

Feature-Independent Neural Coding of Target Detection during Search of Natural Scenes

Fei Guo, Tim J. Preston, Koel Das, Barry Giesbrecht, and Miguel P. Eckstein

Department of Psychological and Brain Sciences, and Institute for Collaborative Biotechnologies, University of California, Santa Barbara, Santa Barbara, California 93106-9660

Visual search requires humans to detect a great variety of target objects in scenes cluttered by other objects or the natural environment. It is unknown whether there is a general purpose neural detection mechanism in the brain that codes the presence of a wide variety of categories of objects embedded in natural scenes. We provide evidence for a feature-independent coding mechanism for detecting behaviorally relevant targets in natural scenes in the dorsal frontoparietal network. Pattern classifiers using single-trial fMRI responses in the dorsal frontoparietal network reliably predicted the presence of 368 different target objects and also the observer's choices. Other vision-related areas such as the primary visual cortex, lateral occipital complex, the parahippocampal, and the fusiform gyri did not predict target presence, while high-level association areas related to general purpose decision making, including the dorsolateral prefrontal cortex and anterior cingulate, did. Activity in the intraparietal sulcus, a main area in the dorsal frontoparietal network, correlated with observers' decision confidence and with the task difficulty of individual images. These results cannot be explained by physical differences across images or eye movements. Thus, the dorsal frontoparietal network detects behaviorally relevant targets in natural scenes independent of their defining visual features and may be the human analog of the priority map in monkey lateral intraparietal cortex.

Introduction

During visual search, objects relevant to behavioral goals must be located in the midst of a cluttered visual environment. Does the brain have specialized mechanisms dedicated to efficiently detect search targets in natural scenes? Single-unit studies have identified that the lateral intraparietal area (LIP) (Bisley and Goldberg, 2003, 2006, 2010; Ipata et al., 2006, 2009; Thomas and Paré, 2007) and the frontal eye fields (FEF) (Schall, 2002; Juan et al., 2004; Thompson et al., 2005; Buschman and Miller, 2007) are regions in the frontoparietal network involved in visual search in synthetic displays. Both regions show elevated activity to feature properties defining the search target. LIP also contributes to attending to visual features (e.g., stimulus shape) independent of eye movements (Sereno and Maunsell, 1998). Neuroimaging studies have identified activations within the human intraparietal sulcus (IPS) and FEF during covert shifts of attention in expectation of a target (Kastner et al., 1999; Corbetta et al., 2000; Hopfinger et al., 2000; Giesbrecht et al., 2003; Giesbrecht and

Mangun, 2005; Kincade et al., 2005). An important property of the frontoparietal network is feature-independent coding (Shulman et al., 2002; Giesbrecht et al., 2003; Giesbrecht and Mangun, 2005; Slagter et al., 2007; Greenberg et al., 2010; Ptak, 2011). Unlike earlier visual areas, which respond to fixed visual attributes (orientation, motion direction, etc.), LIP and FEF respond to behaviorally defined target properties. This property might allow for feature-independent coding of search targets in natural scenes within the frontoparietal network.

However, search in natural scenes, unlike the synthetic displays often used to study search, involves more complex operations such as object and scene categorization and thus might employ mechanisms related to object and scene recognition. A separate literature has investigated the neural correlates of scene perception (Thorpe et al., 1996; Fabre-Thorpe et al., 1998; VanRullen and Thorpe, 2001; Codispoti et al., 2006; Peelen et al., 2009; Walther et al., 2009; Ossandón et al., 2010; Phillips and Segraves, 2010). The parahippocampal place area (PPA), retrosplenial cortex (RSC), and lateral occipital complex (LOC) have been shown to contribute to scene perception (Peelen et al., 2009; Walther et al., 2009; Park et al., 2011): the PPA response primarily reflects spatial boundaries (Kravitz et al., 2011; Park et al., 2011), whereas the LOC response contains information about the scene content (Peelen et al., 2009). However, studies investigating scene categorization typically use a restricted set of object categories [e.g., animal vs nonanimal, cars vs people (Thorpe et al., 1996; VanRullen and Thorpe, 2001; Peelen et al., 2009)]; thus, it is unknown whether these areas exhibit feature-independent coding of search targets in scenes regardless of their categories and visual features.

Received Nov. 23, 2011; revised May 9, 2012; accepted May 17, 2012.

Author contributions: F.G., T.J.P., B.G., and M.P.E. designed research; F.G. and T.J.P. performed research; K.D. contributed unpublished reagents/analytic tools; F.G. and T.J.P. analyzed data; F.G., B.G., and M.P.E. wrote the paper.

Funding for this project was graciously provided by Army Grant W911NF-09-D-0001. Koel Das was supported by Army Grant W911NF-09-D-0001 and the Ramalingaswami Fellowship of the Department of Biotechnology, Government of India.

The authors declare no conflicting financial interests.

Correspondence should be addressed to Miguel P. Eckstein, Department of Psychological and Brain Sciences, University of California, Santa Barbara, Santa Barbara, CA 93106-9660. E-mail: miguel.eckstein@psych.ucsb.edu.

K. Das's present address: Indian Institute of Science Education and Research, Kolkata, India.

DOI:10.1523/JNEUROSCI.5876-11.2012

Copyright © 2012 the authors 0270-6474/12/329499-12\$15.00/0

We investigated the neural correlates mediating visual search in natural scenes for a large set of target objects from a range of categories using multivariate pattern classifiers to analyze single-trial fMRI activity. Our main goals were to identify brain areas mediating the detection of arbitrary targets embedded in natural scenes and to determine whether the identified target-related neural activity correlates with observers' trial-to-trial decisions.

Materials and Methods

Subjects

Twelve observers (4 male and 8 female; mean age 24; range 19–31) from the University of California, Santa Barbara participated in the main experiment. Six additional observers (4 male and 2 female; mean age 22; range 21–25) participated in a control experiment. All observers were naive to the purpose of the experiments. All observers had normal or corrected to normal vision (based on self-report). Before participation, participants provided written informed consent that had been approved by the Ethical Committee of the University of California, Santa Barbara. The observers were reimbursed for their participation with course credit or money.

Experimental design

Visual stimuli. Stimuli consisted of a set of 640 images of natural scenes (both indoor and outdoor). Each image was 727×727 pixels subtending a viewing angle of 17.5° . Images were displayed on a rear projection screen behind the head-coil inside the magnet bore using an LCD video projector (Hitachi CPX505, 1024×768 resolution) within a faraday cage in the scanner room (peak luminance: 455 cd/m^2 , mean luminance 224 cd/m^2 , minimum luminance: 5.1 cd/m^2). Observers viewed the screen via a mirror angled at 45° above their heads. The viewing distance was 108 cm providing a 24.6° field of view.

Visual search experiment. Each image was paired with a cue word that corresponded to an object present in the image or an object absent from the image. All targets (for both for target-present and target-absent conditions) were semantically consistent with the scene. The spatial location and appearance of the target objects in the scenes were highly variable. The observers had no prior information of the size or position of the target object in a single image, and the presence/absence of a target object was not predictable. Among the target-present images, $\sim 50\%$ of the images contained the target on the left side and 50% on the right side. The stimulus set consisted of two main experimental conditions: target-present (in the left or right visual field); target-absent. In addition, another condition consisted of a period of central fixation for the duration of the trial (these fixation trials were included to add jitter to the onsets of the experimental trials). A single run of the fMRI experiment lasted 6 min and 46.62 s (251 TRs) consisting of 4 TRs of initial and final fixation, 32 trials of target-present images, 32 trials of target-absent images, and 16 trials of fixation (i.e., null-events used to facilitate deconvolution of the hemodynamic response). The order of the trials was pseudorandom constrained so that each trial type had a matched trial history one trial back (over all trials in the run) (Buracas and Boynton, 2002). Thus, an additional trial was included for one of the conditions at the start of each run to give the first real trial a matched history (on each experimental run one trial type had 17 repetitions with the first trial being discarded).

Figure 1A shows the temporal structure of the trials. The duration of each trial was 4.86 s (3 TRs). Each trial began with the 400 ms presentation of the cue word followed by 250 ms of fixation and then the 250 ms presentation of the test scene stimulus. Following presentation of the stimulus was a period of fixation lasting 250–750 ms followed by display of the response cue. After the response cue was removed, subjects were presented with a central fixation dot.

