

# UC San Diego

## UC San Diego Previously Published Works

### Title

Feature-Selective Attentional Modulations in Human Frontoparietal Cortex

### Permalink

<https://escholarship.org/uc/item/0vn1r536>

### Journal

Journal of Neuroscience, 36(31)

### ISSN

0270-6474

### Authors

Ester, Edward F  
Sutterer, David W  
Serences, John T  
et al.

### Publication Date

2016-08-03

### DOI

10.1523/jneurosci.3935-15.2016

Peer reviewed

# Feature-Selective Attentional Modulations in Human Frontoparietal Cortex

Edward F. Ester,<sup>1,2</sup> David W. Sutterer,<sup>1,4</sup> John T. Serences,<sup>2,3</sup> and Edward Awh<sup>1,4</sup>

<sup>1</sup>Department of Psychology, University of Oregon, Eugene, Oregon 97403, <sup>2</sup>Department of Psychology and <sup>3</sup>Neurosciences Graduate Program, University of California, San Diego, La Jolla, California 92093, and <sup>4</sup>Department of Psychology and Institute for Mind and Biology, University of Chicago, Chicago, Illinois 60637

Control over visual selection has long been framed in terms of a dichotomy between “source” and “site,” where top-down feedback signals originating in frontoparietal cortical areas modulate or bias sensory processing in posterior visual areas. This distinction is motivated in part by observations that frontoparietal cortical areas encode task-level variables (e.g., what stimulus is currently relevant or what motor outputs are appropriate), while posterior sensory areas encode continuous or analog feature representations. Here, we present evidence that challenges this distinction. We used fMRI, a roving searchlight analysis, and an inverted encoding model to examine representations of an elementary feature property (orientation) across the entire human cortical sheet while participants attended either the orientation or luminance of a peripheral grating. Orientation-selective representations were present in a multitude of visual, parietal, and prefrontal cortical areas, including portions of the medial occipital cortex, the lateral parietal cortex, and the superior precentral sulcus (thought to contain the human homolog of the macaque frontal eye fields). Additionally, representations in many—but not all—of these regions were stronger when participants were instructed to attend orientation relative to luminance. Collectively, these findings challenge models that posit a strict segregation between sources and sites of attentional control on the basis of representational properties by demonstrating that simple feature values are encoded by cortical regions throughout the visual processing hierarchy, and that representations in many of these areas are modulated by attention.

**Key words:** frontoparietal cortex; functional neuroimaging; visual attention; visual cortex

## Significance Statement

Influential models of visual attention posit a distinction between top-down control and bottom-up sensory processing networks. These models are motivated in part by demonstrations showing that frontoparietal cortical areas associated with top-down control represent abstract or categorical stimulus information, while visual areas encode parametric feature information. Here, we show that multivariate activity in human visual, parietal, and frontal cortical areas encode representations of a simple feature property (orientation). Moreover, representations in several (though not all) of these areas were modulated by feature-based attention in a similar fashion. These results provide an important challenge to models that posit dissociable top-down control and sensory processing networks on the basis of representational properties.

## Introduction

Behavioral (Treisman and Gelade, 1980; Wolfe, 1994), electrophysiological (Müller et al., 2006; Andersen et al., 2008; Zhang and Luck, 2009), and functional neuroimaging studies (Corbetta

et al., 1990; Chawla et al., 1999; Saenz et al., 2002; Liu et al., 2003, 2007; Polk et al., 2008) indicate that attention can selectively enhance representations of task-relevant features (e.g., color, orientation, shape, or direction) regardless of their location(s) in a visual scene. This form of feature-based attention (FBA) is integral to many everyday tasks, as we often know the defining features of a target (e.g., my coffee mug is green) but not its location (e.g., my coffee mug is on the table next to my desk).

Invasive electrophysiological recordings in nonhuman primates (McAdams and Maunsell, 1999; Treue and Martinez-Trujillo, 1999; Martinez-Trujillo and Treue, 2004) and functional neuroimaging studies in humans (Serences and Boynton, 2007; Jehee et al., 2011) suggest that FBA enhances cortical representations of behaviorally relevant visual features in early visual areas. These enhancements are

Received Oct. 30, 2015; revised May 18, 2016; accepted June 20, 2016.

Author contributions: E.F.E., J.T.S., and E.A. designed research; E.F.E. and D.W.S. performed research; E.F.E. analyzed data; E.F.E., J.T.S., and E.A. wrote the paper.

This work was supported by National Institutes of Health Grant R01 MH087214 (E.A.).

Correspondence should be addressed to either of the following: Edward F. Ester, Department of Psychology, University of California, San Diego, 9500 Gilman Drive, MC 0109, La Jolla, CA 92093. E-mail: eester@ucsd.edu; or Edward Awh, Department of Psychology and Institute for Mind and Biology, University of Chicago, 940 East 57th Street, Chicago, IL 60637.

DOI:10.1523/JNEUROSCI.3935-15.2016

Copyright © 2016 the authors 0270-6474/16/368188-12\$15.00/0

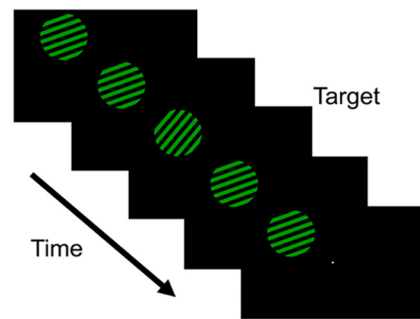
thought to result from top-down feedback signals originating in frontoparietal cortical areas (Serences et al., 2004; Kelley et al., 2008; Greenberg et al., 2010; Zhou and Desimone, 2011; Baldauf and Desimone, 2014; Gregoriou et al., 2014). For example, Baldauf and Desimone (2014) showed human participants displays containing semitransparent, spatially overlapping images of faces and houses. Magnetoencephalographic recordings revealed increased gamma band synchrony between portions of the inferior frontal junction (IFJ) and the location-selective parahippocampal place area (PPA) when participants attended the house image and increased synchrony between the IFJ and the face-selective fusiform face area (FFA) when participants attended the face image. Moreover, gamma phases were advanced in the IFJ relative to either FFA or PPA, suggesting that this region was the driver of changes in synchrony.

At present, it is unclear whether frontoparietal cortical areas implicated in attentional control also encode parametric sensory representations and, if so, whether these representations vary with behavioral relevance. On the one hand, single-unit recording studies suggest that many frontoparietal cortical areas encode task-level variables, such as decision criteria (Kim and Shadlen, 1999) or abstract rules (Wallis et al., 2001). However, other studies have observed parametric sensory representations in frontoparietal cortical areas during perception and working-memory storage (Buschman et al., 2011; Meyer et al., 2011; Mendoza-Halliday et al., 2014; Ester et al., 2015). Moreover, recent evidence from primate and rodent electrophysiology suggests that frontoparietal cortical areas may encode both task variables and parametric sensory representations in a high-dimensional state space (Mante et al., 2013; Rigotti et al., 2013; Raposo et al., 2014).

Motivated by these findings, the current study was designed to examine whether frontoparietal cortical areas typically implicated in attentional control contain continuous or categorical representations of task-relevant sensory parameters (e.g., orientation) and, if so, whether these representations are modulated by FBA. Using functional neuroimaging, we combined a roving “searchlight” analysis (Kriegeskorte et al., 2006) with an inverted encoding model (Brouwer and Heeger, 2009, 2011) to reconstruct and quantify representations of orientation in local neighborhoods centered on every gray matter voxel in the human cortical sheet while participants attended either the orientation or luminance of a stimulus. We observed robust representations of orientation in multiple frontal and parietal cortical areas previously associated with top-down control. Moreover, representations in many—though not all—of these regions were stronger (higher amplitude) when participants were instructed to attend orientation relative to when they were instructed to attend luminance. Collectively, our results indicate that several frontoparietal cortical regions typically implicated in top-down control also encode simple feature properties, such as orientation, and that representations in many of these regions are subject to attentional modulations similar to those seen in posterior visual areas. These results challenge models of selective attention that dissociate “top-down control” from “sensory processing” regions based on the type of information that they encode.

## Materials and Methods

**Participants.** Twenty-one neurologically intact volunteers from the University of Oregon (ages 19–33 years, nine females) participated in a single 2 h scanning session. All participants reported normal or corrected-to-normal visual acuity and were remunerated at a rate of \$20/h. All experimental procedures were approved by the local institutional review board, and all participants gave both written and oral informed consent. Data from three participants were discarded due to excessive head-



**Figure 1.** Stimulus displays. Participants viewed displays containing a single square-wave grating in the upper left or right visual field. On each trial, the grating was assigned one of nine orientations (0–160° in 20° increments) and one of two colors (yellow or green). In separate scans, participants were instructed to attend either the orientation or luminance of the grating. During attend-luminance scans, participants discriminated the direction (clockwise or anti-clockwise) of brief and unpredictable perturbations in stimulus orientation. During attend-orientation scans, participants discriminated the direction (brighter or dimmer) of brief and unpredictable perturbations in stimulus luminance.

motion artifacts (translation or rotation >2 mm in >25% of scans); the data reported here reflect the remaining 18 participants.

**Experimental setup.** Stimulus displays were generated in Matlab using Psychophysics Toolbox software (Brainard, 1997; Pelli, 1997) and back-projected onto a screen located at the base of the magnet bore. Participants were positioned ~58 cm from the screen and viewed stimulus displays via a mirror attached to the MR head coil. Behavioral responses were made using an MR-compatible button box.

**Behavioral tasks.** Participants viewed displays containing a full-contrast, square-wave grating (5° radius, 1 cycle/°) in the upper left or right visual field (horizontal and vertical eccentricity of ±7° and +5° relative to fixation, respectively; Fig. 1). On each trial, the grating was rendered in one of two colors (green or yellow) and assigned one of nine orientations (0–160° in 20° increments). The grating flickered at 3 Hz (i.e., 167 ms on, 167 ms off) for the entire duration of each 10 s trial. The spatial phase of the grating was randomized on every cycle to attenuate the potency of retinal afterimages. In separate scans, participants were instructed to attend either the luminance or orientation of the grating (here, “scan” refers to a continuous block of 36 trials lasting 432 s). Stimulus color and orientation were fully crossed within each scan. Trials were separated by a 2 s blank interval.

