximum possible independence between the two dimensions of interest. Similar techniques investigating “miss” trials are not uncommon in, for example, memory research (Henson, Hornberger, & Rugg, 2005; Eldridge, Knowlton, Furmanski, Bookheimer, & Engel, 2000). Subsequent analyses showed that



including all trials (not just missed trials) yields the same evolution of position coding from retina centered to percept centered (Figure 5).

An assumption of our analysis is that subjects' incorrect responses actually reflect the perceived stimulus position on a trial-by-trial basis rather than simply reflecting errors in motor execution or random guesses (lapses). Although lapses undoubtedly resulted in some missed trials, if missed trials reflected only noise, then modeling that noise would have driven the position discrimination slopes toward zero rather than producing the dramatic improvements that we found in higher level areas. Indeed, the bootstrapping analysis presented in Figure 5 showed that in every higher level visual area we tested, subjects' actual responses were far better at predicting changes in the BOLD than in the random responses that were equally well correlated with the physical positions of the stimuli. Moreover, within each subject, we used the same set of five "percept" activity maps to test position discrimination in every ROI. Any influence of errors and guessing would be uniform across all of the areas we tested and could not account for the opposite effects that we found in control (early) areas versus higher level areas.

How do we know that some other, nonretinotopic coordinate frame does not better account for the results presented here? Could it be that object, face, and place regions code position information in saccade-centered, head-centered, world-centered, or some other physically defined coordinate frame? We can rule out any single physically defined coordinate frame because all of the precision improvements that we found in our perceived position analysis were driven by missed trials, in which subjects misperceived the stimulus location. Despite this, there was improved systematic coding in every high-level ROI we tested—objects perceived to be located in the same position, regardless of where they were physically located, produced the most similar patterns of activity. As noted above, perceived position coding must incorporate a combination of many coordinate frames.

Although the activity we measured in primary visual cortex (V1) was strongly tied to the retinal positions of the stimuli, we found that there was also some unique information about perceived position represented there (Figure 5). This is consistent with the results of a recent study by Murray, Boyaci, and Kersten (2006), which investigated the influence of perceived object size on representations in V1. The authors presented a stimulus of constant retinal size, but they manipulated its perceived size by changing the apparent depth of the object with contextual cues. They found that when the object was perceived as being larger, it activated a larger region of cortex in V1; the pattern of activation was similar to that observed for a physical increase in stimulus size. It is not ble to say whether the information about perceived size in V1 shown by Murray et al. (2006) and the information about perceived position in V1 found in the prese reflect feedback input from higher level areas that possess strong percept-centered coding or whether some preliminary percept-related information is computed in V1 itself. Although these findings and others (Tong, 2003) show that even V1 cannot be said to be strictly retinotopic, in our present results, activity in V1 encoded retinal position much more precisely than perceived position (Figure 5), and there was a steady accumulation of percept-centered information at higher stages in the visual processing hierarchy.

Consistent with the general principle that the receptive fields of individual neurons increase in size at successive stages of visual processing (Desimone & Duncan, 1995), we found steadily decreasing precision of physical position coding as we ascended the visual processing hierarchy. Could the superior coding of perceived stimulus position that we measured in higher level areas simply be a result of coarser retinotopic maps? We would actually expect just the opposite: the capacity for coding perceived position should also be degraded by increasing RF size, unless there are additional sources of information present that are contributing to perceived position coding. Our results show that the precision of retinotopic coding does not reflect an absolute limitation on the capacity of an area for carrying position information. higher level

areas can clearly support more precise position discrimination than is revealed by measuring retinotopic coding, when measured along the perceived position dimension, which incorporates additional information sources. In this sense, measuring physical position coding (e.g., retinotopy) only taps into a portion of the potential capacity for position information in higher level areas. This fact is most apparent in Figure 5: Although there is very little unique retinotopic information in higher level visual areas, there is substantial information about perceived position.

Higher level visual areas are, in general, more strongly modulated by attention (Maunsell, 2004). Could the pattern of results that we found be due to attentional influences that manifest at higher levels? As highlighted in the introduction, attention is indeed one of the factors that can displace an object's perceived position from its retinal position (Suzuki & Cavanagh, 1997), so we would expect attentional influences to contribute to the construction of a percept-centered framework. However, because we used the exact same trials to measure perceived versus physical position coding in our experiment, the influence of attention on the BOLD was the same for each analysis. As with the host of factors that contribute to perceived position, attention could only contribute to a difference between perceived and physical coding in our analysis by virtue of modulating the BOLD in a manner that is correlated with perceived rather than physical stimulus position. The combined influence of such factors is exactly what we aimed to measure in higher level visual areas.

The position discrimination analysis is designed to characterize the nature of position coding within a given ROI. For a single ROI, factors such as number of voxels, signal strength, study scanner noise, and motion artifacts are held constant across the physical and the percept analyses, so the comparison of physical and perceived position coding reflects only the difference in how the GLM predictors were coded. Across ROIs, those factors confound direct precision comparisons. For this reason, we focused on the relative precisions of perceived and physical coding within each visual area to show differences in the underlying nature of position coding across the visual processing hierarchy. Similarly, on an absolute scale, V1 showed more precise coding of perceived position than did any higher level area. However, the bootstrapped distribution in Figure 5 shows that we would expect strong discrimination of perceived position in V1 based solely on the combination of precise retinotopy and the correlation between subjects' responses and the physical stimulus positions (Figure 1). By considering the relative precision of physical versus percept coding ( $-(Z_{\text{percept}} - Z_{\text{physical}})$ ), it becomes clear that coding in V1 is strongly biased toward retinal position, whereas coding in higher level areas is dominated by percept information.

The cross-correlation analysis that we used in this study is a simple yet powerful way to measure stimulus discrimination within an ROI. Recently, a number of studies have demonstrated the promise of multivariate pattern analyses in detecting stimulus discrimination in the BOLD response (Bressler et al., 2007; Haynes & Rees, 2005; Kamitani & Tong, 2005; Carlson, Schrater, & He, 2003; Haxby et al., 2001). Although the sensitivity of all of these approaches derives from their multivariate nature, our correlation analysis is particularly well suited to studying position discrimination because it allows us to track systematic changes in the BOLD response corresponding to incremental, parametric manipulations of the stimulus location. When the correlation analysis yields a significant position discrimination fit within an ROI, it implies not only that different stimulus positions produced detectably different patterns of activity in that ROI but also that the similarity of the activity patterns was dependent on the similarity of the underlying stimuli. The analysis is sensitive to information encoded in complex coordinate frames, but it requires topography, not just different responses to different stimuli.