Observers were required to indicate how confident they were that the cued target object was present or absent in the scene using an 8-point confidence scale by pressing one of the buttons of the response boxes (Lumina fMRI Response Pads, Cedrus Corporation). The 8-point rating scale consisted of the four buttons on the response pad for each hand. To control for the effects of motor activity associated with the subjects' behavioral response, the response mapping between present or absent

and each hand was determined by which response cue was displayed. The response cues consisted of either a red square or a green triangle and were randomized among trials (equal probability for either cue). The association between the response hand and the response cue was counterbalanced across observers.

Saliency control experiment. To control for the low-level physical properties across the target-present and target-absent image sets, we devised a control experiment using the image sets but with the cue word replaced by a fixation that persisted until the scene stimulus was displayed. Instead of searching for a target object, observers were instructed to search for saliency regions within the image and reported which side of the scene stimulus (left vs right) they believed contained the most salient feature. Responses were recorded with a single button press either with the left or the right hand. As with the main experiment, the hand the observers used to respond was determined by a response cue presented after the scene stimulus. The saliency search task was conducted to control for the possibility of a correlation between saliency and the target object in the main study.

fMRI data acquisition. Data were collected at the UCSB Brain Imaging Center using a 3T TIM Trio Siemens Magnetom with a 12 channel phased-array head coil. An echo-planar (EPI) sequence was used to measure BOLD contrast (TR = 1620 ms; TE = 30 ms; flip angle = 65° , FOV = 192 mm; slice thickness = 3.5 mm, matrix = 78×78 , 29 coronal slices) for experimental runs. For one subject we used a sequence with a 210 mm FOV (with all other parameters remaining constant) to accommodate the subject's larger brain. Localizer scans used a higher resolution sequence (TR = 2000 ms; TE = 35 ms; flip angle = 70° , FOV = 192 mm; slice thickness = 2.5 mm, matrix = 78×78 , 30 coronal slices). A high resolution T1-weighted MPRAGE scan (1 mm^3) was also acquired for each participant (TR = 2300 ms; TE = 2.98 ms; flip angle = 9° , FOV = 256 mm; slice thickness = 1.1 mm, matrix = 256×256).

Localization of regions of interest

Retinotopic regions. We identified regions of interest (ROIs) for each subject individually using standard localizer scans in conjunction with each subject's anatomical scan, which were both acquired in a single scanning session lasting $\sim 1 \text{ h } 45 \text{ min}$. These ROIs included areas in early visual cortex, high-level visual cortex and the frontoparietal attention network. The ROIs in retinotopic visual cortex were defined using a rotating wedge and expanding concentric ring checkerboard stimulus in two separate scans (Sereno et al., 1995). Each rotating wedge and concentric ring-mapping scan lasted 512 s, consisting of 16 s of initial and final fixation with 8 full rotations of the wedge or 8 full expansions of the ring stimulus, each lasting 60 s. Wedge stimuli in the polar-mapping scan had a radius of 14° and subtended 75° consisted of alternating black and white squares in a checkerboard pattern. The colors of the checkerboard squares flickered rapidly between black and white at 4 Hz to provide constant visual stimulation. To ensure that subjects maintained central fixation, a gray dot at fixation darkened for 200 ms at pseudo-random intervals, and subjects were required to indicate with a button press when this occurred. The eccentricity mapping procedure was the same except that a ring expanding from fixation (width 4°) was used instead of a rotating wedge. By correlating the BOLD response resulting from the activation caused by the wedge and ring stimuli, we determined which voxels responded most strongly to regions in visual space, and produced both polar and eccentricity maps on the surface of visual cortex that were mapped onto meshes of the individual subject's anatomy.

The center of V1 overlaps the calcarine sulcus and represents the whole contralateral visual field with its edges defined by changes in polar map field signs designating the start of V2d and V2v. Areas V2d, V2v, V3d, and V3v all contain quarter-field representations with V3d adjacent to V2d and V3v adjacent to V2v. Area V4 contains a full field representation and shares a border with V3v (Tootell and Hadjikhani, 2001; Tyler et al., 2005). The anterior borders of these regions were defined using the eccentricity maps. Area V3A and V3B are dorsal and anterior to area V3d, with which they share a border, and contain a full hemifield representation of visual space. To separate them it is necessary to refer to the eccentricity map that shows a second foveal confluence at the border between V3A and V3B (Tyler et al., 2005).

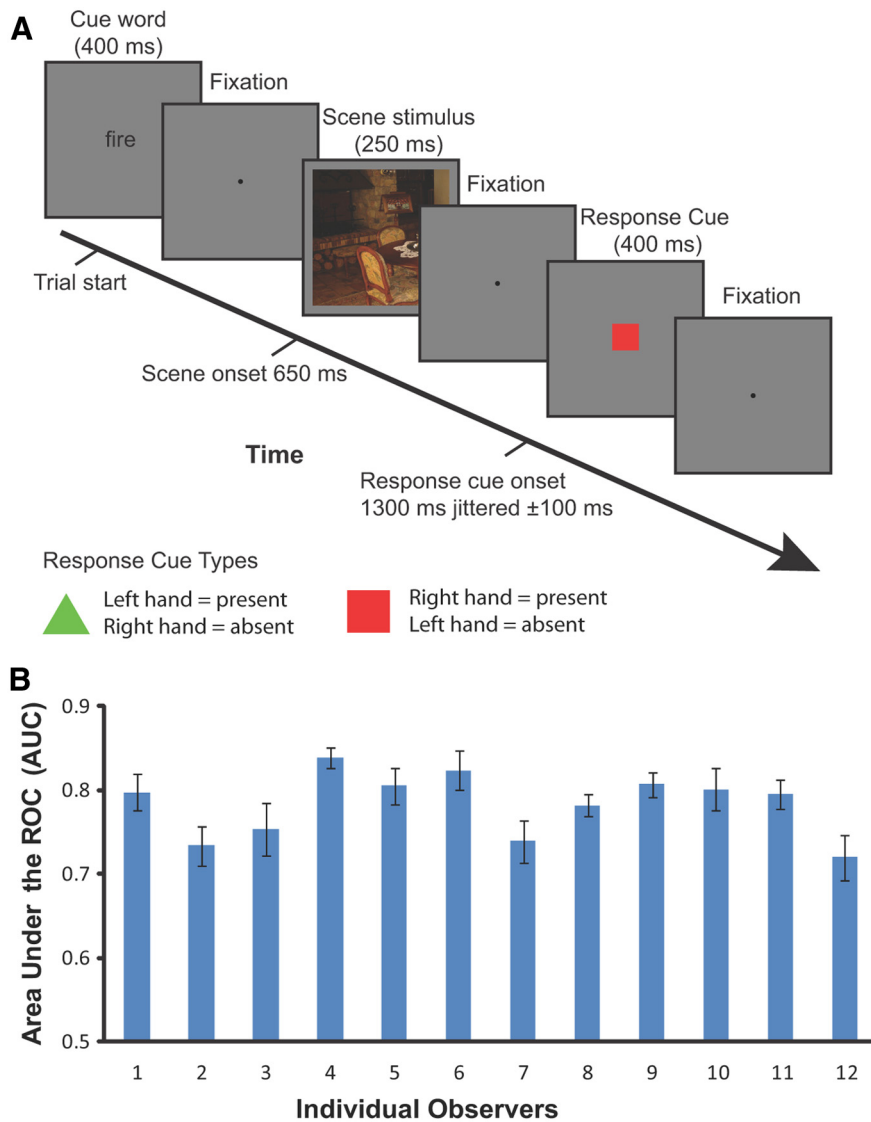


Figure 1. Overview of the experimental paradigm and behavioral results. **A**, Timeline of events in a trial: each trial was 4.86 s (3 TRs); the trial began with the 400 ms presentation of the cue word followed by 250 ms of fixation and then the 250 ms presentation of the scene stimulus. Following presentation of the stimulus was a period of fixation lasting 250–750 ms followed by 400 ms display of the response cue. After the response cue was removed, subjects were presented with a central fixation dot. Observers responded expressing their confidence about the presence or absence of the cued target object using an 8-point confidence scale. Observers were trained to swap hands when pressing the response buttons depending on the response cue that was randomized among trials. **B**, The behavioral performances (quantified by the AUC values) of individual observers detecting the presence of target objects in natural scenes during the visual search study. Error bars correspond to SEs of the mean across sessions.