During attend-orientation scans, participants were required to discriminate the direction (clockwise or anticlockwise) of brief (1 stimulus cycle) and temporally unpredictable perturbations in stimulus orientation. A total of four perturbations occurred on each trial, with the constraint that no targets appeared during the first and last second of each trial, and that targets were separated by ≥2 stimulus cycles (i.e., 667 ms). During attend-luminance scans, participants were required to report small upward (i.e., “brighter”) or downward (i.e., “dimmer”) perturbations in stimulus luminance. Changes in luminance were subject to the same constraints as the attend-orientation task. Each participant completed 3–5 scans in the attend-orientation and attend-luminance tasks.

To ensure that both tasks were sufficiently challenging, we computed orientation and luminance discrimination thresholds for each participant during a separate 1 h behavioral testing session completed 1–3 d before scanning. Participants performed the luminance and orientation discrimination tasks described above, but the magnitudes of luminance and orientation perturbations were continually adjusted using a “three up, one down” adaptive staircase until a criterion accuracy of 75% was reached. The resulting thresholds were used to control the magnitudes of orientation and luminance perturbations and remained constant during scanning.

**fMRI data acquisition and preprocessing.** fMRI data were collected using a 3T Siemens Allegra system located at the Robert and Beverly Lewis Center for Neuroimaging at the University of Oregon. We acquired whole-brain echo-planar images (EPIs) with a voxel size of 3 × 3 × 3.5 mm (33 transverse slices with no gap, 192 × 192 field of view, 64 × 64

image matrix, 90° flip angle, 2000 ms repetition time, 30 ms echo time). EPIs were slice-time corrected, motion corrected (within and between scans), high-pass filtered (3 cycles/scan, including linear trend removal), and aligned to a T1-weighted anatomical scan (1 × 1 × 1 mm resolution) collected during the same scanning session. Preprocessed EPIs were transformed to Talairach space. Finally, the entire time series of each voxel in retinotopically organized visual areas V1, V2v, V3v, and hV4v (see below, Retinotopic mapping) and searchlight neighborhood (see below, Searchlight analysis) was normalized (z-score) on a scan-by-scan basis.

**fMRI eye tracking.** Continuous eye-position data were recorded for seven participants via an MR-compatible ASL 6000 infrared eye-tracking system and digitized at 60 Hz. The eyetracker was calibrated at the beginning of each scan session and recalibrated as needed (e.g., following a loss of the corneal reflection; this typically occurred 0–1 times per scan session). Recordings were filtered for blinks and corrected for linear drift off-line.

**Retinotopic mapping.** Each participant completed a single retinotopic mapping scan lasting 480 s. Visual areas V1, V2v, V3v, and hV4v were identified using standard procedures (Sereno et al., 1995). Participants fixated a small dot at fixation while a phase-reversing (8 Hz) checkerboard wedge subtending 60° of polar angle rotated counterclockwise around the display (period, 60 s). Visual field borders in areas V1, V2v, V3v, and hV4v were identified via a cross-correlation analysis. These data were projected onto a computationally inflated representation of each participant's gray–white matter boundary, and visual field borders were drawn by hand.

**Inverted encoding model.** We used an inverted encoding model to reconstruct representations of stimulus orientation from multivoxel responses in ROIs throughout the cortex. This approach rests on the assumptions that (1) the measured response in a given voxel is a linear sum of underlying neural activity, and (2) at least some of the voxels within an ROI exhibit a nonuniform response profile across orientations (Kamitani and Tong, 2005; Freeman et al., 2011).

We first extracted and averaged the (normalized) responses of each voxel in each visual area or searchlight neighborhood (see below, Searchlight analysis) over a period from 6–10 s after the start of each trial. This specific window was chosen to account for a typical hemodynamic lag of 6 s, but all results reported here generalized across other temporal windows (e.g., 4–10 s or 4–8 s after the start of each trial). Data were averaged across samples and sorted into one of 18 bins based on stimulus orientation (0–160° in 20° increments) and task (attend orientation or attend luminance). We next divided the data into “training” and “test” sets and modeled the measured responses of each voxel in the training set as a linear sum of nine orientation “channels,” each with an idealized response function. Following the terminology of Brouwer and Heeger (2009) let  $B_1$  ( $m$  voxels ×  $n$  trials) be the observed signal in each voxel in each trial,  $C_1$  ( $k$  channels ×  $n$  trials) be a matrix of predicted responses for each information channel on each trial, and  $W$  ( $m$  voxels ×  $k$  channels) be a weight matrix that characterizes the mapping from “channel space” to “voxel space.” The relationship between  $B_1$ ,  $C_1$ , and  $W$  can be described by a general linear model of the following form (Eq. 1):  $B_1 = WC_1$ , where  $C_1$  is a design matrix that contains the predicted response of each channel on each trial. Channel responses were modeled as nine half-wave rectified sinusoids raised to the eighth power and centered at the orientation of the stimulus on each trial (e.g., 0, 20, 40°, etc.). These functions were chosen because they approximate the shape of single-unit orientation tuning functions in V1, where the half-bandwidth of orientation-selective cells has been estimated to be ~20° (though there is substantial variability in bandwidth; Ringach et al., 2002; Gur et al., 2005). Moreover, they act as “steerable filters” that support the computation of channel responses for any possible orientation (Freeman and Adelson, 1991).

Given  $B_1$  and  $C_1$ , we estimated the weight matrix  $\hat{W}$  ( $m$  voxels ×  $k$  channels) using ordinary least-squares regression as follows (Eq. 2):

$$\hat{W} = B_1 C_1^T (C_1 C_1^T)^{-1}$$

Given these weights and voxel responses observed in an independent “test” dataset, we inverted the model to transform the observed test data  $B_2$  ( $m$  voxels ×  $n$  trials) into a set of estimated channel responses,  $C_2$  ( $k$  channels ×  $n$  trials), as follows (Eq. 3):

**Table 1. Searchlight ROIs<sup>a</sup>**

	X	Y	Z	N	Size (voxels)	p
LH iIPS	-25 (±0.76)	-74 (±0.75)	31 (±0.72)	18	635 (±31)	0.058
RH iIPS	28 (±0.64)	-73 (±0.83)	34 (±0.99)	18	736 (±37)	<1e-04
LH sIPS	-23 (±0.84)	-61 (±0.79)	47 (±0.53)	18	666 (±49)	0.004
RH sIPS	24 (±0.86)	-49 (±0.89)	51 (±0.43)	18	721 (±43)	0.013
LH iPCS	-40 (±0.73)	5 (±0.53)	42 (±0.72)	18	518 (±31)	0.008
RH iPCS	45 (±0.76)	4 (±0.96)	30 (±0.64)	17	300 (±30)	<1e-04
LH sPCS	-25 (±0.41)	-3 (±1.37)	56 (±0.53)	18	566 (±37)	0.028
RH sPCS	25 (±0.73)	-3 (±1.53)	56 (±0.60)	18	416 (±39)	0.120
LH Cing	-7 (±1.52)	-2 (±0.48)	37 (±1.04)	18	396 (±117)	0.043
RH Cing	5 (±0.10)	3 (±0.56)	37 (±0.76)	18	337 (±21)	0.043
RH IPL	44 (±0.57)	-44 (±0.92)	44 (±0.48)	18	385 (±23)	0.008
RH IT	35 (±0.32)	2 (±0.80)	-28 (±0.48)	18	140 (±10)	0.168

<sup>a</sup>X, Y, and Z are averaged (±1 SEM) Talairach coordinates across participants. N refers to the number of participants for whom a given ROI could be unambiguously identified. Size refers to the number of voxels (±1 SEM) present in each ROI. p corresponds to the proportion of bootstrap permutations where reconstruction amplitudes were <0 (FDR corrected); thus a p value >0 indicates that a given ROI contained a reliable representation of stimulus orientation. Cing, Cingulate gyrus; iIPS, inferior intraparietal sulcus; iPCS, inferior precentral sulcus; IPL, inferior parietal lobule; IT, inferior temporal cortex; LH, left hemisphere; RH, right hemisphere; sIPS, superior intraparietal sulcus; sPCS, superior precentral sulcus.

**Table 2. Frontoparietal ROIs supporting above-chance decoding of task set (attend-orientation vs attend-luminance)<sup>a</sup>**

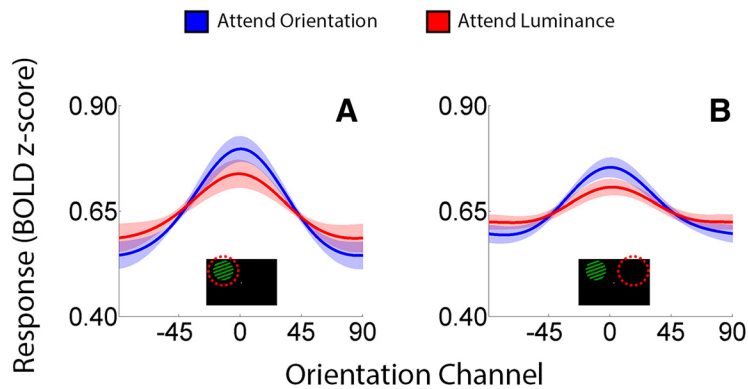
	X	Y	Z	Size (voxels)	p
LH iIPS	-23	-77	28	914	0.145
LH sIPS	-24	-57	47	833	0.010
LH IPL	-57	-26	36	842	0.025
LH SPL	-33	-37	58	550	0.033
LH iPCS	-45	-8	47	527	0.026
LH sPCS	-21	-12	61	687	0.064
LH IT	-39	-21	-24	646	0.774
LH OFC	-22	27	-13	376	0.623
LH IPFC	-47	11	30	352	0.145
RH iIPS	26	-73	34	895	<1e-04
RH sIPS	28	-36	60	981	0.003
RH IPL	48	-20	48	534	0.327
RH SPL	28	-36	60	518	0.534
RH iPCS	50	-1	37	658	0.033
RH sPCS	24	-12	59	494	0.067
RH IT	42	1	-28	314	0.222
RH OFC	23	31	-14	358	0.129
RH vIPFC	47	31	10	1092	0.033
RH dIPFC	13	30	48	461	0.338

<sup>a</sup>X, Y, and Z are Talairach coordinates. We computed a reconstructed representation of stimulus orientation using multivoxel activation patterns measured in each ROI during attend-orientation scans. p refers to the proportion of 10,000 bootstrap permutations where the amplitude of the reconstructed representation was ≤0 (FDR corrected). dIPFC, Dorsolateral prefrontal cortex; IPL, inferior parietal lobule; iIPS, inferior intraparietal sulcus; IT, inferior temporal cortex; LH, left hemisphere; IPFC, lateral prefrontal cortex; OFC, orbitofrontal cortex; RH, right hemisphere; sIPS, superior intraparietal sulcus; SPL, superior parietal lobule; vIPFC, ventrolateral prefrontal cortex.