The perceived location of an object depends on many factors. For example, eye movements, gaze direction, scene and object motion, visual reference frames, and attention all influence

perceived position (Dassonville, Bridgeman, Kaur Bala, Thiem, & Sampanes, 2004; Nijhawan, 2002; Schlag & Schlag-Rey, 2002; Whitney, 2002; Ross et al., 1997; Suzuki & Cavanagh, 1997; Matin et al., 1982). Thus, percept-centered mapping in the visual cortex reflects the accumulation of a copious array of retinal and extraretinal information. In this study, we measured the aggregate information about perceived position in each ROI, but we cannot yet say whether such percept information was supported by a single, highly complex percept-based topographic map or the coexistence of many maps registered to each other but supported by distinct subpopulations of neurons. Complex spatial representations that reflect multiple coordinate frames have been found in parietal and motor cortex (Graziano, 2006; Graziano & Gross, 1998; Andersen, Snyder, Bradley, & Xing, 1997). Our results, demonstrating that position coding in higher level visual areas adapts to reflect an object's perceived position on a trial-by-trial basis, show that a similarly nuanced picture of multiplexed spatial maps is necessary to understand how object position is computed and encoded in the human visual system.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## REFERENCES

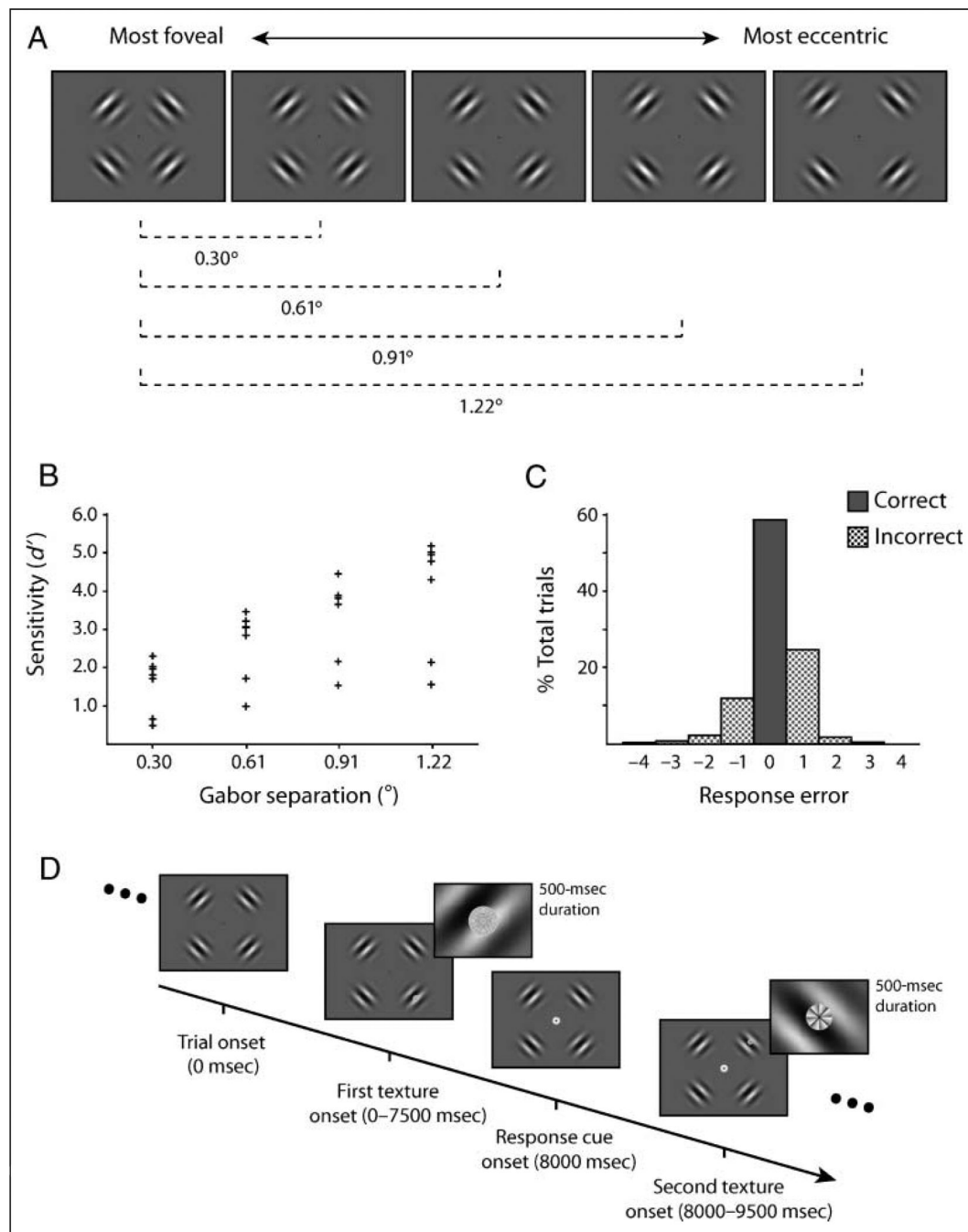
- Altmann CF, Deubelius A, Kourtzi Z. Shape saliency modulates contextual processing in the human lateral occipital complex. *Journal of Cognitive Neuroscience* 2004;16:794–804. [PubMed: 15200707]
- Andersen RA, Snyder LH, Bradley DC, Xing J. Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annual Review of Neuroscience* 1997;20:303–330.
- Avidan G, Hasson U, Hendler T, Zohary E, Malach R. Analysis of the neuronal selectivity underlying low fMRI signals. *Current Biology* 2002;12:964–972. [PubMed: 12123569]
- Benjamini Y, Hochberg Y. Controlling the false discovery rate—A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B, Methodological* 1995;57:289–300.
- Bressler D, Spotswood N, Whitney D. Negative BOLD fMRI response in the visual cortex carries precise stimulus-specific information. *PLoS ONE* 2007;2:e410. [PubMed: 17476332]
- Bridgeman B, Peery S, Anand S. Interaction of cognitive and sensorimotor maps of visual space. *Perception & Psychophysics* 1997;59:456–469. [PubMed: 9136275]
- Cai RH, Pouget A, Schlag-Rey M, Schlag J. Perceived geometrical relationships affected by eye-movement signals. *Nature* 1997;386:601–604. [PubMed: 9121582]
- Carlson TA, Schrater P, He S. Patterns of activity in the categorical representations of objects. *Journal of Cognitive Neuroscience* 2003;15:704–717. [PubMed: 12965044]
- Dassonville P, Bridgeman B, Kaur Bala J, Thiem P, Sampanes A. The induced Roelofs effect: Two visual systems or the shift of a single reference frame? *Vision Research* 2004;44:603–611. [PubMed: 14693187]
- d'Avossa G, Tosetti M, Crespi S, Biagi L, Burr DC, Morrone MC. Spatiotopic selectivity of BOLD responses to visual motion in human area MT. *Nature Neuroscience* 2007;10:249–255.
- De Valois RL, De Valois KK. Vernier acuity with stationary moving Gabors. *Vision Research* 1991;31:1619–1626. [PubMed: 1949630]
- Desimone R, Duncan J. Neural mechanisms of selective visual attention. *Annual Review of Neuroscience* 1995;18:193–222.
- Dukelow SP, DeSouza JF, Culham JC, van den Berg AV, Menon RS, Vilis T. Distinguishing subregions of the human MT+ complex using visual fields and pursuit eye movements. *Journal of Neurophysiology* 2001;86:1991–2000. [PubMed: 11600656]
- Dumoulin SO, Bittar RG, Kabani NJ, Baker CL Jr, Le Goualher G, Bruce Pike G, et al. A new anatomical landmark for reliable identification of human area V5/MT: A quantitative analysis of sulcal patterning. *Cerebral Cortex* 2000;10:454–463. [PubMed: 10847595]