Functional regions. We used functional localizers to identify the area hMT+/V5, LOC, fusiform face area (FFA), and parahippocampal place area (PPA). We also defined areas implicated in the processing of eye-movements and spatial attention: the frontal and supplementary eye-fields (FEF and SEF) and an area in IPS. Area hMT+/V5 was defined as the set of voxels in lateral temporal cortex that responded significantly higher ($p < 10^{-4}$) to a coherently moving array of dots than to a static array of dots (Zeki et al., 1991). The scan lasted for 376 s consisting of 8 s of initial and final fixation with 18 repeats of 20 s blocks. There were three block types consisting of black dots on a mid-gray background viewed through a circular aperture. Dots were randomly distributed within the circular aperture with a radius of 13 degrees and a dot density of 20 dots/deg². During the moving condition, all dots moved in the same direction with a speed of 3 deg/s for 1 s before reversing direction. In the edge condition, strips of dots (width 2 deg) moved with opposite motion to each other creating kinetically defined borders. To define area hMT+/V5, a GLM analysis was performed and the activation resulting from the

contrast, moving > stationary dots, was used to define the region constrained by individual anatomy to an area within the inferior temporal sulcus.

The LOC, FFA, and PPA localizers were combined into a single scan to maximize available scanner time. In the combined localizer scan, referred to as the LFP localizer, the scan duration was 396 s consisting of three, 12 s periods of fixation; at the beginning end and the middle of the run, and five repeats of the 4 experimental conditions each lasting 18 s. These conditions were as follows: intact objects, phase-scrambled objects, intact faces, and intact scenes. During the 18 s presentation period, each stimulus was presented for 300 ms followed by 700 ms of fixation before the next stimulus presentation. To maximize statistical power, the LFP scan was run twice for each individual using different trial sequences. To define areas of functional activity, a GLM analysis was performed on the two localizer scans. Area LOC was defined as the activation revealed by the intact objects > scrambled objects contrast (Kourtzi and Kanwisher, 2001). The FFA region was isolated by the face stimuli > intact object stimuli contrast (Kanwisher et al., 1997). The PPA regions by the scenes > faces + objects contrast (Epstein and Kanwisher, 1998). All localizer regions were guided by the known anatomical features of their areas reported by previous groups.

We localized brain areas in the frontoparietal attention network by adapting an eye-movement task developed by Connolly et al. (2002). The task consisted of eight repeats of two blocks each lasting 20 s: one where a fixation dot was presented centrally and a second where every 500 ms the fixation dot was moved to the opposite side of the screen. During the moving dot condition, the dot could be positioned anywhere along a horizontal line perpendicular to the vertical meridian and between 4 and 15 degrees away from the screen center. Subjects were required to make saccades to keep the moving dot fixated. Contrasting the two conditions in a GLM analysis revealed activation in the FEFs, SEFs, and a region in the IPS (Table 1).

We defined a control ROI centered on each subject's hand motor area from BOLD activity resulting from hand use identified using a GLM contrasting all task trials versus fixation trials. During fixation trials, subjects were not required to respond and so there would be no hand use. The hand area ROIs were then defined about the activity resulting from this contrast in the central sulcus (primary motor cortex), guided by reference to studies directly investigating motor activity resulting from hand or finger use (Lotze et al., 2000; Alkadhi et al., 2002).

fMRI data analysis. We used FreeSurfer (<http://surfer.nmr.mgh.harvard.edu/>) to determine the gray-white matter and gray matter-pial boundaries for each observer from their anatomical scans, which were then used to reconstruct inflated and flattened 3D surfaces. In addition, a group average mesh was constructed using spherical surface-based cortical registration to allow more accurate group analyses. Functional runs were preprocessed using FSL 4.1 (<http://www.fmrib.ox.ac.uk/fsl/>) to perform 3D motion correction, alignment to individual anatomical scans, high-pass filtering (3 cycles per run), and linear trend removal. No spatial smoothing was performed on the functional data used for the multivariate analysis. SPM8 (

Table 1. The mean Talairach coordinates for the regions in the frontoparietal network

| Area | Left hemisphere | | | | Right hemisphere | | | |
|------|-----------------|-------|------|-----------------|------------------|-------|------|-----------------|
| | x | y | z | mm ² | x | y | z | mm ² |
| FEF | −37.4 | −5.7 | 50.9 | 1385.7 | 34.0 | −9.4 | 50.3 | 1565.3 |
| SEF | −7.8 | −3.0 | 58.6 | 1079.6 | 6.2 | −6.6 | 57.4 | 1090.8 |
| IPS | −21.2 | −62.9 | 50.8 | 2034.5 | 20.3 | −61.4 | 54.0 | 1988.7 |

ion.ucl.ac.uk/spm/) was used to perform GLM analyses of the localizer and experimental scans. Regions of interest were defined on the inflated mesh and projected back into the space of the functional data.

Multivoxel pattern analysis. Multivariate pattern classification analysis (MVPA) has been successfully applied to fMRI data to evaluate the information content of multivoxel activation patterns in targeted brain regions (Norman et al., 2006; Weil and Rees, 2010; Serences and Saproo, 2012). Here, we used regularized LDA (Duda et al., 2000) to classify the patterns of fMRI data within each ROI. The regions we investigated included V1, V2d, V2v, V3A, V3B, V3d, V3v, V4, LOC, hMT+/V5, RSC, FFA, PPA, FEF, SEF, and IPS. We used the union of corresponding ROIs in the left and right hemispheres to construct a single region for each ROI. We normalized (*z*-score) each voxel time course separately for each experimental run to minimize baseline differences between runs. The initial data vectors for the multivariate analysis were then generated by shifting the fMRI time series by three TRs to account for the hemodynamic response lag. All classification analyses were performed on the mean of the three data-points collected for each trial with the first data-point taken from the start of the trial when the cue word is presented. For each single trial, the output of the multivariate pattern classifier was a single scalar value generated based on the weighted sum of the input values across all voxels in one specific ROI. We used a leave-one-out cross validation scheme across runs. The classifier learned a function that mapped between voxel activity patterns and experimental conditions from nine of the 10 runs. Given a new pattern of activity from a single trial in the left-out testing run, the trained classifier determined whether this trial belonged to target-present or target-absent condition.

In addition, we used a multivariate “searchlight” analysis (Kriegeskorte et al., 2006) to determine the discrimination performance in areas of the brain for which we had not defined ROIs. The searchlight method estimates the MVPA classification performance of the voxels contained within a sphere of a given radius about a central voxel. This spherical searchlight is then moved systematically over each voxel in the brain and the procedure repeated with the resulting discrimination performance assigned to the voxel at the center of the searchlight. In this way a map of the classification performance over the whole brain can be assessed. We used a searchlight analysis where the bounding sphere had a radius of 9 mm and contained 153 voxels. For each group of voxels we applied the same LDA classification technique as that used for the ROI-based analysis described previously. The searchlight analysis was performed on each individual’s data in their native anatomical space. To produce the group map the individual searchlight maps were transformed into the group average space by using Freesurfer’s cortical surface matching algorithm before statistical testing. In brief, the surface matching procedure involves inflating all subjects’ white matter meshes into unit spheres while retaining the cortical depth information on the 2D surface of the sphere. The cortical geometry of all subjects is then matched by transformation of the spherical surfaces and the resulting registration can then be applied to volumetric data allowing, for example, BOLD activity to be projected into the group average space.

Evaluation using area under ROC. For each individual observer, the performance of both behavioral response and MVPA decision variables were quantified by the area under the receiver operating curve, referred to as AUC (DeLong et al., 1988). The value of AUC was calculated using a non-parametric algorithm that quantifies, for each scalar value (8-point confidence ratings or MVPA decision variables derived the multivariate pattern classifier) corresponding to the target-present trials, the probability that it will exceed the responses to all the values corresponding to the target-absent trials in the test dataset.

$$A_z = \frac{1}{PA} \sum_p^P \sum_a^A \text{step}(\lambda_p - \lambda_a) + \frac{1}{2} \delta(\lambda_p - \lambda_a), \quad (1)$$

where P and A are the total number of target-present and target-absent trials; λ_p and λ_a are the scalar values for the p th target-present trial and a th target-absent trial, respectively; and step is the heavyside step function defined as follows:

$$\text{step}(x) = \begin{cases} 1, & \text{if } x > 0 \\ 0, & \text{if } x < 0 \end{cases}. \quad (2)$$

This function measures the frequency in which a given response to a target-present trial exceeds the response to the target-absent trial.

The impulse function δ quantifies the instance in which the values for the target-present and target-absent trials are a tie. The function δ is defined as follows:

$$\delta(x) = \begin{cases} 1, & \text{if } x = 0 \\ 0, & \text{if } x \neq 0 \end{cases}. \quad (3)$$

Choice probabilities. In addition to discriminating the experimental conditions (target-present vs target-absent), we also conducted a choice probability analysis of the single-trial fMRI data. The trials were labeled based on the observers’ choices of response rather than the type of the stimulus conditions. For the correct trials, the labels based on choice or stimulus conditions were the same; whereas for the incorrect trials, the labels derived from the behavioral choices were used regardless of the stimulus conditions. The same MVPA procedures and evaluation of performance described above were applied for the choice probability.