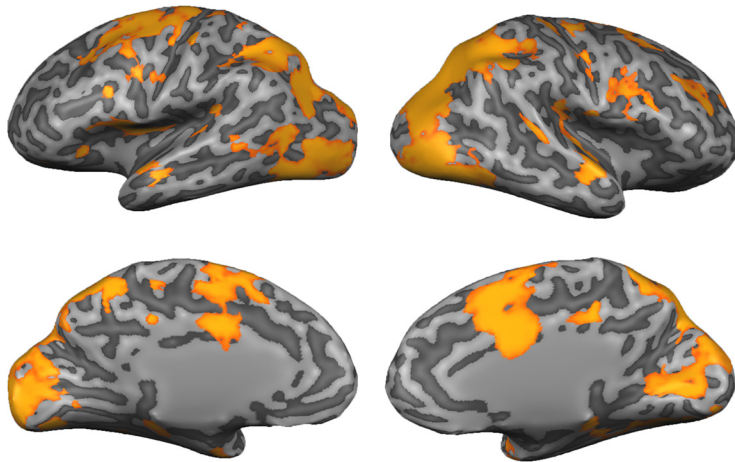
$$C_2 = (\hat{W}^T \hat{W})^{-1} \hat{W}^T B_2$$

This step transforms the data measured on each trial of the test set from voxel space back into stimulus space, such that the pattern of channel responses is a representation of the stimulus presented on each trial. The estimated channel responses on each trial were then circularly shifted to a common center (0°, by convention) and averaged across trials. To generate the smooth, 180-point functions shown, we repeated the encoding model analysis 19 times and shifted the centers of the orientation channels by 1° on each iteration.

Critically, different participants completed different numbers of attend-orientation and attend-luminance scans. We therefore implemented a cross-validation routine where  $B_1$  always contained data from an equal number of attend-orientation and attend-luminance scans. This ensured that the training set was unbiased and that the data used to estimate the weight matrix  $\hat{W}$  ( $B_1$ ) and channel response matrix  $C_2$  ( $B_2$ ) were completely independent. Data from the remaining attend-orientation and attend-luminance scans were designated as the test dataset. This procedure was repeated until all



**Figure 2.** Representations of stimulus orientation in retinotopically organized visual cortex. Data have been averaged across visual areas V1, V2v, V3v, and hV4v and sorted by location relative to the stimulus on each trial (contralateral vs ipsilateral; **A** and **B**, respectively). Shaded areas are  $\pm 1$  within-participant SEM.



**Figure 3.** Searchlight-defined ROIs encoding stimulus orientation. A leave-one-participant-out cross-validation scheme was used to generate an SPM of searchlight neighborhoods containing a robust representation of stimulus orientation for each participant ( $p < 0.01$ , FDR corrected for multiple comparisons; see Materials and Methods, Searchlight definition of ROIs representing stimulus orientation). Here, the SPMs for a representative participant (DM) have been projected onto a computationally inflated representation of his or her cortical sheet. For exposition, neighborhoods containing a robust representation of orientation have been assigned a value of 1 while neighborhoods that did not have been zeroed out. Across participants, robust representations of stimulus orientation were present in a broad network of visual, parietal, and frontal cortical areas (Table 1).

unique combinations of equal numbers of attend-orientation and attend-luminance scans were included in the training set, and reconstructions were averaged across permutations.

**Searchlight analyses.** The primary goal of this experiment was to examine whether frontoparietal regions typically implicated in top-down control encode representations of stimulus orientation and, if so, whether representations encoded by these regions are modulated by feature-based attention. To address these questions, we implemented two separate “searchlight” analyses (Kriegeskorte et al., 2006; Serences and Boynton, 2007; Ester et al., 2015) that allowed us to reconstruct and quantify representations of stimulus orientation across the entirety of the cortex. For both analyses, we defined a spherical neighborhood with a radius of 12 mm around each gray matter voxel in the cortical sheet. We next extracted and averaged the normalized response of each voxel within each neighborhood over a period spanning 6–10 s following the start of each trial. This specific window was chosen to account for a typical hemodynamic lag of 4–6 s, but all results reported here generalized across multiple temporal windows (e.g., 4–8 or 4–10 s following the start of each trial). Data within each searchlight neighborhood were sorted into 18 bins based on stimulus orientation (0–160° in 20° increments) and participants’ task set (i.e., attend orientation vs attend luminance). We made no assumptions regarding the retinotopic preferences of voxels within each neighborhood, as many regions outside the occipital cortex have large spatial

receptive fields and do not exhibit clear retinotopy at the visual eccentricities used in this study ( $\pm 7^\circ$  horizontal and  $+5^\circ$  vertical). Thus, data were combined across stimulus locations (i.e., left vs right visual field).

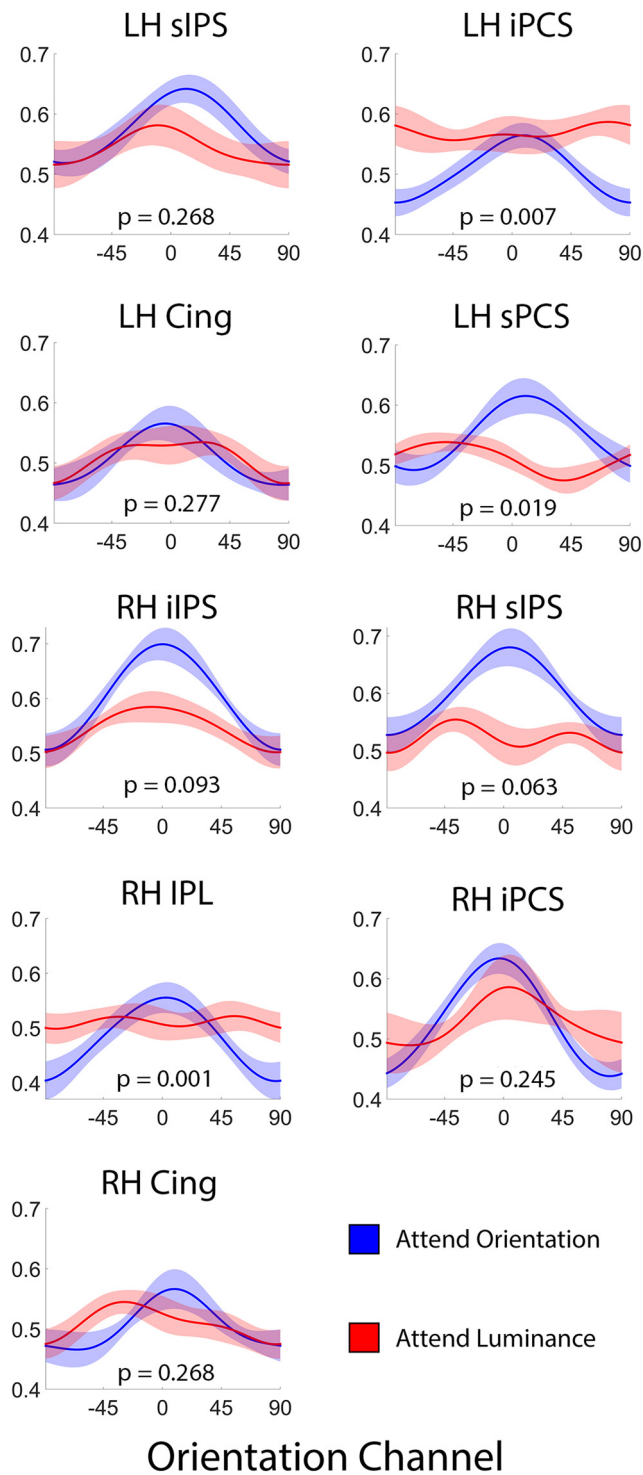
**Searchlight definition of ROIs representing stimulus orientation.** In the first searchlight analysis, we used an inverted encoding model to reconstruct representations of stimulus orientation from multivoxel activation patterns measured within each searchlight neighborhood (reconstructions were pooled and averaged across attend-orientation and attend-luminance scans). Reconstructions within each searchlight neighborhood were fit with an exponentiated cosine function of the following form (Eq. 4):  $f(x) = \alpha(e^{k[\cos(\mu-x)-1]}) + \beta$ .

Here,  $\alpha$  and  $\beta$  control the vertical scaling (i.e., signal over baseline) and baseline of the function, while  $k$  and  $\mu$  control the concentration (the inverse of dispersion; a larger value corresponds to a “tighter” function) and center of the function. No biases in reconstruction centers were expected or observed, so for convenience we fixed  $\mu$  at 0. Fitting was performed by combining a general linear model with a grid search procedure. We first defined a range of plausible  $k$  values (from 1 to 30 in 0.1 increments). For each possible value of  $k$ , we generated a response function using Equation 4 after setting  $\alpha$  to 1 and  $\beta$  to 0. Because trial-by-trial reconstructions of orientation were shifted to a common center at 0°, we fixed  $\mu$  at this value. Next, we generated a design matrix containing the predicted response function and a constant term (i.e., a vector of ones) and used ordinary least-squares regression to obtain estimates of  $\alpha$  and  $\beta$  (defined by the regression coefficient for the response function and constant term, respectively). We then selected the combination of  $k$ ,  $\alpha$ , and  $\beta$  that minimized the sum of squared errors between the observed and predicted reconstructions.