- Eldridge LL, Knowlton BJ, Furmanski CS, Bookheimer SY, Engel SA. Remembering episodes: A selective role for the hippocampus during retrieval. *Nature Neuroscience* 2000;3:1149–1152.
- Epstein R, Graham KS, Downing PE. Viewpoint-specific scene representations in human parahippocampal cortex. *Neuron* 2003;37:865–876. [PubMed: 12628176]
- Epstein R, Harris A, Stanley D, Kanwisher N. The parahippocampal place area: Recognition, navigation, or encoding? *Neuron* 1999;23:115–125. [PubMed: 10402198]
- Fischer J, Whitney D. Attention narrows position tuning of population responses in V1. *Current Biology* 2009a;19:1356–1361. [PubMed: 19631540]
- Fischer J, Whitney D. Precise discrimination of object position in the human pulvinar. *Human Brain Mapping* 2009b;30:101–111. [PubMed: 17990302]
- Gardner JL, Merriam EP, Movshon JA, Heeger DJ. Maps of visual space in human occipital cortex are retinotopic, not spatiotopic. *Journal of Neuroscience* 2008;28:3988–3999. [PubMed: 18400898]
- Gauthier I, Tarr MJ, Anderson AW, Skudlarski P, Gore JC. Activation of the middle fusiform “face area” increases with expertise in recognizing novel objects. *Nature Neuroscience* 1999;2:568–573.
- Graziano MS. Progress in understanding spatial coordinate systems in the primate brain. *Neuron* 2006;51:7–9. [PubMed: 16815326]
- Graziano MS, Gross CG. Spatial maps for the control of movement. *Current Opinion in Neurobiology* 1998;8:195–201. [PubMed: 9635202]
- Grill-Spector K, Knouf N, Kanwisher N. The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience* 2004;7:555–562.
- Grill-Spector K, Kushnir T, Edelman S, Avidan G, Itzhak Y, Malach R. Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* 1999;24:187–203. [PubMed: 10677037]
- Grill-Spector K, Kushnir T, Hendler T, Edelman S, Itzhak Y, Malach R. A sequence of object-processing stages revealed by fMRI in the human occipital lobe. *Human Brain Mapping* 1998;6:316–328. [PubMed: 9704268]
- Grill-Spector K, Malach R. fMR-adaptation: A tool for studying the functional properties of human cortical neurons. *Acta Psychologica (Amsterdam)* 2001;107:293–321.
- Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 2001;293:2425–2430. [PubMed: 11577229]
- Haynes JD, Rees G. Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nature Neuroscience* 2005;8:686–691.
- Hemond CC, Kanwisher NG, Op de Beeck HP. A preference for contralateral stimuli in human object- and face-selective cortex. *PLoS ONE* 2007;2:e574. [PubMed: 17593973]
- Henson RN, Hornberger M, Rugg MD. Further dissociating the processes involved in recognition memory: An fMRI study. *Journal of Cognitive Neuroscience* 2005;17:1058–1073. [PubMed: 16102237]
- Holmes AP, Friston KJ. Generalisability, random effects and population inference. *Neuroimage* 1998;7:S754.
- Huk AC, Dougherty RF, Heeger DJ. Retinotopy and functional subdivision of human areas MT and MST. *Journal of Neuroscience* 2002;22:7195–7205. [PubMed: 12177214]
- Kamitani Y, Tong F. Decoding the visual and subjective contents of the human brain. *Nature Neuroscience* 2005;8:679–685.
- Kanwisher N, McDermott J, Chun MM. The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience* 1997;17:4302–4311. [PubMed: 9151747]
- Kerzel D. Attention maintains mental extrapolation of target position: Irrelevant distractors eliminate forward displacement after implied motion. *Cognition* 2003;88:109–131. [PubMed: 12711155]
- Kourtzi Z, Kanwisher N. Activation in human MT/MST by static images with implied motion. *Journal of Cognitive Neuroscience* 2000;12:48–55. [PubMed: 10769305]
- Larsson J, Heeger DJ. Two retinotopic visual areas in human lateral occipital cortex. *Journal of Neuroscience* 2006;26:13128–13142. [PubMed: 17182764]