Image-specific analysis. For each individual image, we computed the proportion of observers responding to that specific image as target-present, as well as the proportion of MVPAs applied to individual brains categorizing the same image as target-present. We thus obtained two scalar values corresponding to behavior and neural activity, respectively. A scalar value of 0 represented that all of the observers (or MVPA) labeled one particular image as target-absent, while the value of 1 represented that all of the observers (or MVPA) labeled that image to be target-present. Similarly, the value of 0.5 represented that half of the observers (or MVPA) decided the image to be target-present and the other half decided the image to be target-absent condition. The 640 pairs of scaled values corresponding to the 640 images were computed for each ROI, resulting in a correlation coefficient associated with each ROI.

False discovery rate control to correct for multiple comparisons. Unless mentioned otherwise, all multiple hypothesis tests were false discovery rate (FDR) corrected. The FDR control procedure described by Benjamini and Hochberg (1995) was used to correct for multiple comparisons. Let H_1, \dots, H_m be the null hypotheses and P_1, \dots, P_m their corresponding p values. These p values were ordered in ascending order and denoted by $P_{(1)}, \dots, P_{(m)}$. For a given α , find the largest k such that $p_{(k)} \leq \frac{k}{m} \alpha$, then reject all $H_{(i)}$ for $i = 1, \dots, k$. The FDR-corrected p values were computed by dividing each observed p value by its percentile rank to get an estimated FDR.

Eye-tracking data collection and analysis. Nine of the 12 observers had their eye position recorded during the fMRI experimental scans using an Eyelink 1000 eye-tracker (Eyelink). This system uses fiber optics to illuminate the eye with infra-red light and tracks the eye orientation using the pupil position relative to the corneal reflection. Observers completed a nine-point calibration before the first experimental run, and repeated the calibration before subsequent runs if the calibration began to degrade due to head motion. All visual stimuli presented on the screen were within the limits of the calibration region. An eye movement was recorded as a saccade if both velocity and acceleration exceeded a threshold (velocity > 30 °/s; acceleration > 8000 °/s²). A saccade outside an area extending 1° from the fixation was considered an eye-movement away from fixation. Data collected during the whole trial of 4.86 s was analyzed, starting from the frame corresponding to the onset of the stimuli display of one trial until the frame of the end of a trial (before the display onset of the following trial). An interval of 4.86 s was examined for any

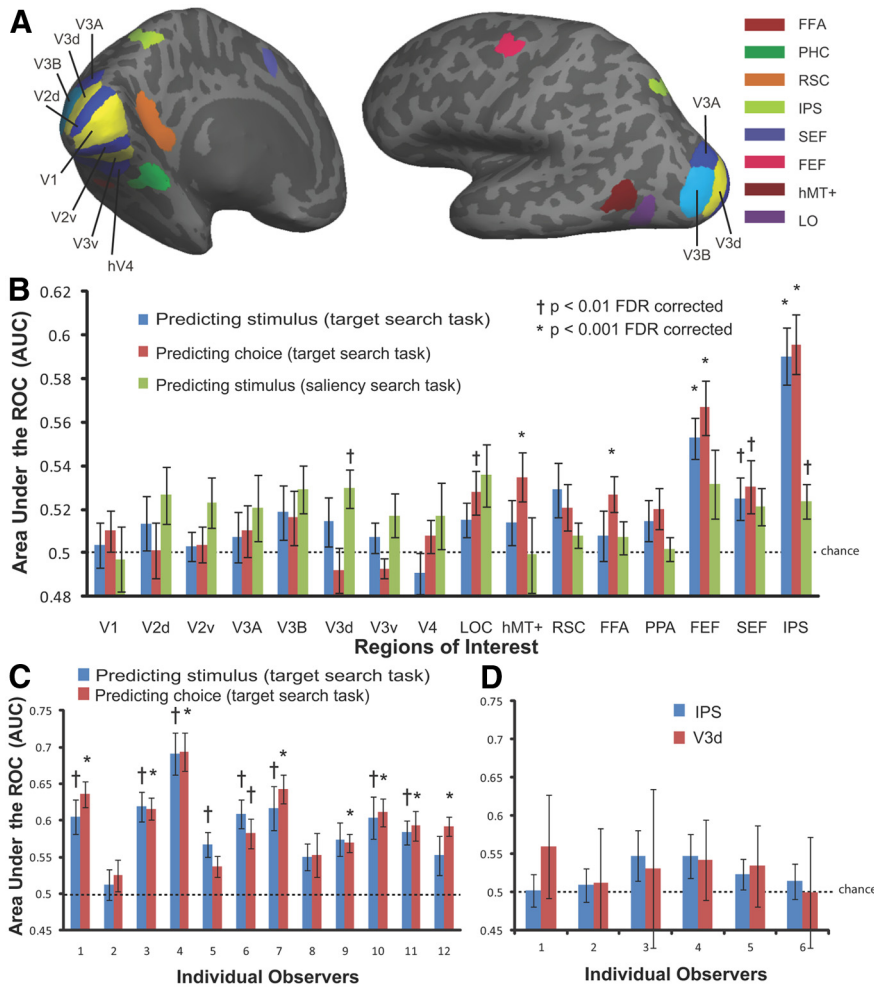


Figure 2. Multi-voxel pattern classifier analysis. **A**, Regions of interest defined by retinotopic and functional localizer scans. **B**, Average MVPA performance (AUC averaged across the 12 observers in the main study) for each ROI predicting stimulus type (target-presence vs target-absence) and observers' choices during the target search task; average MVPA performance (AUC averaged across the 6 observers in the control study) for each ROI predicting stimulus type in the saliency search task. **C**, Individual MVPA performance (IPS) of the 12 observers in the main study predicting stimulus type and observers' choices during the target search task. **D**, Individual MVPA performance (IPS and V3d) of the six observers in the control study predicting stimulus during the saliency search task.

eye movements occurring within the trial that could potentially affect the BOLD signal, starting with the onset of the cue word and including the display and response periods. The first analysis examined three measures across target-present and target-absent conditions: (1) the mean total number of saccades, (2) the mean of saccade amplitude (in degrees of visual angle), and (3) the standard deviation of distance of eye position from central fixation on a trial. The second analysis examined the measure of average distance from eye position to central fixation, which is an absolute value of relevant eye positions. We used the average distance of horizontal and vertical coordinates of eye positions to the horizontal and vertical coordinates of the center fixation as a two-dimensional input to the pattern classifier to discriminate the presence or absence of the target objects in the natural scenes.

Results

Behavioral performance

Figure 1B shows the behavioral performance deciding on the presence versus absence of the target for each individual observer during the scanning session. The AUC values of the 12 observers in the target search study ranged from 0.72 (± 0.03 , SEM) to 0.84 (± 0.01 , SEM), with a mean of 0.78 and SEM of 0.01 across all observers.

To control for the effects of target saliency, a new set of observers were presented the same set of images that were used in the main study and reported a binary decision whether the right or left side of the image appears more salient. If the main target detection task and the saliency control task were orthogonal, then the observer saliency judgments should be at chance predicting the presence of the targets specified in the main study. We used the binary decisions of the saliency judgment (left vs right) corresponding to the target-present and the target-absent image-sets to obtain a target detection AUC. The AUC values for six observers were 0.43, 0.45, 0.47, 0.47, 0.47, 0.45 (mean 0.46, SEM 0.01; if the assignment of right vs left saliency is switched for target presence vs absence, the resulting mean AUC is 0.54). For each of the six observers, the AUC values discriminating target-present versus target-absent were not significantly different from chance (p values of each observer were 0.02, 0.02, 0.23, 0.28, 0.20, 0.05, after FDR correction no p value satisfied $p < 0.05$) confirming that the target search tasks and saliency tasks were close to orthogonal.

Pattern classifier predicting stimulus type and choice probability in ROIs defined by localizers

The first MVPA analysis investigated which cortical areas contained information about the detection of the arbitrary target objects within natural scenes. We restricted our analysis to the ROIs identified by standard localizers. These ROIs include visual areas from V1 to V4, hMT+/V5, LOC, FFA, PPA, RSC, and the frontoparietal network regions including IPS, FEF, and SEF (Fig. 2, Table

1). Within each ROI, we used the pattern classifier to predict from single-trial fMRI data whether the displayed natural image contained the target object or not. Second, we performed a choice probability analysis on a trial-by-trial basis to investigate the contribution of different cortical areas.