We next identified searchlight neighborhoods containing a robust representation of stimulus orientation using a leave-one-participant-out cross-validation procedure (Esterman et al., 2010). The 18 participants were identified as DA,

DB, DC, etc. through to DR. For each participant (e.g., DA), we randomly selected (with replacement) and averaged amplitude estimates from each neighborhood from each of the remaining 17 participants (e.g., DB–DR). This procedure was repeated 1000 times, yielding a set of 1000 amplitude estimates for each neighborhood. We then generated a statistical parametric map (SPM) for the held-out participant (DA) that indexed neighborhoods with amplitude estimates that were  $>0$  on 99% of all permutations [false-discovery-rate (FDR) corrected for multiple comparisons]. Finally, we projected each participant’s SPM—which was generated using data from the remaining 17 participants—onto a computationally inflated representation of his or her gray–white matter boundary and used BrainVoyager’s “Create POIs from Map Clusters” function with an area threshold of 25 mm<sup>2</sup> to identify ROIs containing a robust representation of stimulus orientation (i.e., amplitude,  $>0$ ). Clusters located in the same general anatomical area were combined to create a single ROI. Because of differences in cortical folding patterns, some ROIs could not be unambiguously identified in all 18 participants. Therefore, across participants, we retained all ROIs shared by at least 17 of 18 participants (Table 1).

In a subsequent analysis, we extracted multivoxel activation patterns from each searchlight-defined ROI and computed reconstructions of stimulus orientation during attend-orientation and attend-luminance



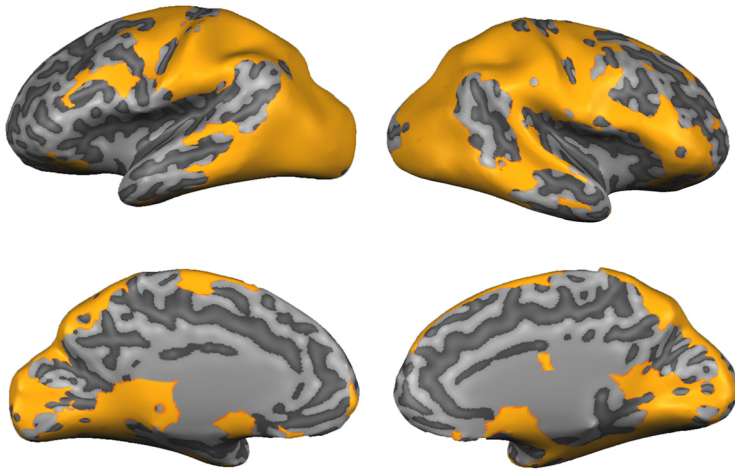
**Figure 4.** Reconstructed representations of orientation in frontoparietal cortex are modulated by task relevance. Each panel plots reconstructed representations of stimulus orientation measured during attend-orientation and attend-luminance scans in searchlight-defined ROIs that contained a robust representation of orientation (Table 1). The  $p$  value in each panel corresponds to the proportion of bootstrap permutations where amplitude estimates were reliably higher during attend-luminance relative to attend-orientation scans (FDR corrected for multiple comparisons); thus, a  $p$  value  $< 0.05$  indicates that amplitude estimates were reliably larger during attend-orientation scans relative to attend-luminance scans. Shaded regions are  $\pm 1$  within-participant SEM. Cing, Cingulate gyrus; iIPS, inferior intraparietal sulcus; iPCS, inferior precentral sulcus; IPL, inferior parietal lobule; LH, left hemisphere; RH, right hemisphere; sIPS, superior intraparietal sulcus.

scans using the inverted encoding model approach described above. Note that each participant's ROIs were defined using data from the remaining 17 participants; this ensured that participant-level reconstructions of orientation computed from data within each ROI remained statistically independent of the reconstruction used to define these ROIs in the first place (Kriegeskorte et al., 2009; Vul et al., 2009; Esterman et al., 2010). We first identified all ROIs that contained a robust representation of stimulus orientation during attend-orientation scans (we restricted our analyses to attend-orientation scans to maximize sensitivity; as shown in Figs. 4 and 6, many ROIs contained a robust representation of orientation only when this feature was relevant). Specifically, for each participant searchlight ROI we computed a reconstruction of stimulus orientation using data from attend-orientation scans. Within each ROI, we randomly selected (with replacement) and averaged reconstructions across our 18 participants. This step was repeated 10,000 times, yielding 10,000 unique stimulus reconstructions. We then estimated the amplitude of each reconstruction and computed the proportion of permutations where an amplitude estimate  $\leq 0$  was obtained (FDR corrected across ROIs).

For each ROI containing a robust representation of stimulus orientation during attend-orientation runs (defined as  $p < 0.05$ , corrected), we also computed reconstructions of stimulus orientation using data from attend-luminance scans. We then compared reconstructions across attend-orientation and attend-luminance scans using a permutation test. Specifically, for each ROI we randomly selected (with replacement) and averaged attend-orientation and attend-luminance reconstructions from our 18 participants. Each averaged reconstruction was fit with the exponentiated cosine function described by Equation 4, yielding a single amplitude, baseline, and concentration estimate for attend-orientation and attend-luminance reconstructions. This procedure was repeated 10,000 times, yielding 10,000 element vectors of parameter estimates for each task. Finally, we compared parameter estimates across tasks by computing the proportion of permutations where a larger amplitude, baseline, or higher-concentration estimate was observed during attend-luminance scans relative to attend-orientation scans ( $p < 0.05$ , FDR corrected across ROIs; Fig. 4). No reliable differences in reconstruction concentrations were observed in any of the ROIs we examined; thus we focus on amplitude and baseline estimates throughout this manuscript.

**Searchlight definition of task-selective ROIs.** Although the searchlight analysis described in the preceding section allowed us to identify ROIs encoding a robust representation of stimulus orientation throughout the cortex, it did not allow us to establish whether these ROIs were engaged in top-down control. We therefore conducted a second (independent) searchlight analysis where we trained a linear classifier to decode the participants' task set (i.e., attend-orientation vs attend-luminance) from multivoxel activation patterns measured within each searchlight neighborhood. This approach rests on the assumption that ROIs engaged in top-down control (1) should encode a representation of what feature participants are asked to attend and (2) can be used to identify neural sources of cognitive control, as has been demonstrated in several previous studies (Esterman et al., 2009; Liu et al., 2011; Riggall and Postle, 2012).

We trained a linear support vector machine (SVM; LIBSVM implementation; Chang and Lin, 2011) to discriminate between attend-orientation and attend-luminance scans based on multivoxel activation patterns measured in searchlight neighborhoods centered on every gray matter voxel in the cortical sheet. Note that the question of whether a given ROI contains task-selective information is orthogonal to the question of whether the same ROI contains orientation-selective information. Specifically, the dataset used to train the SVM always contained data from an equal number of attend-orientation and attend-luminance scans, and each attend-orientation and attend-luminance scan contained an equal number of trials for each possible stimulus orientation and location. Finally, data from each voxel in the attend-orientation and attend-luminance scans were independently  $z$ -scored (on a scan-by-scan basis) before being assigned to training or test datasets; this ensured that the overall response of each voxel



**Figure 5.** Searchlight-defined ROIs encoding task set. We combined a roving searchlight analysis with an SVM to identify cortical regions representing participants' task set (i.e., attend orientation vs attend luminance;  $p < 0.01$ , FDR corrected for multiple comparisons). Here, the resulting map has been projected onto a computationally inflated image of a representative participant's brain (DM). For exposition, searchlight neighborhoods containing a robust representation of orientation have been assigned a value of 1 while neighborhoods that did not have been zeroed out. From this map, we manually defined a set of 19 frontal, parietal, and inferior temporal ROIs that encoded task set (Table 2).

during each task had the same mean and SD. The trained SVM was used to predict participants' task set (i.e., attend-orientation vs attend-luminance) on each trial of the test set. To generate a single estimate of classifier performance for each searchlight neighborhood, we computed the proportion of test trials for which the trained SVM accurately predicted the participant's task.

This procedure was repeated until all unique combinations of equal numbers of attend-orientation and attend-luminance scans had been combined in the training set, and the results were averaged across permutations. To find ROIs containing a representation of participants' task set, we randomly selected (with replacement) and averaged decoding accuracies from each searchlight neighborhood across all 18 participants. This procedure was repeated 1000 times, yielding a set of 1000 decoding accuracy estimates for each neighborhood. We then generated an SPM marking neighborhoods where participant-averaged classifier performance exceeded chance-level decoding performance (50%) on 99% of permutations (FDR corrected for multiple comparisons). This SPM was used to identify a set of candidate ROIs that encoded task set (again via BrainVoyager's "Create POIs from Map Clusters" function with a cluster threshold  $25 \text{ mm}^2$ ; Table 2).

Although this procedure reveals ROIs that support reliable above-chance decoding performance, it does not reveal which factor(s) are responsible for robust decoding. For example, above-chance decoding could be driven by unique patterns of activity associated with participants' task set (i.e., attend-orientation or attend-luminance) or unique patterns of activity associated with some other factor (e.g., the spatial location of the stimulus on each trial). In the current study, all relevant experimental conditions (stimulus location and orientation) were fully crossed within each scan, and the dataset used to train the classifier always contained data from an equal number of attend-orientation and attend-luminance scans. It is therefore unlikely that robust decoding performance was driven by a factor other than participants' task set. Nevertheless, to provide a stronger test of this hypothesis, we extracted multivoxel activation patterns from each ROI that supported robust decoding. For each participant and ROI, we computed a null distribution of 1000 decoding accuracies after randomly shuffling the task condition labels in the training dataset. This procedure eliminates any dependence between multivoxel activation patterns and task condition labels and allowed us to estimate an upper bound on decoding performance that could be achieved by other experimental factors or fortuitous noise. For each ROI and participant, we computed the 99th percentile of this null distribution. Finally, for each ROI we compared averaged observed decoding per-

formance (obtained without shuffling) with averaged decoding performance at the 99th percentile across participants. Empirically observed decoding accuracies exceeded this criterion in all 19 ROIs we examined, confirming that these regions do indeed support robust decoding of participants' task set.

Multivoxel activation patterns from each task-selective ROI were used to reconstruct representations of stimulus orientation during attend-orientation and attend-luminance scans using the same method described above. Note that this analysis does not constitute "double dipping," as (1) the classifier was trained to discriminate between task sets regardless of stimulus orientation, (2) an equal number of attend-orientation and attend-luminance scans were used to train the classifier, and (3) data from attend-orientation and attend-luminance scans were independently z-scored before training, thereby ensuring that decoding performance could not be attributed to overall differences in response amplitudes or variability across tasks.