- Levy I, Hasson U, Avidan G, Hendler T, Malach R. Center-periphery organization of human object areas. *Nature Neuroscience* 2001;4:533–539.
- MacEvoy SP, Epstein RA. Position selectivity in scene and object-responsive occipitotemporal regions. *Journal of Neurophysiology* 2007;98:2089–2098. [PubMed: 17652421]
- MacMillan, NA.; Creelman, CD. *Detection theory: A user's guide*. 2nd ed.. Erlbaum; Mahwah, NJ: 2004.
- Matin L, Picoult E, Stevens JK, Edwards MW Jr, Young D, MacArthur R. Oculoparalytic illusion: Visual-field dependent spatial mislocalizations by humans partially paralyzed with curare. *Science* 1982;216:198–201. [PubMed: 7063881]
- Maunsell, JHR. The role of attention in visual cerebral cortex.. In: Werner, JS.; Chalupa, LM., editors. *The visual neurosciences*. MIT Press; Cambridge, MA: 2004. p. 1538-1545.
- McKee SP, Silverman GH, Nakayama K. Precise velocity discrimination despite random variations in temporal frequency and contrast. *Vision Research* 1986;26:609–619. [PubMed: 3739236]
- McKyton A, Zohary E. Beyond retinotopic mapping: The spatial representation of objects in the human lateral occipital complex. *Cerebral Cortex* 2007;17:1164–1172. [PubMed: 16818474]
- Murray SO, Boyaci H, Kersten D. The representation of perceived angular size in human primary visual cortex. *Nature Neuroscience* 2006;9:429–434.
- Nijhawan R. Neural delays, visual motion and the flash-lag effect. *Trends in Cognitive Sciences* 2002;6:387. [PubMed: 12200181]
- Norman KA, Polyn SM, Detre GJ, Haxby JV. Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences* 2006;10:424–430. [PubMed: 16899397]
- Ramachandran VS, Anstis SM. Illusory displacement of equiluminous kinetic edges. *Perception* 1990;19:611–616. [PubMed: 2102995]
- Roelofs C. Optische localisation. *Arch Augenheilkunde* 1935;109:395–415.
- Ross J, Morrone MC, Burr DC. Compression of visual space before saccades. *Nature* 1997;386:598–601. [PubMed: 9121581]
- Ross J, Morrone MC, Goldberg ME, Burr DC. Changes in visual perception at the time of saccades. *Trends in Neurosciences* 2001;24:113–121. [PubMed: 11164942]
- Schlag J, Schlag-Rey M. Through the eye, slowly: Delays and localization errors in the visual system. *Nature Reviews Neuroscience* 2002;3:191–215.
- Schwarzlose RF, Swisher JD, Dang S, Kanwisher N. The distribution of category and location information across object-selective regions in human visual cortex. *Proceedings of the National Academy of Sciences, U.S.A* 2008;105:4447–4452.
- Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, et al. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 1995;268:889–893. [PubMed: 7754376]
- Snedecor, GW.; Cochran, WG. *Statistical methods*. 7th ed.. Iowa State University Press; Ames: 1980.
- Suzuki S, Cavanagh P. Focused attention distorts visual space: An attentional repulsion effect. *Journal of Experimental Psychology: Human Perception and Performance* 1997;23:443–463. [PubMed: 9104004]
- Talairach, J.; Tournoux, P. *Co-planar stereotaxic atlas of the human brain: 3-Dimensional proportional system: An approach to cerebral imaging*. Thieme Medical Publishers; New York: 1988.
- Tong F. Primary visual cortex and visual awareness. *Nature Reviews Neuroscience* 2003;4:219–229.
- Tootell RB, Reppas JB, Kwong KK, Malach R, Born RT, Brady TJ, et al. Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *Journal of Neuroscience* 1995;15:3215–3230. [PubMed: 7722658]
- Tootell RB, Tsao D, Vanduffel W. Neuroimaging weighs in: Humans meet macaques in “primate” visual cortex. *Journal of Neuroscience* 2003;23:3981–3989. [PubMed: 12764082]
- Volkman J. The method of single stimuli. *American Journal of Psychology* 1932;44:808–809.
- Watson JD, Myers R, Frackowiak RS, Hajnal JV, Woods RP, Mazziotta JC, et al. Area V5 of the human brain: Evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cerebral Cortex* 1993;3:79–94. [PubMed: 8490322]
- Whitaker D, McGraw PV. The effect of suprathreshold contrast on stimulus centroid and its implications for the perceived location of objects. *Vision Research* 1998;38:3591–3599. [PubMed: 9893792]



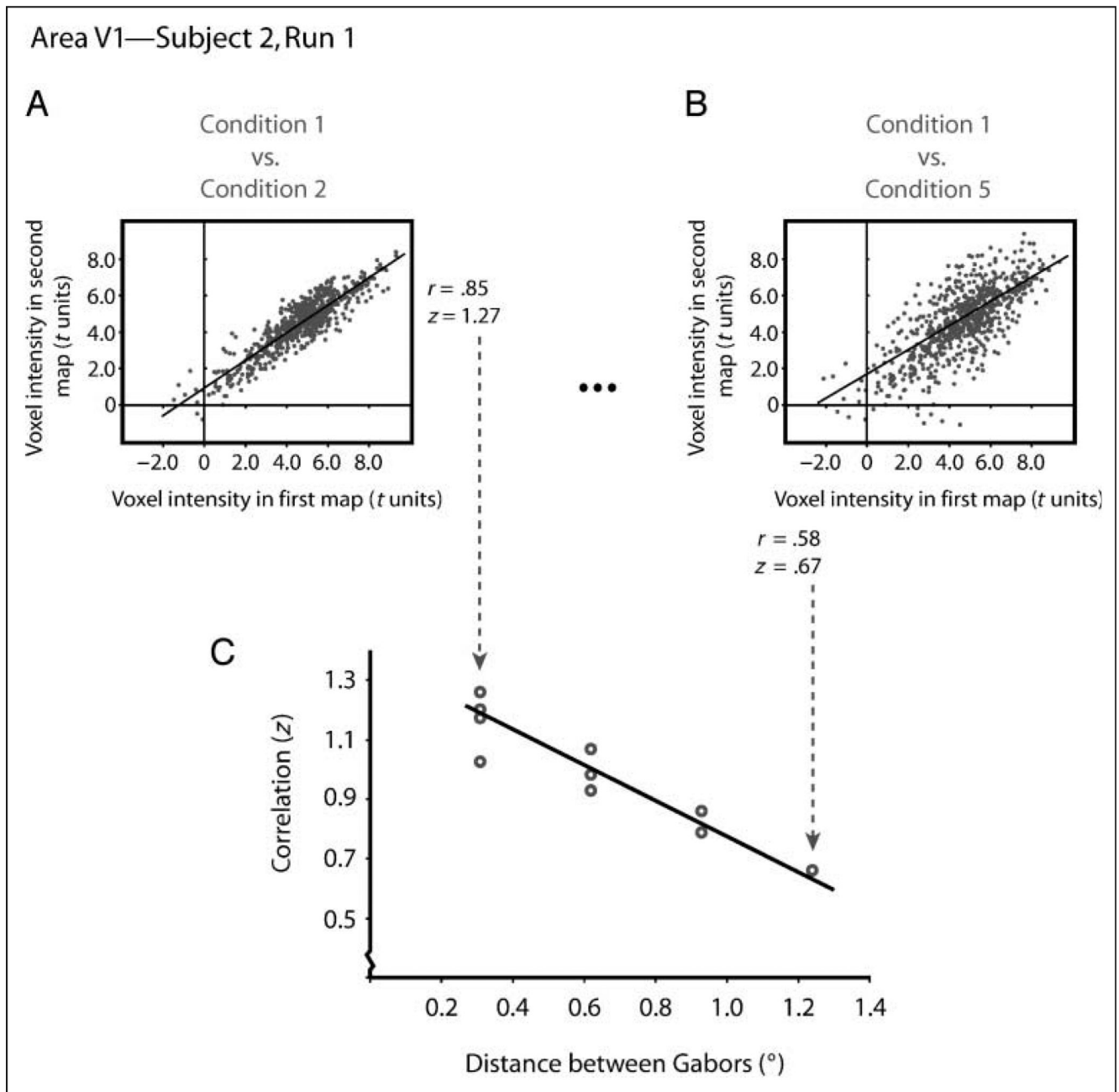
- Whitaker D, McGraw PV, Levi DM. The influence of adaptation on perceived visual location. *Vision Research* 1997;37:2207–2216. [PubMed: 9578903]
- Whitaker D, McGraw PV, Pacey I, Barrett BT. Centroid analysis predicts visual localization of first- and second-order stimuli. *Vision Research* 1996;36:2957–2970. [PubMed: 8917796]
- Whitney D. The influence of visual motion on perceived position. *Trends in Cognitive Sciences* 2002;6:211–216. [PubMed: 11983584]
- Whitney D, Bressler DW. Spatially asymmetric response to moving patterns in the visual cortex: Re-examining the local sign hypothesis. *Vision Research* 2007;47:50–59. [PubMed: 17049580]
- Yi DJ, Kelley TA, Marois R, Chun MM. Attentional modulation of repetition attenuation is anatomically dissociable for scenes and faces. *Brain Research* 2006;1080:53–62. [PubMed: 16507300]