Figure 2B displays the AUC value averaged across 12 observers for each ROI separately. The average AUC is significantly higher than chance in IPS and FEF ($p < 0.001$) for discriminating between the target-present versus target-absent experimental conditions. MVPA performance for all the low visual areas, the object-, face-, and scene-selective areas could not predict target presence above chance level (the FDR-corrected p values of these regions were all larger than 0.01). Similarly, the classifier performance was highest in IPS and FEF for discriminating single-trial behavioral choices of observers. A repeated-measures ANOVA showed a significant difference across ROIs for both classifying stimulus ($F_{(15,165)} = 6.05, p < 0.001$) and behavioral choice ($F_{(15,165)} = 6.85, p < 0.001$), but no significant difference between stimulus and choice ($F_{(1,352)} = 0.72, p = 0.40$). The average classifier performance discriminating choice was slightly higher

than discriminating stimulus type in IPS and FEF, but not statistically significant ($p = 0.33$). Overall, IPS was the best predictor, with an average AUC of 0.59 ± 0.01 for discriminating target-present versus target-absent trials and AUC 0.60 ± 0.01 for predicting observers' single trial choice.

The classifier performances discriminating stimulus and choice using fMRI activity in the region of IPS for each individual observer are shown in Figure 2C. In general, the classification analysis revealed similar trends across the 12 observers when discriminating stimulus and choice but with considerable variation in the classifier performance (Fig. 2C).

Saliency control task to quantify contributions of physical differences between target-present and target-absent image sets

Using the same set of 640 images, observers in the control study performed a saliency search task that was orthogonal to the target detection task in the main study. To assess the possible contributions of the low-level physical differences across the target-absent and target-present image sets to the differences in BOLD activity in the main study, we conducted the same pattern classification analysis discriminating target presence/absence with the saliency control data. Figure 2B shows classifier performances for each region of interest (third column). We applied the classifier to the single-trial fMRI data of the saliency control study to predict the whether an image contained the target-present or not (as specified in the main study). Figure 2B shows that for the saliency task, no classifier performance for ROIs discriminating target-present versus target-absent conditions was significantly above chance at the criteria of $p < 0.001$ FDR corrected. The AUC values of two regions, V3d (AUC = 0.53, $p = 0.006$) and IPS (AUC = 0.52, 0.007), did show statistical significance at the criteria of $p < 0.01$ FDR corrected. To investigate whether the significance of group average performance was driven by specific observers, the AUC values for each individual observers were shown in Figure 2D. The classifier performance for each individual observer using fMRI activity in the IPS and V3d in the saliency search task was not significantly different from chance (the FDR corrected p values of all observers were larger than 0.05). Thus, the results indicate that the effect of the neural activity in the frontoparietal network discriminating target presence/absence in target search task cannot be attributed to differing low-level physical properties across the image sets.

Correlation between MVPA decision variable and behavioral confidence rating

The choice probability analysis in the previous section shows that neural activity can classify the binarized behavioral decisions (target-present or target-absent) of the human observers. We also investigated whether there was a relationship between the classifier scalar decision variable as extracted with the MVPA and the

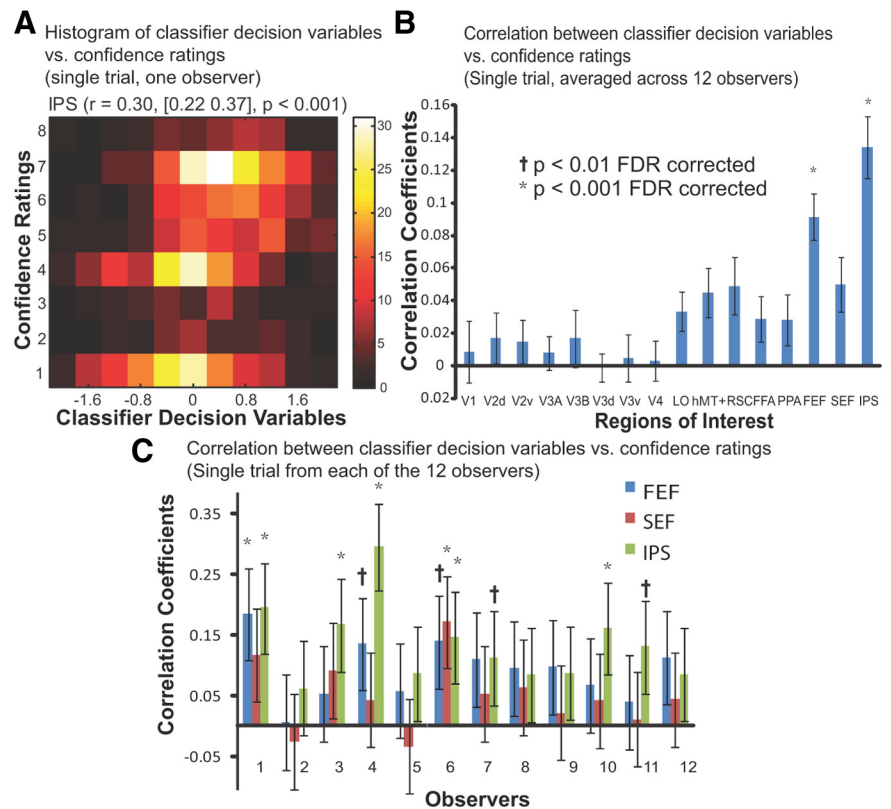


Figure 3. Correlation between MVPA decision variables and behavioral confidence ratings (single-trial analysis). **A**, The two-variable histogram of the behavioral confidence ratings versus the classifier decision variables computed from fMRI response in IPS from single trials of one representative observer. The correlation coefficient between the two variables was 0.30, with a confidence interval of [0.22 0.37]. **B**, Average correlation coefficients across 12 observers between the behavioral confidence ratings versus the classifier decision variables computed from fMRI activity of each ROI on a trial-by-trial basis. **C**, The correlation coefficients in regions within the frontoparietal network for each individual observer. Error bars show the 95% confidence interval.

confidence level of the observers' decisions. If the measured classifier decision variable indeed is related to sensory evidence for the presence/absence of a target object, then a significant positive correlation between the two variables should exist.

Figure 3A depicts the two-variable histogram of the confidence ratings versus the classifier decision variables computed from fMRI response in IPS on a trial-by-trial basis for one representative observer. The color of the two-variable histogram codes the number of trials falling into each frequency bin. The correlation between the two variables is significant, with a correlation coefficient of 0.30 (95% confidence interval [0.22 0.37], $p < 0.001$). In Figure 3B, the correlation coefficients for all the ROIs of each individual observer were computed and then averaged across all the observers to show the relative strength of correlation among regions. Consistent with our stimulus classification and choice probability results, FEF and IPS resulted in the strongest correlation [$r = 0.09$ (FEF), $r = 0.13$ (IPS), $p < 0.001$]. Furthermore, Figure 3C shows the correlation coefficients resulting from the three regions within the frontoparietal network (FEF, SEF, IPS), for each individual observer. There is a large amount of variance in the strength of the correlation and significance levels across the observers (Fig. 3C).

Image-specific relationship between behavioral performance and neural activity

We conducted an image-specific analysis to determine whether there is a relationship between the observers' behavioral difficulty in finding a target with specific images and the neural activity