## Results

### *Reconstructions of stimulus orientation in retinotopically organized visual cortex*

Multiple single-unit (McAdams and Maunsell, 1999; Treue and Martinez-Trujillo, 1999) and human neuroimaging (Saenz et al., 2002; Serences and Boynton, 2007; Scolari et al., 2012) studies have documented feature-based attentional modulations in retinotopically organized subregions of the visual cortex. We therefore began by examining whether and how feature-based attention modulated representations of orientation in these areas. Reconstructed representations from retinotopically organized visual areas are plotted as a function of visual area and task (attend orientation vs luminance) in Figure 2. Reconstructions have been averaged over visual areas V1, V2v, V3v, and hV4v as similar results were obtained when each region was examined separately. Next, we estimated the amplitude, baseline, and concentration of each participant's reconstructions and compared these values across attend-orientation and attend-luminance scans using permutation tests (see Materials and Methods, Quantification and comparison of reconstructed representations). Reconstruction amplitudes were reliably higher in contralateral visual areas during attend-orientation scans relative to attend-luminance scans ( $p = 0.039$ ). Conversely, amplitude estimates did not differ across tasks in ipsilateral visual areas ( $p = 0.183$ ). No differences in reconstruction baseline or concentration estimates were observed in either contralateral or ipsilateral ROIs.

### *Reconstructions of stimulus orientation in searchlight-defined orientation-selective ROIs*

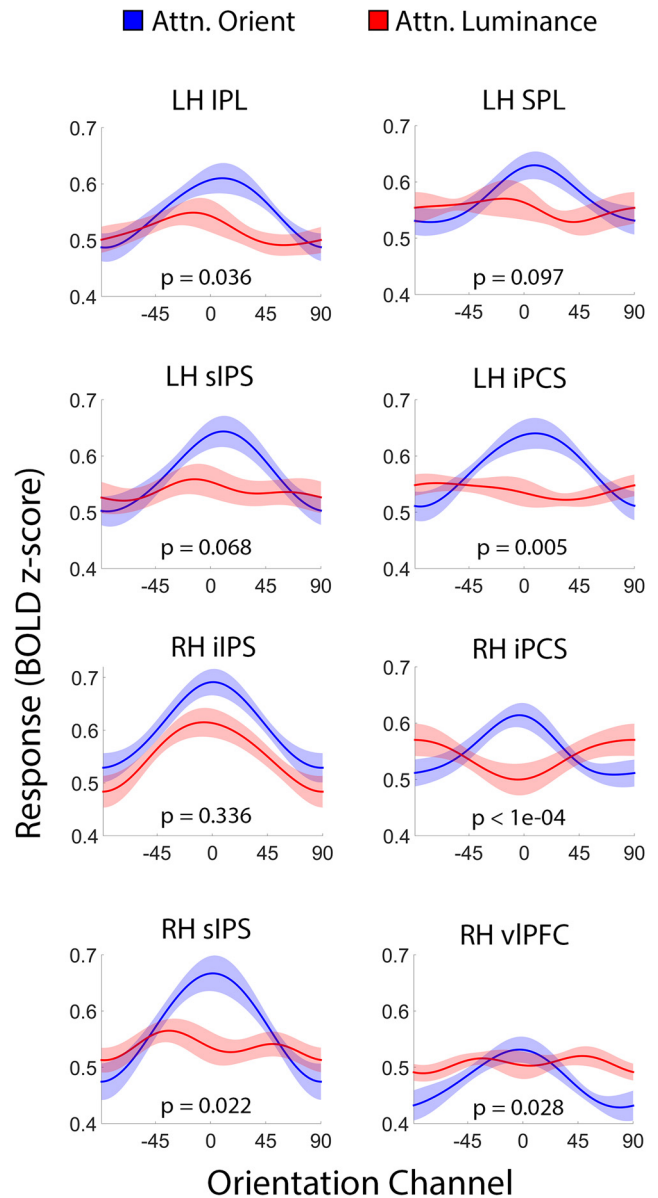
We combined a roving searchlight with an inverted encoding model to identify ROIs encoding stimulus orientation throughout the cortical sheet (see Materials and Methods, Searchlight definition of ROIs representing stimulus orientation). Across participants, robust representations of stimulus orientation were observed in a broad network of occipital, parietal, and frontal cortical areas, including the bilateral occipital cortex, the medial and superior parietal cortex, the superior precentral sulcus, and the dorsolateral prefrontal cortex (Fig. 3; Table 1). Next, we ex-

amined whether (and how) representations of stimulus orientation were modulated by task goals (i.e., attend-orientation vs attend-luminance). To do so, we extracted multivoxel activation patterns from ROIs containing a robust representation of stimulus orientation (Table 1) and generated separate reconstructions of stimulus orientation for attend-orientation and attend-luminance scans. These are shown in Figure 4. Reconstruction amplitudes were reliably larger during attend-orientation scans relative to attend-luminance scans in the left inferior and superior precentral sulcus (iPCS and sPCS, respectively; sPCS is thought to contain the human homolog of the macaque frontal eye fields; Fig. 4) and the right inferior parietal lobule (IPL), with similar trends present in right inferior and superior precentral sulcus. Additionally, reconstruction baselines were reliably higher during attend-luminance scans relative to attend-orientation scans in left iPCS and right IPL (both FDR-corrected  $p$  values  $< 1e-04$ ). Finally, attention had no effect on reconstruction concentration estimates (where a larger value corresponds to a “tighter” reconstruction) in any of the regions we examined (all FDR-corrected  $p$  values  $> 0.60$ ). These results indicate that feature-specific representations encoded by cortical areas typically regarded as “sources” of top-down control are also modulated by attention.

#### Representations of orientation in “task-selective” ROIs

Although many of the searchlight-defined frontoparietal ROIs discussed in the preceding section have been previously implicated in cognitive control (Koechlin et al., 2003; Bressler et al., 2008; Esterman et al., 2009; Bichot et al., 2015; Marshall et al., 2015), it is unclear what role(s) they serve in the current experiment. Based on earlier work (Esterman et al., 2009; Liu et al., 2011; Liu, 2016; Riggall and Postle, 2012), we reasoned that regions engaged in top-down control over visual selection would contain a representation of what task participants were instructed to perform, i.e., attend orientation vs. attend luminance. To this end, we trained a linear SVM to discriminate what task participants were instructed to perform (i.e., attend-orientation vs attend-luminance) from multivoxel activation patterns measured in searchlight neighborhoods centered on each gray matter voxel in the cortical sheet (see Materials and Methods, Searchlight definition of task-selective ROIs). As shown in Figure 5, task-selective signals were present in a broad network of bilateral visual, parietal, inferior temporal, superior precentral, and lateral prefrontal cortical regions. Here, we focus on ROIs located in the frontal and parietal cortex as the searchlight-based stimulus reconstruction approach described in the preceding section failed to identify any ROIs in the temporal cortex that were shared by a majority of participants. A complete summary of these ROIs is available in Table 2.

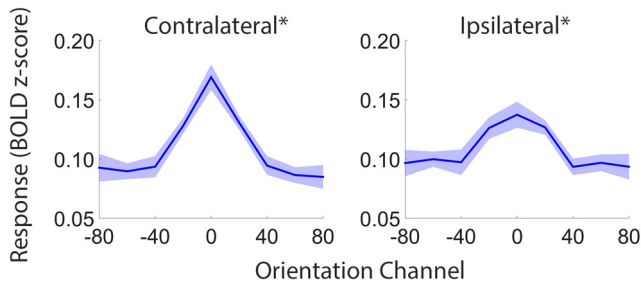
Reconstructions computed from each task-selective ROI containing a robust representation of stimulus orientation during attend-orientation scans (Table 2) are plotted as a function of task (attend-orientation vs attend-luminance) in Figure 6. Direct comparisons between reconstruction parameters within each ROI revealed higher-amplitude reconstructions during attend-orientation scans relative to attend-luminance scans in several areas, including the left inferior parietal lobule, the bilateral inferior precentral sulcus, the right superior intraparietal sulcus, and the right ventromedial prefrontal cortex (Fig. 6). Similar trends were observed in the left superior intraparietal sulcus and the left superior parietal lobule ( $p < 0.10$ ). Reconstruction baseline estimates were reliably larger during attend-luminance scans relative to attend-orientation scans in the right inferior precentral



**Figure 6.** Attentional modulations in task-selective frontoparietal ROIs. Each panel plots reconstructed representations of stimulus orientation from searchlight-defined ROIs containing a robust representation of participants’ task set (i.e., attend orientation or attend luminance; see Fig. 5 and Table 2). The  $p$  value in each panel corresponds to the proportion of bootstrap permutations where amplitude estimates were reliably higher during attend-luminance relative to attend-orientation scans (FDR corrected for multiple comparisons); thus, a  $p$  value  $< 0.05$  indicates that amplitude estimates were reliably larger during attend-orientation scans relative to attend-luminance scans. Shaded regions are  $\pm 1$  within-participant SEM. iIPS, Inferior intraparietal sulcus; iPCS, inferior precentral sulcus; IPL, inferior parietal lobule; LH, left hemisphere; RH, right hemisphere; sIPS, superior intraparietal sulcus; SPL, superior parietal lobule; vIPFC, ventrolateral prefrontal cortex.

sulcus and the right ventrolateral prefrontal cortex ( $p < 1e-04$  and  $p = 0.006$ , respectively;  $p$  values for all other regions  $> 0.17$ ). Task had no effect on reconstruction concentrations in any of the regions shown in Figure 6 (all  $p$  values  $> 0.74$ ). These results dovetail with the results of the searchlight-based reconstruction analysis described above, and thus provide converging evidence that representations of stimulus orientation in several—but not all—ROIs implicated in top-down control over visual selection were systematically modulated by task set.