**Figure 1.**

Main experimental stimuli and behavioral results. (A) On each trial, Gabor stimuli appeared at one of five possible eccentricity conditions, ranging from  $8.43^\circ$  to  $9.65^\circ$  from a central fixation point. The range of eccentricities pictured is exaggerated for visualization purposes. The Gabors flickered in counterphase at 7.5 Hz and were present for the duration of each 10 second trial. A white annulus around the fixation point during the last 2 sec of each trial cued subjects to respond, indicating which of the five conditions was present. In a sixth baseline condition, only the fixation point was present for 10 sec. The behavioral and the correlation analyses refer to the separations between the five stimulus conditions—these are indicated below the stimuli (in degrees visual angle). (B) Subjects' sensitivity (in  $d'$  units, see MacMillan

& Creelman, 2004) is plotted against stimulus separation. The positive trend in this plot indicates that the five stimulus positions sampled the dynamic range of subjects' discrimination sensitivity. The overall hit rate was 58%. (C) Histogram of the response errors for all trials in the main experiment, binned by the distance between the correct and the actual responses. For example, if, on a given trial, the Gabors were presented in position 4 ( $9.34^\circ$  from fixation) but the subject responded with position 3, that trial is binned in response error  $(-1)$ . The trials in solid gray at response error 0 are the trials in which subjects responded correctly; we exclusively used the *missed* trials (dotted bars; 42% of total trials), in which perceived and physical positions were dissociated in the analysis of the main experiment. (D) An example trial. The Gabor stimuli were present for the duration of each 10-sec trial. At a random time during the first 8 sec of each trial, a small texture (either a radial or a concentric grating) appeared for 500 msec at  $9.04^\circ$  from fixation in a random quadrant. During the last 2 sec of the trial, a second texture appeared at the same eccentricity in a random quadrant; the type of texture matched the first on 50% of the trials. At the end of the trial, subjects gave two responses, indicating the positions of the Gabors (5AFC) and whether the two presented textures were the same (2AFC).