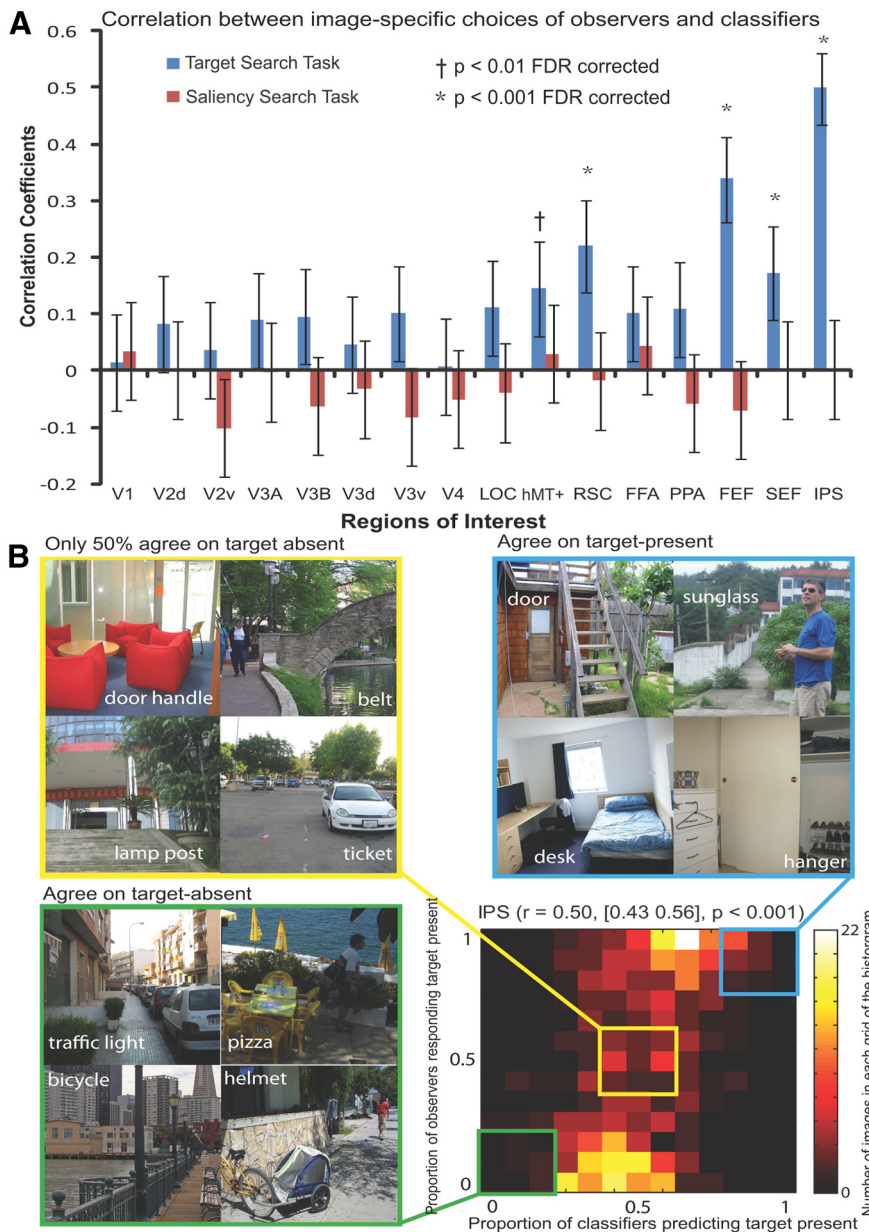


Figure 4. Image-specific relationship between human behavior and neural activity. **A**, For each image, the proportion of observers responding to that image as target-present and the proportion of pattern classifiers predicting from neural activity the same image as target-present were quantified, resulting in a pair of scalar values. The 640 pairs of scaled values in each ROI were correlated to compute a correlation coefficient for each ROI. **B**, The 640 pairs of scaled values in the region of IPS were plotted in a grid figure, in which the color indicates the number of images that are in each grid. Three groups of images were simultaneously illustrated: images that observers or pattern classifiers highly agreed to be target-present, those highly agreed to be target-absent, and those highly disagreed to be target-present or target-absent.

elicited in observers’ brains to that image. Note that a significant correlation between individual observers’ single trial confidence rating and the MVPA decision variable does not necessarily imply a relationship between observers’ ensemble behavioral decisions for an image and their collective neural responses. For example, if each image elicits idiosyncratic neural responses and behavioral decisions across observers, then there should not be any relationship between the MVPA decision variable and ensemble observer behavior for a specific image.

We investigated the image-specific correlation between the observers’ choices and the choices based on pattern classifiers using single-trial fMRI activity. For each of the 640 images, we

quantified the proportion of observers responding to that image as target-present and the proportion of pattern classifiers classifying that same image as target-present. The 640 pairs of scaled values of each image were computed for each ROI, resulting in a correlation coefficient associated with each ROI (Fig. 4A, first column). We found significant positive correlation ($p < 0.001$) in several regions including IPS, SEF, FEF, RSC, and hMT+/V5; still, IPS had the strongest correlation ($r = 0.50$, 95% confidence interval $[0.43\ 0.56]$, $p < 0.001$) compared with other regions. For IPS, we plotted the pairs of scaled values of each image as a two-point histogram (Fig. 4B), in which the horizontal-axis and vertical-axis correspond to the neural and behavioral relative frequencies predicting target-present. The color indicates the number of images within each bin. The image-specific analysis indicates that the proportion of human observers behaviorally choosing an image as target-present and the proportion of observers predicted by the MVPA from fMRI activity was significantly correlated in the frontoparietal attention network, indicating a strong relationship between image-specific human behavior and neural activity.

We also investigated the image-specific correlation between averaged confidence ratings and the classifier decision variables averaged for each individual images. Figure 5B shows a two-variable scatter plot of the behavioral confidence ratings versus the classifier decision variables computed from fMRI response in IPS for each individual image averaged across 12 observers (expected maximum 12 trials), revealing a significant positive correlation between the two variables ($r = 0.53$, 95% confidence interval $[0.46\ 0.59]$, $p < 0.001$). Similar analyses for all the ROIs (Fig. 5A) indicate significant correlation ($p < 0.001$) in several regions including IPS, SEF, FEF, PPA, RSC, and hMT+/V5, although IPS has the strongest correlation compared with other ROIs.

The image-wise analysis allows us to scrutinize individual sample images corresponding to different bins in the histogram. Figure 4B shows three groups of natural images, in which the semantically associated target objects were indicated by the cue words. The three groups illustrated correspond to the following: images that all observers and all pattern classifiers highly agreed to be target-present (Fig. 4B, top right), images that all observers and all pattern classifiers highly agreed to be target-absent (Fig. 4B, bottom left), and images that for which there was high disagreement among observers and among pattern classifier about its class: target-present or target-absent (Fig. 4B, top left). To prevent selection bias, the four sample images associated with each group were randomly picked from all the images within each of the

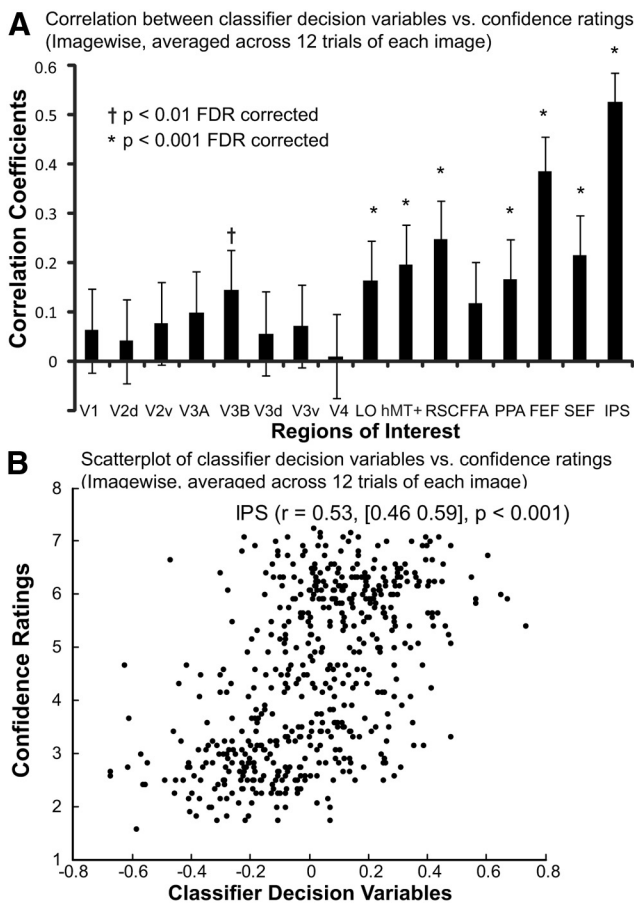


Figure 5. Image-specific correlation between MVPA decision variables and behavioral confidence ratings (averaged across observers for each individual image). **A**, Correlation coefficients between the behavioral confidence ratings versus the classifier decision variables computed from fMRI activity of each ROI for each individual image averaged across 12 observers (expected maximum 12 trials). **B**, The two-variable scatter plot of the behavioral confidence ratings versus the classifier decision variables computed from fMRI response in IPS for each individual image. The correlation coefficient between the two variables was 0.53, with a confidence interval of [0.46 0.59].

three cells in the two-dimensional histogram. Visual scrutiny of the images for the three groups provides a qualitative impression that high agreement across observers and classifiers about target presence or absence is associated with relatively “easier” images. In contrast, high decision disagreement among observers and pattern classifiers is corresponding to relatively “difficult” images.

We conducted a similar image-wise analysis with the saliency control data to rule out the possible effect of low-level physical differences across the target-present and target-absent image sets. Figure 4A shows (second column in each bar figure) the correlation coefficients computed from the saliency control fMRI data for each ROI and the behavioral decisions from the main study. None of the ROIs resulted in statistically significant correlation (the FDR corrected p values of all ROIs were all larger than 0.05) between the behavioral choices of observers in the main study and the choices based on pattern classifier decision variables using fMRI activity recording during the saliency task.