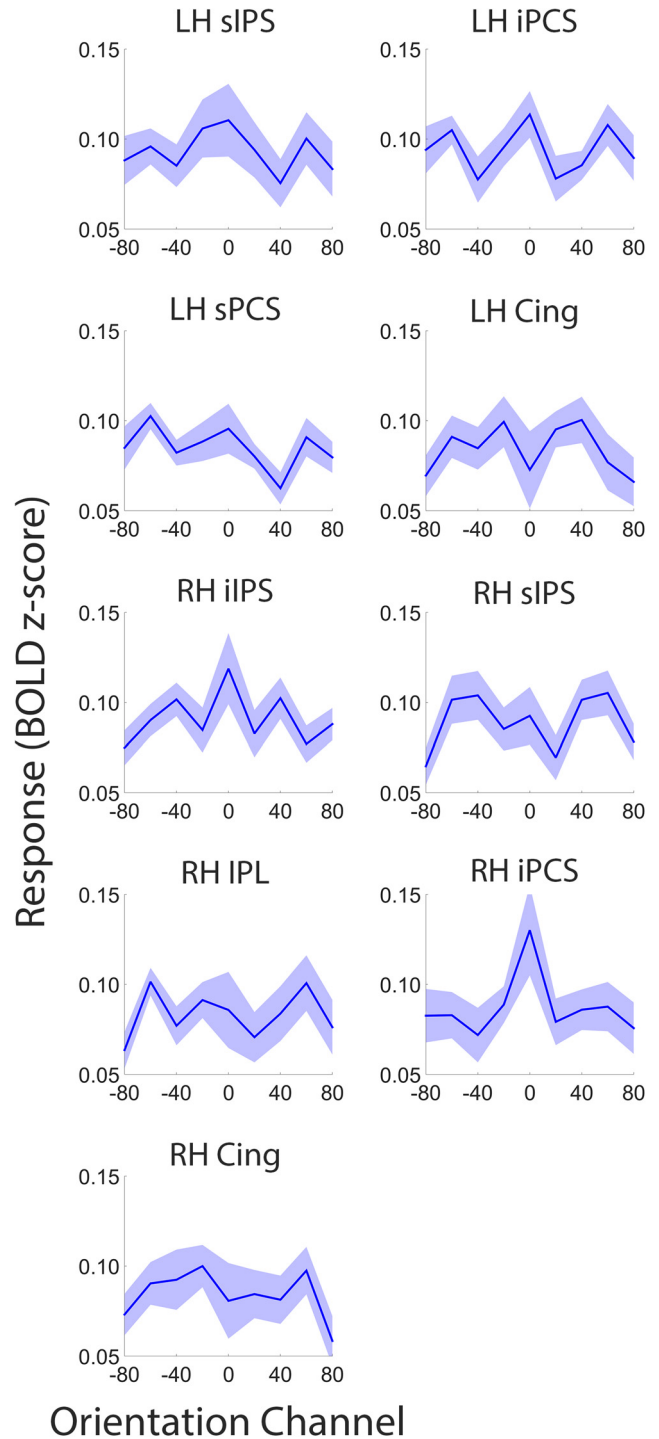
**Figure 7.** Continuous versus categorical representations in visual cortical ROIs. To examine whether the orientation-selective representations plotted in Figure 2 are continuous, we re-computed reconstructions of stimulus orientation from activation patterns measured in contralateral and ipsilateral visual areas during attend-orientation scans using a basis set of nonoverlapping delta functions. If the representation encoded by a given ROI is discrete or categorical, then the reconstructed representation computed using this approach should exhibit a sharp peak at the target orientation. We therefore computed the slope of the reconstructed representation in each ROI (see text for details). A  $p$  value  $< 0.05$  indicates a positive slope and is consistent with a continuous rather than categorical or discrete representation. Shaded regions are  $\pm 1$  within-participant SEM. \* $p < 0.05$  and  $\wedge p < 0.10$ , FDR corrected for multiple comparisons across ROIs.

*Categorical versus continuous representations of orientation*

It is well known that portions of the parietal and prefrontal cortex encode categorical information (Freedman et al., 2001). Although the smooth reconstructions shown in Figure 4 are nominally consistent with a continuous or analog representation, recall that they were generated using a basis set of nine overlapping sinusoids. This overlap ensures the responses of neighboring points along each curve are correlated, and will confer smoothness to the reconstructions even if the underlying feature representation is categorical or discrete. We therefore re-computed reconstructions of stimulus orientation measured during attend-orientation scans using a basis set containing nine orthogonal Kronecker delta functions, where each function was centered on one of the nine possible stimulus orientations (i.e., 0–160°; Saproo and Serences, 2014; Ester et al., 2015). The resulting reconstructions are plotted for visual (compare Figs. 2, 7), searchlight amplitude (compare Figs. 4, 8), and task-selective (compare Figs. 6, 9) ROIs. We reasoned that if representations of stimulus orientation are categorical, then reconstructed representations should exhibit a sharp peak at the stimulus’ orientation and a uniformly small response to all other orientations. To examine this possibility, we randomly selected (with replacement) and averaged participant-level reconstructions from ROIs containing a robust representation of stimulus orientation ( $p < 0.05$ ; Tables 1, 2). We then subtracted the average responses of orientation channels located three and four steps away from the target orientation (i.e.,  $\pm 60^\circ$  and  $\pm 80^\circ$ ) from the averaged responses of orientation channels adjacent to the target (i.e.,  $\pm 20^\circ$  and  $\pm 40^\circ$ ), yielding an estimate of reconstruction slope. This procedure was repeated 10,000 times, yielding a distribution of reconstruction slopes for each ROI. Slope estimates for many visual and task-selective ROIs were reliably  $> 0$ , consistent with a continuous rather than discrete or categorical representation (Figs. 7, 9). We were unable to reconstruct robust representations of stimulus orientation in many searchlight amplitude ROIs (Fig. 8). Slope estimates in each of these regions were indistinguishable from 0 (all  $p$  values  $> 0.40$ ).

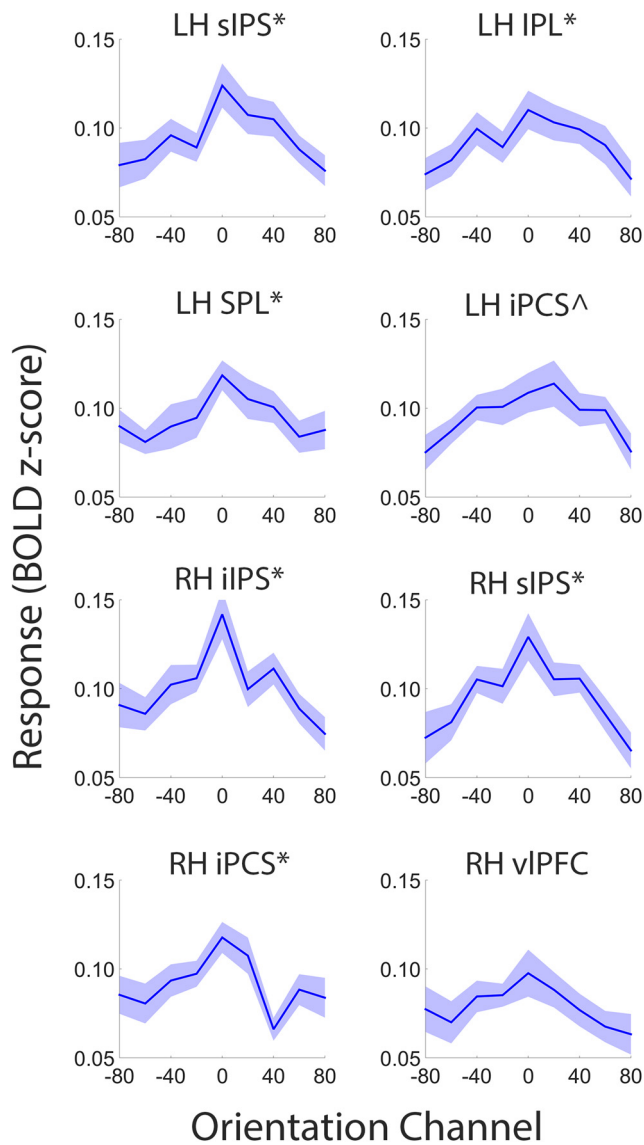
*Eye-tracking control analysis*

To assess compliance with fixation instructions, we recorded continuous eye-position data for seven participants. We identified all stable fixations (defined as a 200 ms epoch during which



**Figure 8.** Continuous versus categorical representations in searchlight amplitude ROIs. Compare with attend-orientation reconstructions in Figure 4. Conventions are in Figure 7. We were unable to reconstruct a representation of stimulus orientation in many ROIs. Shaded regions are  $\pm 1$  within-participant SEM \* $p < 0.05$  and  $\wedge p < 0.10$ , FDR corrected for multiple comparisons across ROIs.

eye position did not deviate  $> 0.25^\circ$ ) that occurred outside of a  $0.5^\circ$  centered on fixation during the course of each 10 s trial. We then compared the endpoints, polar angles, and polar distances of these fixations as a function of stimulus location (i.e., left or right visual field) and stimulus orientation (repeated-measures ANOVA with stimulus location and stimulus orientation as within-participant factors). We observed no main effects or in-



**Figure 9.** Continuous versus categorical representations in task-selective ROIs. Compare with attend-orientation reconstructions in Figure 6. Conventions are in Figures 7 and 8. Shaded regions are  $\pm 1$  within-participant SEM. \* $p < 0.05$  and  $\wedge p < 0.10$ , FDR corrected for multiple comparisons across ROIs.

teractions between these factors on either saccade endpoints or saccade vectors (all FDR-corrected  $p$  values  $> 0.27$ ). We also directly compared stimulus reconstructions across participants from whom eye-position data were ( $N = 7$ ) or were not ( $N = 11$ ) collected in each ROI that contained a robust representation of stimulus orientation (Figs. 2, 4, 6). Specifically, for each ROI we randomly selected (with replacement) and averaged seven attend-orientation reconstructions from the seven participants who underwent eye tracking while in the scanner and 7 of the 11 participants who did not. We then estimated and compared reconstruction amplitudes across these groups. This procedure was repeated 10,000 times, yielding a 10,000 element vector of group amplitude differences for each ROI. Finally, we estimated an empirical (FDR corrected)  $p$  value for amplitude differences within each ROI by computing proportion of permutations where the amplitude difference was  $\geq 0$ . A  $p$  value  $< 0.025$  indicates that reconstruction amplitudes were reliably smaller in the group of participants who underwent eye tracking, while a  $p$  value  $> 0.975$