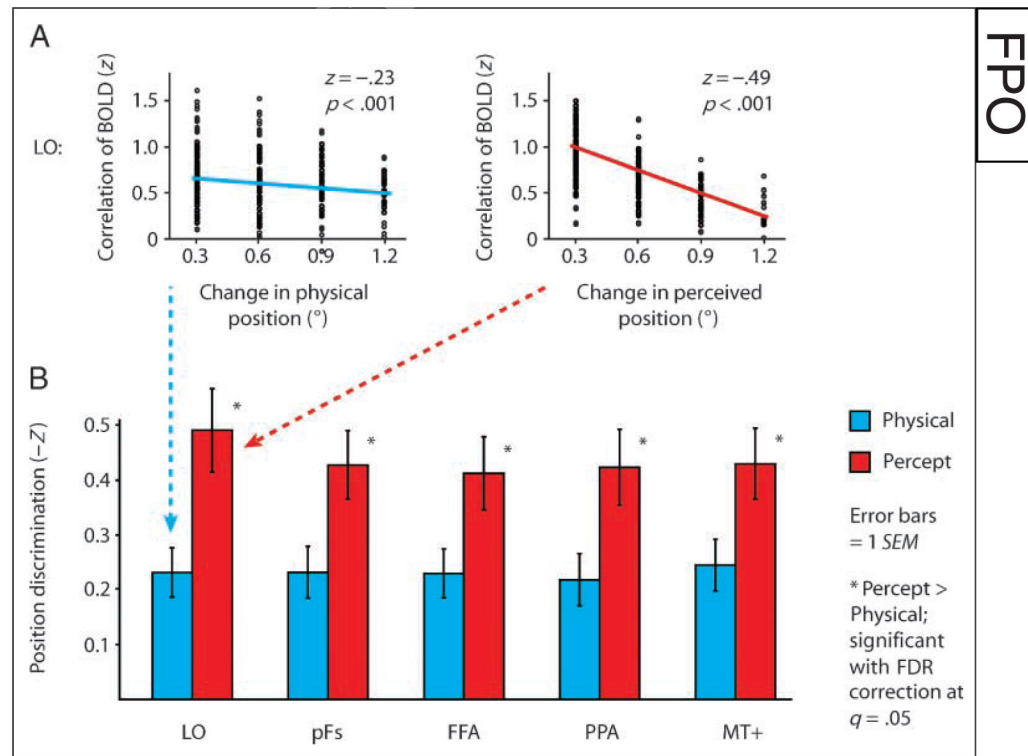


**Figure 2.**

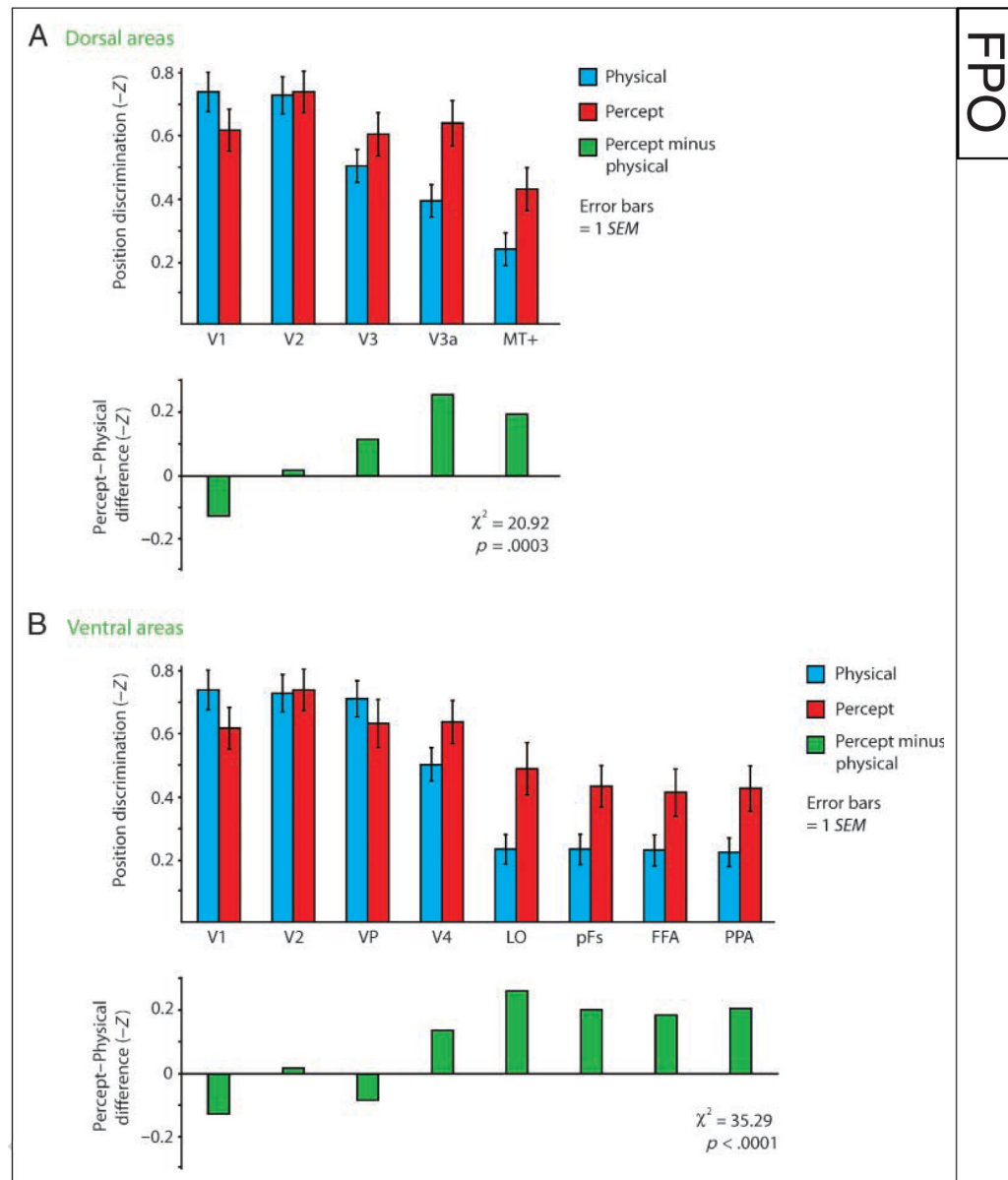
Construction of a position discrimination plot. To measure position selectivity within an ROI, we created a position discrimination plot, in which the similarity between the patterns of activity produced by any two of the five stimulus conditions is plotted against the spatial separation between the stimuli presented in those conditions. We used the correlation (Pearson's  $r$ ) between two patterns of BOLD response as the measure of their similarity. To compute the correlation between a given pair of activity maps, we plotted the intensity values ( $t$  units) from one map against those from the other map on a voxel-by-voxel basis and fit a linear regression to the plot. We transformed the resulting  $r$  value to a Fisher  $z$  to allow for a linear comparison among multiple correlations measured in this way. Two such correlations computed within V1 for an example subject are shown in panels A and B. The plot in panel A

corresponds to two adjacent stimulus conditions, whereas the plot in panel B corresponds to the two furthest separated stimulus eccentricities. Note that the correlation in panel A is substantially stronger than that in panel B. The fact that retinotopically proximal stimuli produce similar patterns of BOLD response whereas more distant stimuli produce less similar patterns of BOLD is an indication that activity in the ROI is position selective. To evaluate the precision of position selectivity in the BOLD response, we plotted each of the 10 correlations against the distance between the stimuli that produced it and fit a linear regression to the resulting plot (C). Panel C shows the correlations from a single run; we computed the correlations on a run-by-run basis, so a full position discrimination plot for one subject had fifty total points (see Methods). A significant negative trend in this position discrimination plot indicates that activity in the ROI is sensitive to the parametric position manipulation (Fischer & Whitney, 2009a, 2009b; Bressler et al., 2007).