Quantifying the effect of eye movements using fixation patterns and classifier analysis

Even though observers were instructed to maintain steady fixation during our studies, involuntary eye movements could con-

taminate our results. To assess the possible contributions of eye movements to the pattern classifier’s ability to discriminate the presence of targets and observer choices, we analyzed the eye-tracking data of the nine observers (the first three observers’ eye-tracking data were not available due to technical issues). We found no significant differences in the mean total number of saccades, the mean saccade amplitude, or the SD of distance from eye position to central fixation on a trial between target-present and target-absent conditions ($p = 0.54$, $p = 0.33$, $p = 0.60$ for the three metrics, respectively). The various eye movement metrics are summarized in Table 2. In addition, we conducted a multivariate pattern classifier analysis based on the two-variable measure of average distance of horizontal and vertical coordinates of eye positions to the two coordinates of the central fixation. Table 3 shows the classifier performance discriminating target-present versus target-absent trials was not significantly above chance (AUC = 0.50, $p = 0.74$), using the average distance of eye position from central fixation on a trial. The AUC values of all the nine observers were also not higher than chance (the FDR corrected p values of all ROIs were all larger than 0.05), indicating that the ability of the neural activity in IPS and FEF to discriminate target presence/absence cannot be driven by consistent differences in eye movement patterns.

Quantifying the effects of motor activity associated with the behavioral response

To avoid a consistent stimulus-response mapping, the response mapping between present or absent and each hand was determined by which response cue (randomized among trials) was displayed. In addition, the association between the response hand and the response cue was counterbalanced across observers. To further rule out the alternative explanation that motor activity influenced classification performance, we defined a bilateral ROI from the hand regions of left and right primary motor cortex for each observer to estimate the effectiveness of the control. We conducted pattern classifier analyses using neural activity in the hand regions of the primary motor cortex (combined left and right regions) and showed that the MVPA performance was not significantly different from chance discriminating target-present versus target-absent trials (group averaged AUC = 0.52, SEM = 0.01, $p = 0.11$), indicating that the ability of the neural activity in the frontoparietal network to discriminate target presence/absence cannot be explained by motor activity.

Quantifying the effects of response time differences between target-present and target-absent conditions

Many visual search studies have shown that the RT for target-absent trials is longer than that for target-present trials (Wolfe, 1998). We found the same trend of RT differences in our visual search experiment: the group average RT for target-present trials was 1.03 s while the group average RT for target-absent trials was 1.20 s. The RT difference of 0.17 s between target-present and target-absent conditions was statistically significant ($p < 0.05$). To investigate whether the RT differences between experimental conditions can account for the neural activity in frontoparietal network discriminating target-presence versus target-absence, we conducted an MVPA analysis using neural activity in each ROI to classify short-RT trials versus long-RT trials, regardless of the target presence/absence conditions. If RT differences were the only factors that account for the MVPA performance classifying target-present versus target-absent trials, then we should be able to classify short-RT trials versus long-RT trials with a comparable performance. We defined short-RT trials as those trials with

Table 2. Measures of eye movements

| Sub | Mean no. of saccades (SEM) | | Mean saccades amplitude (SEM) | | Standard deviation of distance from fixation (SEM) | |
|-------|----------------------------|----------------|-------------------------------|----------------|--|----------------|
| | Target-absent | Target-present | Target-absent | Target-present | Target-absent | Target-present |
| Sub4 | 6.54 (0.60) | 6.23 (0.62) | 2.58 (0.32) | 2.62 (0.30) | 1.46 (0.35) | 1.54 (0.37) |
| Sub5 | 0.52 (0.31) | 0.49 (0.28) | 1.81 (0.43) | 1.66 (0.26) | 0.56 (0.53) | 0.19 (0.10) |
| Sub6 | 2.76 (0.75) | 2.52 (0.70) | 2.22 (0.39) | 2.20 (0.38) | 0.85 (0.31) | 0.80 (0.28) |
| Sub7 | 3.93 (0.74) | 3.83 (0.70) | 2.09 (0.26) | 2.14 (0.26) | 0.87 (0.31) | 0.88 (0.31) |
| Sub8 | 3.88 (0.57) | 4.02 (0.56) | 1.88 (0.20) | 2.05 (0.22) | 0.73 (0.21) | 0.89 (0.22) |
| Sub9 | 6.56 (0.82) | 6.65 (0.80) | 3.79 (0.52) | 3.80 (0.48) | 2.67 (0.49) | 2.75 (0.48) |
| Sub10 | 4.77 (0.92) | 4.68 (0.87) | 1.73 (0.21) | 1.74 (0.23) | 0.54 (0.19) | 0.54 (0.20) |
| Sub11 | 2.77 (0.57) | 2.92 (0.52) | 3.37 (0.57) | 3.74 (0.50) | 1.38 (0.42) | 1.44 (0.39) |
| Sub12 | 4.10 (0.86) | 4.16 (0.86) | 4.85 (0.53) | 4.80 (0.53) | 2.97 (0.57) | 2.72 (0.57) |
| Mean | 3.98 (0.63) | 3.94 (0.63) | 2.70 (0.36) | 2.75 (0.37) | 1.34 (0.30) | 1.31 (0.30) |

Table 3. Classifier discriminating target presence/absence using average absolute distance of eye position from fixation on a trial

| | Sub4 | Sub5 | Sub6 | Sub7 | Sub8 | Sub9 | Sub10 | Sub11 | Sub12 | Mean |
|------------------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Classifier performance (SEM) | 0.53 (0.02) | 0.51 (0.04) | 0.53 (0.02) | 0.55 (0.02) | 0.47 (0.02) | 0.46 (0.02) | 0.48 (0.03) | 0.47 (0.03) | 0.53 (0.02) | 0.50 (0.01) |

Table 4. Group average classifier performance discriminating short-RT versus long-RT trials

| ROI | Classifier performance (SEM) |
|---------|------------------------------|
| V1 | 0.50 (0.01) |
| V2d | 0.51 (0.01) |
| V2v | 0.52 (0.01) |
| V3A | 0.52 (0.01) |
| V3B | 0.52 (0.01) |
| V3d | 0.50 (0.01) |
| V3v | 0.51 (0.01) |
| V4 | 0.52 (0.01) |
| LOC | 0.51 (0.01) |
| hMT+/V5 | 0.50 (0.01) |
| RSC | 0.49 (0.01) |
| FFA | 0.51 (0.01) |
| PPA | 0.50 (0.01) |
| FEF | 0.51 (0.01) |
| SEF | 0.50 (0.01) |
| IPS | 0.51 (0.01) |

shorter RTs than the median RT and long-RT trials as those with times longer than the median RT. Table 4 shows that no ROI resulted in above chance performance in discriminating short-RT from long-RT trials (the FDR corrected p values of all ROIs were all larger than 0.05), indicating that the RT differences cannot account for the effects in the frontoparietal network discriminating target presence/absence. In addition, if differences in RTs across conditions account for the ability of the MVPA to discriminate target present versus absent trials, then we might also expect a positive correlation between RT differences and MVPA performance discriminating target present/absent trials: observers with high RT differences might have higher associated MVPA AUC discriminating target present versus absent trials. This was not the case. The correlation between the RT differences and the AUC values in the IPS ($r = 0.07$, $p = 0.84$), FEF ($r = 0.07$, $p = 0.83$), SEF ($r = 0.00$, $p = 0.99$), and other ROIs ($p > 0.05$) were not statistically significant.

Searchlight analysis

To further investigate whether other brain regions aside from the ones identified by our functional localizers (e.g., the frontoparietal network) were involved in the visual search task, we conducted a spherical search light analysis measuring patterns

discriminating target-presence versus target-absence. Figure 6 shows the results for the classification analysis displayed on a flattened mesh constructed from the average mesh of the 12 observers. Classification accuracy was significantly above chance ($p < 0.005$ uncorrected; t test across all subjects against chance) in three regions covered by our ROIs, the frontal eye fields, supplementary eye fields, and the intraparietal sulcus. However, significant discrimination performance was also found in left ventral premotor cortex (PMv), right dorsolateral prefrontal cortex (DLPFC), and anterior insular cortex (AIC). The Talairach coordinates of the above areas are shown in Table 5.