**Table 3.** Eye-movement control analyses<sup>a</sup>

	Nontracked $\alpha$ –tracked $\alpha$
Visual cortex (Fig. 2)	
Contralateral	0.946
Ipsilateral	0.841
Searchlight ROIs representing orientation (Table 1; Fig. 4)	
RH iIPS	0.030
LH sIPS	0.114
RH sIPS	0.310
LH iPCS	0.106
RH iPCS	0.225
LH sPCS	0.036
LH Cing	0.059
RH Cing	0.225
RH IPL	0.567
Searchlight ROIs representing task set (Table 2; Fig. 6)	
LH sIPS	0.125
LH IPL	0.195
LH SPL	0.165
LH iPCS	0.165
RH iIPS	0.125
RH sIPS	0.165
RH iPCS	0.125
RH vIPFC	0.125

<sup>a</sup>For each ROI shown in Figures 2, 4, and 6, we computed, modeled, estimated, and compared the amplitudes of reconstructed representations measured during attend-orientation scans across participants who did versus those who did not undergo eye-tracking while in the scanner. Nontracked  $\alpha$ –tracked  $\alpha$  corresponds to the proportion of 10,000 bootstrap permutations where reconstruction amplitudes in each ROI were larger for participants who were not tracked relative to those who were (FDR corrected). No group-level differences in reconstruction amplitudes were observed in any of these ROIs, suggesting that the attentional modulations shown in Figures 2, 4, and 6 cannot be attributed to different eye-movement strategies. Cing, Cingulate gyrus; IPL, inferior parietal lobule; iIPS, inferior intraparietal sulcus; iPCS, inferior precentral sulcus; LH, left hemisphere; RH, right hemisphere; sIPS, superior intraparietal sulcus; sPCS, superior precentral sulcus; SPL, superior parietal lobule; vIPFC, ventrolateral prefrontal cortex.

indicates the converse (two-tailed). With few exceptions (the right IPS and left sPCS ROIs defined using the searchlight-based reconstruction procedure; Table 3) were well within these boundaries. Thus, the reconstructions shown in Figures 2, 4, and 6 cannot be explained by eye movements.

## Discussion

Here, we used an inverted encoding model and a roving searchlight analysis to reconstruct and quantify representations of orientation from population-level activity across the entire human cortical sheet. We observed robust, continuous representations of orientation in a multitude of visual, parietal, and frontal cortical areas. Moreover, orientation-selective representations in many of these areas were enhanced during attend-orientation relative to attend-luminance scans. Collectively, our results suggest several frontoparietal cortical regions—long thought to provide the source of attentional control signals—encode continuous representations of sensory information, and that the representations encoded by many (but not all) of these areas are modulated by attention.

In a recent study, we reported that multiple frontoparietal cortical areas encode precise, analog representations of orientation during a visual working memory task (Ester et al., 2015). The current study builds upon these findings in several important ways. In our earlier study, we post-cued participants to remember one of two lateralized gratings over a brief delay interval and found representations of the cued, but not the uncued, grating in multiple regions of posterior sensory and frontoparietal cortical areas. In the current study, participants were instructed to attend either the orientation or luminance of a lateralized grating. According to “object-based” models of attention, selecting one feature of an object enhances cortical representations of that feature,

along with all other features of the same object (Duncan, 1984; Egly et al., 1994; Roelfsema et al., 1998; O'Craven et al., 1999). However, we found stronger representations of orientation during attend-orientation scans relative to attend-luminance scans in multiple posterior sensory and frontoparietal cortical areas. This finding dovetails with other reports suggesting that feature-based attention can selectively enhance representations of task-relevant features without enhancing representations of task-relevant features that are part of the same object (Serences et al., 2009; Xu, 2010; Jehee et al., 2011), and demonstrates that feature-based attentional modulations are distributed across the visual processing hierarchy, including regions typically associated with attentional control rather than sensory processing.

Traditionally, “sources” and “targets” of attentional control signals have been distinguished on the basis of univariate response properties (e.g., averaged single-unit/population spike rates or averaged fMRI activation). For example, individual neurons in many frontoparietal cortical areas regarded as sources of attentional control signals often exhibit selectivity for multiple task-level variables (e.g., which of multiple stimuli should be attended or what motor outputs are appropriate given the current context) but not sensory variables or feature properties. Conversely, neurons in posterior sensory cortical areas regarded as targets of attentional control signals exhibit strong selectivity for specific feature properties, but not other task-level variables. In the current study, we show that parametric sensory information is encoded within multivariate activation patterns in posterior sensory and frontoparietal cortical areas (Mante et al., 2013; Rigotti et al., 2013; Raposo et al., 2014; Ester et al., 2015), suggesting that sources and targets of attentional control signals cannot be fully dissociated on the basis of their representational properties.

Sources and targets of attentional control signals can also be distinguished by examining functional interactions between cortical areas. For example, several studies have reported that feature-based attentional modulations observed in posterior sensory cortical areas lag similar modulations observed in frontoparietal cortical areas by several dozen milliseconds (Buschman and Miller, 2007; Zhou and Desimone, 2011; Siegel et al., 2015). Others have reported direct links between activity in frontoparietal cortical areas and feature-based attentional modulations in posterior sensory areas. In one example, Baldauf and Desimone (2014) reported increased gamma band synchrony between posterior sensory areas and the IFJ during an object-based attention task. Critically, gamma phases were advanced in the IFJ relative to posterior sensory areas, suggesting that this region was the driver of changes in synchrony. In a second example, Bichot et al. (2015) showed that neurons in the ventral prearcuate (VPA) region of the prefrontal cortex exhibited feature selectivity during a visual search task. Feature-selective signals in this region emerged before feature-selective signals in the frontal eye fields (FEFs) or inferotemporal cortex, and transient deactivation of VPA abolished feature selectivity in the FEF. Thus, while our findings argue against a clear divide between source and target based on representational properties, the timing and order of neural activity in different regions may still point to a broad distinction between these two aspects of attentional function.

Single-unit recording studies suggest that feature-based attentional modulations in visual areas V4 and MT are well described by a feature-similarity gain model, where attention increases the gain of neurons preferring the attended orientation and decreases the gain of neurons preferring orthogonal orientation (Treue and Martinez-Trujillo, 1999; Martinez-Trujillo and Treue, 2004). These gain changes, in turn, lead to a decrease in the bandwidth of

population-level feature representations. In the current study, we found that reconstructions of stimulus orientation had a larger amplitude during attend-orientation scans relative to attend-luminance scans, but no differences in bandwidth. This likely reflects important differences between the displays used here and those used in other studies. For example, Martinez-Trujillo and Treue (2004) recorded from cells retinotopically mapped to the location of a task-irrelevant stimulus located in the visual hemifield opposite the target. The critical finding was that the responses of these neurons were contingent on the similarity between the features of this task-irrelevant stimulus and the target in the opposite hemifield. Thus, when the features of the target and task-relevant stimulus matched, responses to the task-irrelevant stimulus increased. Conversely, when the features of the target and task-irrelevant stimulus did not match, responses to the latter were suppressed. In the current study, we presented a single grating in the upper left or right visual field, with no stimuli in the opposite hemifield. Thus, there was never a task-irrelevant sensory signal that needed to be enhanced or suppressed. We speculate that feature-similarity gain modulations (i.e., an increase in the responses of orientation channels preferring the stimulus' orientation coupled with a decrease in the responses of orientation channels preferring the orthogonal orientation) would manifest if a task-irrelevant distractor was present in the visual hemifield opposite the target.

Many influential models of visual processing postulate functionally and anatomically segregated “top-down control” and sensory processing areas, with perception ultimately dependent on the coordination of signals originating in these areas. Here, we show that several frontoparietal cortical regions typically associated with top-down attentional control encode parametric representations of sensory stimuli similar to those observed in posterior sensory cortical areas. Moreover, we show that these representations are modulated by task demands in many (though not all) frontoparietal cortical areas. These findings are inconsistent with classic models of selective attention and cognitive control that postulate segregated attentional control and sensory processing networks. However, they are readily accommodated by generative (e.g., predictive coding; Friston, 2008) or dynamic inference models where sensory signals are passed between hierarchically organized cortical systems to compute a probabilistic representation of the external environment. In these models, attention optimizes the process of perceptual inference by reducing uncertainty about the likely state of the world. At the level of single neurons or local cortical circuits, this process can be achieved by selectively increasing the gain of neurons carrying the most information about a stimulus (Feldman and Friston, 2010).

## References

- Andersen SK, Hillyard SA, Müller MM (2008) Attention facilitates multiple stimulus features in parallel in human visual cortex. *Curr Biol* 18:1006–1009. [CrossRef Medline](#)
- Baldauf D, Desimone R (2014) Neural mechanisms of object-based attention. *Science* 344:424–427. [CrossRef Medline](#)
- Bichot NP, Heard MT, DeGennaro EM, Desimone R (2015) A source for feature-based attention in the prefrontal cortex. *Neuron* 88:832–844. [CrossRef Medline](#)
- Brainard D (1997) The psychophysics toolbox. *Spat Vis* 10:433–436. [CrossRef Medline](#)
- Bressler SL, Tang W, Sylvester CM, Shulman GL, Corbetta M (2008) Top-down control of human visual cortex by frontal and parietal cortex in anticipatory visual spatial attention. *J Neurosci* 28:10056–10061. [CrossRef Medline](#)
- Brouwer GJ, Heeger DJ (2009) Decoding and reconstructing color from