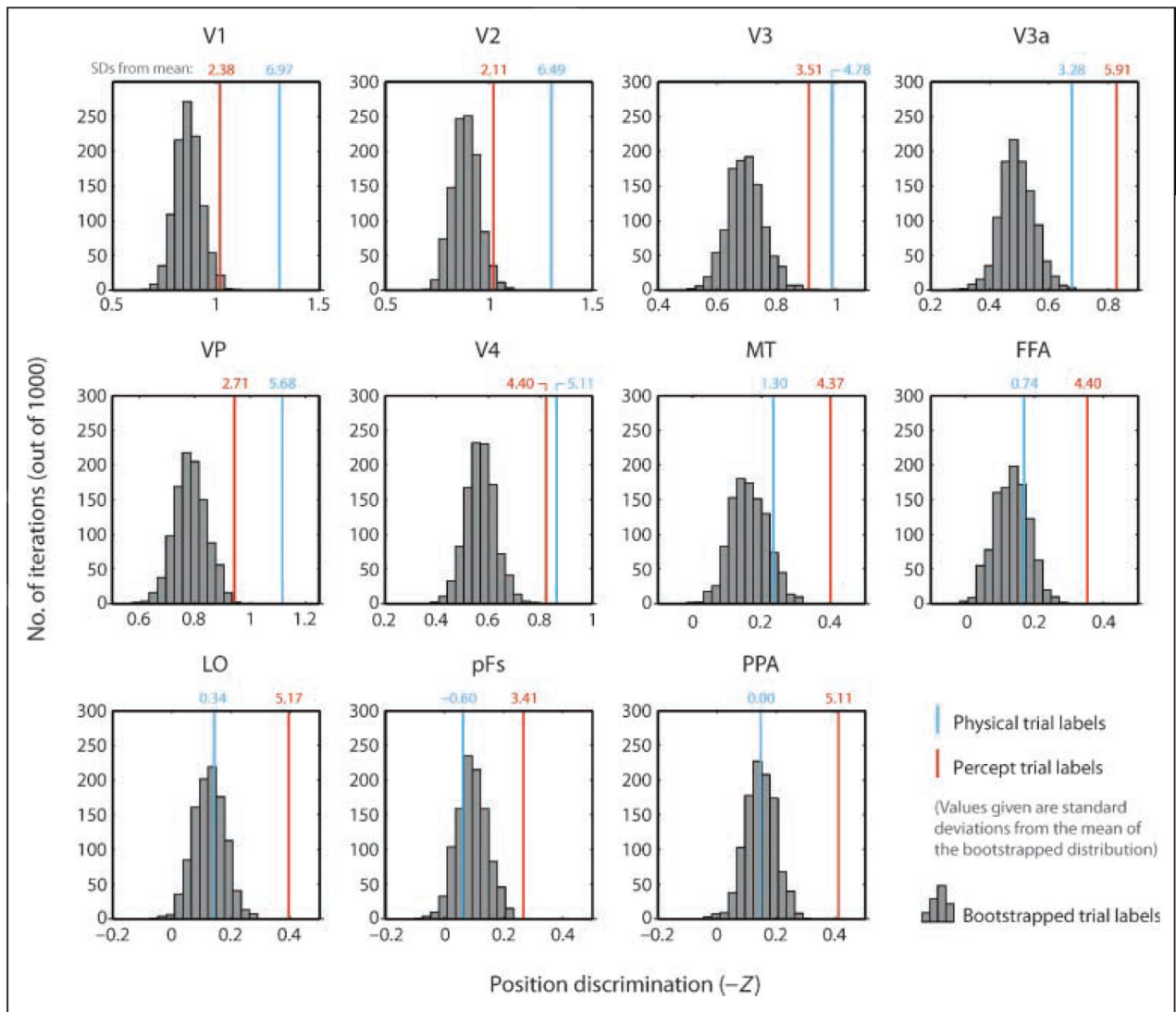
**Figure 3.**

Comparison of physical and perceived position discrimination in higher level visual areas. (A) For each ROI, we constructed a position discrimination plot, as outlined in Figure 2, separately for physical position (panel A, left) and perceived position (panel A, right). These plots are based on the exact same trials; the only difference between the two is whether the trials were coded according to their physical position, or according to subjects' responses, in the GLM analyses. The plots in panel A show data from all subjects; to perform a group-level analysis, we fit a linear regression to all subjects' data taken together and included a random effect of subject in the regression model to account for between-subject variance (see Methods). The goodness of fit of the linear regression captures how tightly changes in BOLD were coupled with changes in physical or perceived stimulus position, hence serving as an index of the precision of position coding. (B) The precision of physical (blue) and perceived (red) position coding is plotted side by side for each of the five higher level visual areas. Y axis units are in  $-Z$ , so a taller bar indicates a more strongly negative correlation on the position discrimination plot. Error bars indicate  $\pm 1$  SEM. Coding of perceived position was significantly more precise than coding of physical position in every higher level visual area we tested (see Table 1 for statistics).

**Figure 4.**

Comparison of position discrimination in higher level visual with that in early visual areas. (A) Precision of physical (blue) and perceived (red) position discrimination is shown for the dorsal visual areas in ascending order. Coding preference is captured in the green bars: they show the within-area comparison of physical versus perceived position discrimination for each visual area, obtained by subtracting the precision of physical coding from that of percept coding  $-(Z_{\text{percept}} - Z_{\text{physical}})$ . Positive values in the green bars indicate a preferential coding of perceived position, whereas negative values indicate a preferential coding of retinal position. Although activity in MT+ reflects perceived position more precisely than physical position, in earlier areas this bias is diminished or reversed. A chi-square test revealed that the nature of position coding differed significantly among these dorsal areas ( $\chi^2_{\text{dorsal}} = 20.92, p = .0003$ ). Panel B shows a comparison of perceived and physical position discrimination in the ventral visual stream. Here, too, there was significant variability between areas in the bias for representing physical or perceived position ( $\chi^2_{\text{ventral}} = 35.29, p < .0001$ ). To test for a systematic

progression in the nature of position coding across areas, we ranked all 11 areas according to their locations in the visual processing stream (see Methods) and computed a Spearman rho rank correlation between the visual area ordering and the position coding bias  $-(Z_{\text{percept}} - Z_{\text{physical}})$ . The correlation was highly significant ( $\rho = .80$ ;  $p = .003$ ). To test this a priori ordering against all other possible rankings of visual areas, we computed a 25,000-sample bootstrapped distribution of rank correlations, with a randomly drawn ranking of areas for each sample. The correlation of .80 obtained with the a priori ranking was larger in absolute value than 99.5% of the bootstrapped samples, indicating that our ranking based on functional anatomy is a good match to the independently measured position discrimination estimates for each area. The strong correlation between an area's position in the visual processing stream and its bias for representing physical or perceived position reinforces the idea that the nature of position coding evolves as information progresses through the visual processing hierarchy, becoming relatively more strongly tied to perception in higher level areas.

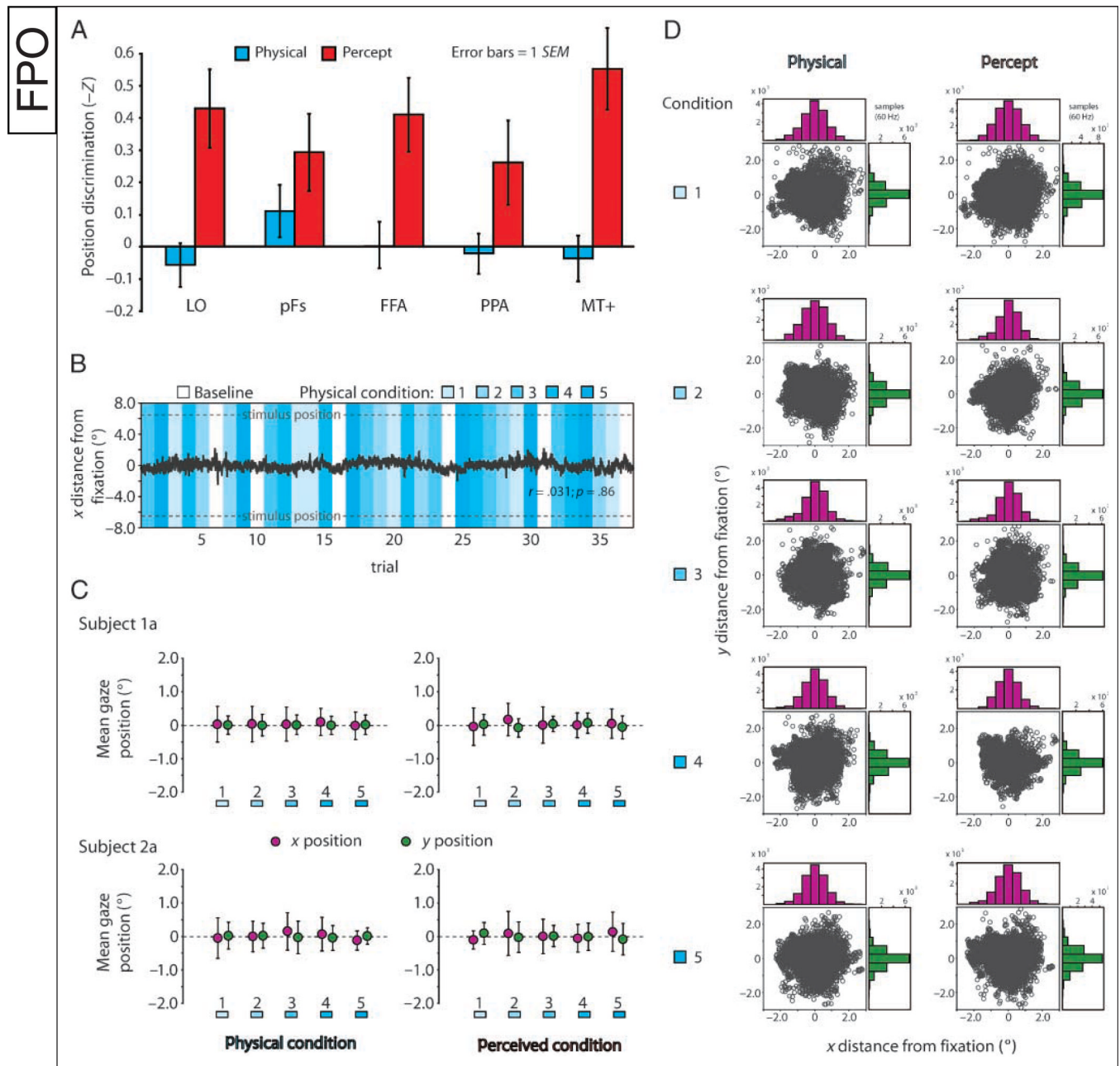


**Figure 5.**

Bootstrapped position discrimination analysis. To evaluate how uniquely predictive subjects' responses were of changes in the pattern of BOLD response, relative to other possible codings of the GLM predictors, we performed the position discrimination analysis on a bootstrapped sample of 1000 sets of trial labels per subject. To generate each bootstrapped sample, we started with the physical stimulus positions and added "errors" sampled from the distribution of errors that subjects actually made during the experiment (Figure 1C). Thus, over the course of all 1000 iterations, the average errors in the GLM design matrices matched the distribution in Figure 1C. On each iteration, we obtained a group-wise position discrimination score ( $-Z$ ) for each ROI; the collected position discrimination scores are plotted in the gray histograms. Of particular interest was whether subjects' responses performed better at predicting the pattern of BOLD in higher level areas than did the random errors in the bootstrapped trial labels, which were, on average, equally well correlated with the physical stimulus positions as subjects' responses were. In fact, in every higher level area, the percept encoding was an extreme outlier from the bootstrapped distribution (least significant was pFs;  $-Z = 3.41$ ,  $p < .001$ ), indicating that the precise errors that subjects made were uniquely predictive of changes in the pattern of BOLD in higher level areas. Likewise, the fact that the physical trial labels were extreme

outliers from the bootstrapped distributions in lower level areas indicated that even a slight perturbation of the precise retinotopic labeling of the trials resulted in a dramatic decrease in the ability to predict changes in the pattern of BOLD response.





**Figure 6.**

Control analysis with eye tracking during scanning. (A) Position discrimination data for two additional subjects scanned on the same task as in the main experiment, with eye tracking. Error bars represent  $\pm 1$  SEM. The data are consistent with the results of the main experiment; the position discrimination estimates from these subjects fell within the range of estimates acquired for subjects in the main experiment in each visual area (most significant differences were MT+:  $z_{\text{physical}} = 1.81, p = .07$  and PPA:  $Z_{\text{percept}} = 0.87, p = .39$ , for physical and percept discrimination scores, respectively). Discrimination of physical position hovered near zero for these subjects; this pattern is not atypical compared with the individual subject data from the main experiment, although there were also subjects that showed substantial discrimination of physical position (see Supplementary Figure 3). (B) Sample eye trace for one run from Subject

1a. The  $x$  position of gaze (sampled at 60 Hz) is plotted for the duration of the run, and the presentation of the five position conditions is indicated behind the trace in shades of blue. The correlation between eye position and the stimulus conditions for this run was  $r = .031$ ,  $p = .86$ . The largest correlation for any run was  $r = .048$ ,  $p = .78$ . The mean stimulus position, indicated with gray dashed lines, was at  $\pm 6.4^\circ$  from fixation. (C) Mean values for the  $x$  position (purple) and  $y$  position (green) are shown for each of the five physical positions (left plots) and perceived positions (right plots) for the two control subjects. Eye position was not correlated with the stimulus position or the responses made for either subject (see Results for statistics). There was also no correlation between physical or perceived position and variability of eye position (see Results for statistics). (D) Collections of the recorded gaze positions corresponding to each of the physical and perceived conditions for Subject 1a. The fixation point fell at position (0,0) in the center of each plot; the scatterplots show the collected position measurements sampled during the presentation of each condition (physical plots on the left) as well as the collected eye positions corresponding to each reported condition (percept plots on the right). The adjoining histograms show the  $x$  and  $y$  distributions of recorded eye positions composing each scatterplot.

**Table 1**

The Precision of Physical and Perceived Position Coding in Each higher level ROI

	Physical Position Discrimination ( $-Z$ )	Perceived Position Discrimination ( $-Z$ )	Z Test (Percept > Physical)	p
LO	.231	.487	-2.90	.0038*
pFs	.229	.426	-2.23	.0259*
FFA	.227	.409	-2.06	.0396*
PPA	.217	.424	-2.34	.0195*
MT +	.241	.428	-2.12	.0339*

The precision estimates were computed by fitting a linear regression to the position discrimination plot (grouped data) of each ROI. An  $r$  value significantly different from zero indicates significant position selectivity within the ROI, and a more negative  $r$  implies more precise position coding (Fischer & Whitney, 2009b). We applied a Fisher  $z$ -transform to each precision estimate ( $r$  value) for the sake of linear comparison, and we present the precision estimates in negative  $z$  units so that larger values indicate more precise position selectivity. We performed a within-area comparison of perceived and physical position coding for each ROI using a Z test (see Methods). Encoding of perceived position was more precise than that of physical position in every higher level area we tested. We corrected for multiple comparisons by controlling the false discovery rate to 0.05 (Benjamini & Hochberg, 1995).

\* Significant with FDR correction for multiple comparisons at  $q = 0.05$ .