- responses in human visual cortex. *J Neurosci* 29:13992–14003. [CrossRef Medline](#)
- Brouwer GJ, Heeger DJ (2011) Cross-orientation suppression in human visual cortex. *J Neurophysiol* 106:2108–2119. [CrossRef Medline](#)
- Buschman TJ, Miller EK (2007) Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* 315:1860–1862. [CrossRef Medline](#)
- Buschman TJ, Siegel M, Roy JE, Miller EK (2011) Neural substrates of cognitive capacity limitations. *Proc Natl Acad Sci U S A* 108:11252–11255. [CrossRef Medline](#)
- Chang CC, Lin CJ (2011) LIBSVM: a library for support vector machines. *ACM Trans Intell Syst Technol* 2:1–27.
- Chawla D, Rees G, Friston KJ (1999) The physiological basis of attentional modulation in extrastriate visual areas. *Nat Neurosci* 2:671–676. [CrossRef Medline](#)
- Corbetta M, Miezin FM, Dobmeyer S, Shulman GL, Petersen SE (1990) Attentional modulation of neural processing of shape, color, and velocity in humans. *Science* 248:1556–1559. [CrossRef Medline](#)
- Duncan J (1984) Selective attention and the organization of visual information. *J Exp Psychol Gen* 113:501–517. [CrossRef Medline](#)
- Egly R, Driver J, Rafal RD (1994) Shifting visual attention between objects and locations: evidence from normal and parietal lesion subjects. *J Exp Psychol Gen* 123:161–177. [CrossRef Medline](#)
- Ester EF, Sprague TC, Serences JT (2015) Parietal and frontal cortex encode stimulus-specific mnemonic representations during visual working memory. *Neuron* 87:893–905. [CrossRef Medline](#)
- Esterman M, Chiu YC, Tamber-Rosenau BJ, Yantis S (2009) Decoding cognitive control in human parietal cortex. *Proc Natl Acad Sci U S A* 106:17974–17979. [CrossRef Medline](#)
- Esterman M, Tamber-Rosenau BJ, Chiu YC, Yantis S (2010) Avoiding non-independence in fMRI data analysis: leave one subject out. *Neuroimage* 50:572–576. [CrossRef Medline](#)
- Feldman H, Friston KJ (2010) Attention, uncertainty, and free energy. *Front Hum Neurosci* 4:215. [CrossRef Medline](#)
- Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2011) Categorical representation of visual stimuli in the primate prefrontal cortex. *Science* 291:312–316. [CrossRef Medline](#)
- Freeman J, Brouwer GJ, Heeger DJ, Merriam EP (2011) Orientation decoding depends on maps, not columns. *J Neurosci* 31:4792–4804. [CrossRef Medline](#)
- Freeman WT, Adelson EH (1991) The design and use of steerable filters. *IEEE T Pattern Anal* 13:891–906. [CrossRef](#)
- Friston K (2008) Hierarchical models in the brain. *PLoS Comput Biol* 4:e10000211. [CrossRef Medline](#)
- Greenberg AS, Esterman M, Wilson D, Serences JT, Yantis S (2010) Control of spatial and feature-based attention in frontoparietal cortex. *J Neurosci* 30:14330–14339. [CrossRef Medline](#)
- Gregoriou GG, Rossi AF, Ungerleider LG, Desimone R (2014) Lesions of prefrontal cortex reduce attentional modulation of neuronal responses and synchrony in V4. *Nat Neurosci* 17:1003–1011. [CrossRef Medline](#)
- Gur M, Kagan I, Snodderly DM (2005) Orientation and direction selectivity of neurons in V1 of alert monkeys: functional relationships and laminar dependence. *Cereb Cortex* 15:1207–1221. [Medline](#)
- Jehee JF, Brady DK, Tong F (2011) Attention improves encoding of task-relevant features in human visual cortex. *J Neurosci* 31:8210–8219. [CrossRef Medline](#)
- Kamitani Y, Tong F (2005) Decoding the visual and subjective contents of the human brain. *Nat Neurosci* 8:679–685. [CrossRef Medline](#)
- Kelley TA, Serences JT, Giesbrecht B, Yantis S (2008) Cortical mechanisms for shifting and holding visuospatial attention. *Cereb Cortex* 18:114–125. [CrossRef Medline](#)
- Kim JN, Shadlen MN (1999) Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nat Neurosci* 2:176–185. [CrossRef Medline](#)
- Koechlin E, Ody C, Kouneiher F (2003) The architecture of cognitive control in the human prefrontal cortex. *Science* 302:1181–1185. [CrossRef Medline](#)
- Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping. *Proc Natl Acad Sci U S A* 103:3863–3868. [CrossRef Medline](#)
- Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI (2009) Circular analysis in systems neuroscience: The dangers of double dipping. *Nat Neurosci* 12:535–540. [CrossRef Medline](#)
- Liu T (2016) Neural representation of object-specific attentional priority. *Neuroimage* 129:15–24. [CrossRef Medline](#)
- Liu T, Slotnick SD, Serences JT, Yantis S (2003) Cortical mechanisms of feature-based attentional control. *Cereb Cortex* 13:1334–1343. [CrossRef Medline](#)
- Liu T, Larsson J, Carrasco M (2007) Feature-based attention modulates orientation-selective responses in human visual cortex. *Neuron* 55:313–323. [CrossRef Medline](#)
- Liu T, Hospadaruk L, Zhu DC, Gardner JL (2011) Feature-specific attentional priority signals in human cortex. *J Neurosci* 31:4484–4495. [CrossRef Medline](#)
- Mante V, Sussillo D, Shenoy KV, Newsome WT (2013) Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* 503:78–84. [CrossRef Medline](#)
- Marshall TR, O’Shea J, Jensen O, Bergmann TO (2015) Frontal eye fields control attentional modulation of alpha and gamma oscillations in contralateral occipital cortex. *J Neurosci* 35:1638–1647. [CrossRef Medline](#)
- Martinez-Trujillo JC, Treue S (2004) Feature-based attention increases the selectivity of population responses in primate visual cortex. *Curr Biol* 14:744–751. [CrossRef Medline](#)
- McAdams CJ, Maunsell JH (1999) Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J Neurosci* 19:431–441. [Medline](#)
- Mendoza-Halliday D, Torres S, Martinez-Trujillo JC (2014) Sharp emergence of feature-selective sustained activity along the dorsal visual pathway. *Nat Neurosci* 17:1255–1262. [CrossRef Medline](#)
- Meyer T, Qi XL, Stanford TR, Constantinidis C (2011) Stimulus selectivity in dorsal and ventral prefrontal cortex after training in working memory tasks. *J Neurosci* 31:6266–6276. [CrossRef Medline](#)
- Müller MM, Andersen S, Trujillo NJ, Valdés-Sosa P, Malinowski P, Hillyard SA (2006) Feature-selective attention enhances color signals in early visual areas of the human brain. *Proc Natl Acad Sci U S A* 103:14250–14254. [CrossRef Medline](#)
- O’Craven KM, Downing PE, Kanwisher N (1999) fMRI evidence for objects as the units of attentional selection. *Nature* 401:584–587. [CrossRef Medline](#)
- Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* 10:437–442. [CrossRef Medline](#)
- Polk TA, Drake RM, Jonides JJ, Smith MR, Smith EE (2008) Attention enhances the neural processing of relevant features and suppresses the processing of irrelevant features in humans: a functional magnetic resonance imaging study of the Stroop task. *J Neurosci* 28:13786–13792. [CrossRef Medline](#)
- Raposo D, Kaufman MT, Churchland AK (2014) A category-free neural population supports evolving demands during decision making. *Nat Neurosci* 17:1784–1792. [CrossRef Medline](#)
- Riggall AC, Postle BR (2012) The relationship between working memory storage and elevated activity as measured with functional magnetic resonance imaging. *J Neurosci* 32:12990–12998. [CrossRef Medline](#)
- Rigotti M, Barak O, Warden MR, Wang XJ, Daw ND, Miller EK, Fusi S (2013) The importance of mixed selectivity in complex cognitive tasks. *Nature* 497:585–590. [CrossRef Medline](#)
- Ringach DL, Shapley RM, Hawken MJ (2002) Orientation selectivity in Macaque V1: diversity and laminar dependence. *J Neurosci* 22:5639–5651. [Medline](#)
- Roelfsema PR, Lamme VA, Spekreijse H (1998) Object-based attention in the primary visual cortex of the macaque monkey. *Nature* 395:376–381. [CrossRef Medline](#)
- Saenz M, Buracas GT, Boynton GM (2002) Global effects of feature-based attention in human visual cortex. *Nat Neurosci* 5:631–632. [CrossRef Medline](#)
- Saproo S, Serences JT (2014) Attention improves transfer of motion information between V1 and MT. *J Neurosci* 34:3586–3596. [CrossRef Medline](#)
- Scolari M, Byers A, Serences JT (2012) Optimal deployment of attentional gain during fine discriminations. *J Neurosci* 32:7723–7733. [CrossRef Medline](#)
- Serences JT, Boynton GM (2007) Feature-based attentional modulations in the absence of direct visual stimulation. *Neuron* 55:301–312. [CrossRef Medline](#)

- Serences JT, Schwarzbach J, Courtney SM, Golay X, Yantis S (2004) Control of object-based attention in human cortex. *Cereb Cortex* 14:1346–1357. [CrossRef Medline](#)
- Serences JT, Ester EF, Vogel EK, Awh E (2009) Stimulus-specific delay activity in human primary visual cortex. *Psychol Sci* 20:207–214. [CrossRef Medline](#)
- Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RBH (1995) Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268:889–892. [CrossRef Medline](#)
- Siegel M, Buschman TJ, Miller EK (2015) Cortical information flow during flexible sensorimotor decisions. *Science* 348:1352–1355. [CrossRef Medline](#)
- Treisman AM, Gelade G (1980) A feature-integration theory of attention. *Cogn Psychol* 12:97–136. [CrossRef Medline](#)
- Treue S, Martínez-Trujillo JC (1999) Feature-based attention influences motion processing gain in macaque visual cortex. *Nature* 399:575–579. [CrossRef Medline](#)
- Vul E, Harris C, Winkielman P, Pashler H (2009) Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspect Psychol Sci* 4:274–290. [CrossRef Medline](#)
- Wallis JD, Anderson KC, Miller EK (2001) Single neurons in prefrontal cortex encode abstract rules. *Nature* 411:953–956. [CrossRef Medline](#)
- Wolfe J (1994) Guided search 2.0: a revised model of visual search. *Psychon B Rev* 1:202–238. [CrossRef](#)
- Xu Y (2010) The neural fate of task-irrelevant features in object based processing. *J Neurosci* 30:14020–14028. [CrossRef Medline](#)
- Zhang W, Luck SJ (2009) Feature-based attention modulates feedforward visual processing. *Nat Neurosci* 12:24–25. [CrossRef Medline](#)
- Zhou H, Desimone R (2011) Feature-based attention in the frontal eye field and area V4 during visual search. *Neuron* 70:1205–1217. [CrossRef Medline](#)