Discussion

Feature-independent coding of target detection

Neuroimaging studies (Peelen et al., 2009; Walther et al., 2009; Park et al., 2011) have identified two mechanisms of scene perception. One is the representation of spatial boundaries of scenes in PPA (Kravitz et al., 2011; Park et al., 2011); the other is the representation of content and/or category in LOC (Peelen et al., 2009; Park et al., 2011). The scene categorization tasks are effectively visual search tasks, restricted to detecting a subset category of target objects in natural scenes (e.g., animal, face, car). Unknown is whether the brain regions revealed by scene categorization tasks also mediate the detection of arbitrary target objects in scenes.

Our results showed that pattern classifiers using single-trial fMRI activity in the frontoparietal network can discriminate the presence of 368 different target objects in natural scenes. Furthermore, the neural activity correlated with observers' trial-by-trial decisions and decision confidence. Together, the results suggest the existence of a neural mechanism for feature-independent coding of search targets in natural scenes in the frontoparietal network. Our results do not imply that object-related areas do not code target objects in natural scenes. Instead, we suggest that different targets do not elicit a similar pattern of activity in the object areas. If each target elicited a unique object-specific pattern of activity in LOC, then it is unlikely that MVPA would reliably detect the presence of a target across the variety of objects in our experiment. In contrast, a study with a single target object type would elicit consistent activity in LOC that could be decoded by MVPA.

The role of the frontoparietal network in feature-independent target detection

Neuroimaging studies have shown that the categorical properties of natural scenes are encoded in object-selective cortex (Peelen et al., 2009) and that perceptual decisions are mediated by the lateral intraparietal area and the dorsolateral prefrontal cortex (Heekeren et al., 2008). We found above chance MVPA performance for discriminating target presence/absence in the frontoparietal network but not in early visual areas, object-, face-, and scene-selective areas. Our findings converge with studies implicating the frontoparietal network in visual search. Bisley and colleagues proposed that LIP acts as a priority map of behavioral relevance (Bisley and Goldberg, 2010; Bisley, 2011). FEF can signal the target location in visual search (Schall, 2002; Juan et al., 2004; Thompson et al., 2005; Buschman and Miller, 2007), and this neural selection process does not necessarily require saccade execution (Thompson et al., 1997). Our findings provide evidence that the frontoparietal network is not only involved in search of synthetic displays, but also the detection of target objects in scenes. Moreover, our results suggest the existence of a neural mechanism for feature-independent coding of targets during search of scenes regardless of their category. This is also consistent with the feature-independent coding property of the frontoparietal network with synthetic displays (Shulman et al., 2002; Giesbrecht et al., 2003; Giesbrecht and Mangun, 2005; Slagter et al., 2007; Greenberg et al., 2010; Ptak, 2011).

Critically, the saliency control, the eye movement controls, the motor activity control, and the analysis of RT differences rule out important alternative interpretations of our results. Saliency is a concern because the images sets between conditions were physically different. To rule out effects of low-level physical differences across image sets (VanRullen and Thorpe, 2001), we conducted a control experiment using the same stimuli and an orthogonal saliency search task. The results showed that no ROI was significantly above chance at discriminating target-present versus target-absent image sets, indicating that the effect of the neural activity in the frontoparietal network classifying target presence/absence cannot be attributed to low-level physical properties across the image sets.

Another concern is that the MVPA results were driven by systematic differences in eye movements between conditions. Ruling out this alternative is particularly important because FEF and IPS are known to be activated by eye movements and attention (Corbetta, 1998; Gitelman et al., 2000; Bisley, 2011; Ptak, 2011). We minimized eye movements by using brief image presentations, instructing observers to maintain fixation, and monitoring observers' eye movements. The statistical analyses and a separate MVPA analysis of the eye movement data did not reveal significant differences or discriminatory information across target-present and target-absent conditions. These analyses indicate that the MVPA results classifying target presence cannot be accounted for by differential eye-movement patterns.

Motor activity is another potential concern. The pattern classifier analysis using neural activity in the primary motor cortex showed that motor activity cannot discriminate target-presence versus target-absence. This control analysis, combined with the response cue procedure that randomized the response hand on a trial-by-trial basis, support the notion that the ability of the neu-

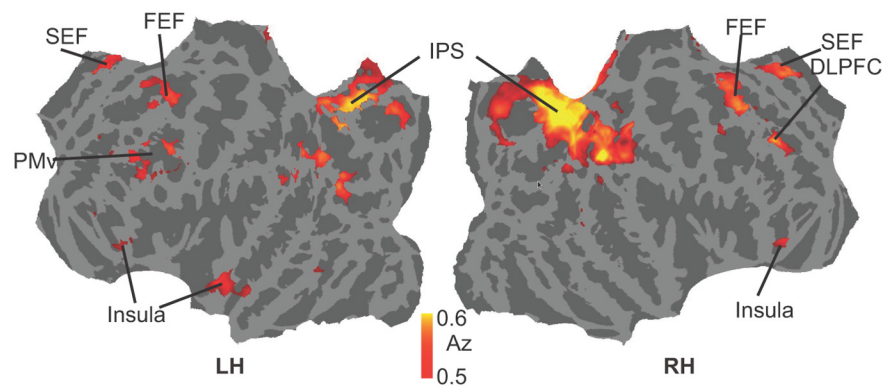


Figure 6. Group-average searchlight analysis discriminating target-presence versus target-absence at $p < 0.005$ (uncorrected).

Table 5. Talairach coordinates of areas discriminating target-presence versus target-absence significantly above chance for the group

| Area | Left hemisphere | | | Right hemisphere | | |
|---------------|-----------------|-------|------|------------------|-------|-------|
| | x | y | z | x | y | z |
| FEF | −27.4 | −2.5 | 42 | 21.1 | −1.4 | 49 |
| SEF | −9.1 | 6.9 | 47 | 7.6 | 13.3 | 44.7 |
| IPS | −17.7 | −68.1 | 48.7 | 18.4 | −64.2 | 48.4 |
| PMv | −34.1 | 13.9 | 22.4 | NA | NA | NA |
| DLPFC | NA | NA | NA | 33.6 | 35.7 | 26.5 |
| Insula cortex | −54.3 | 7.7 | −9.5 | 20.9 | 21.4 | −13.3 |

$p < 0.005$ uncorrected; t test across all subjects against chance.

ral activity in the frontoparietal network to discriminate target presence/absence cannot be explained by motor activity.

Another concern is that the MVPA performance could potentially be explained by RT differences between the target-present and target-absent trials. A review of the RT effects on BOLD activity (Yarkoni et al., 2009) revealed the areas (lateral and ventral prefrontal cortex, anterior insular cortex, and anterior cingulate cortex) whose BOLD amplitude significantly correlated with response time. Notably, the areas affected did not include IPS or FEF (Yarkoni et al., 2009). A control MVPA analysis showed that no ROI discriminated between short-RT and long-RT trials, indicating that the RT differences cannot account for the MVPA performance in the frontoparietal network. There is the possibility that the measured RTs mostly reflected the retrieval of the appropriate response selection rule corresponding to the color response cue rather than time on task related to the search task. Such a case would leave open the possibility that a difference in time on task across target present and absent trials might account for our results. However, even in the scenario that time on task is resulting in differential activity across target present versus absent trials, it would be unclear as to why the differential pattern of activity would be observed only in IPS and FEF and not in other vision areas.

Other association areas in target detection

The searchlight analysis we conducted revealed brain regions not included in our ROI analysis that could discriminate between target-present and target-absent trials (Fig. 6). These were left PMv, right DLPFC, and AIC. PMv is well placed to combine information necessary for perceptual decision taking input from visual cortex (Markowitsch et al., 1987), SEF (Ghosh and Gattera, 1995), parietal areas (Borra et al., 2008), and prefrontal cortex (Boussaoud et al., 2005). Its role in motor planning and connec-

tion to primary motor cortex (Ghosh and Gattera, 1995) suggest a direct link between decision making and response selection. Its ability to discriminate between conditions in our task may reflect a general perceptual decision-making process rather than one specific to visual search. The discrimination performance observed in right DLPFC may also be more related to a general perceptual decision-making process than one specific to visual search (Heekeren et al., 2004). AIC is activated in many studies investigating decision making and has been implicated in a range of processes including pain perception (Ploghaus et al., 1999), self recognition (Devue et al., 2007), conscious awareness (Craig, 2009), and lapses of attention (Weissman et al., 2006). The wide range of processes in which AIC is activated suggests that it is involved in more general decision-making processes rather than visual search.

Neural activity and decision confidence

Our results show that the classifier decision variables extracted from the frontoparietal network not only relate to binary decisions but also correlate with the confidence of human decisions when searching for targets in natural scenes. This finding agrees with recent monkey neurophysiology work showing that LIP activity represents formation of the decision and the degree of certainty underlying the decision (Kiani and Shadlen, 2009).

Predicting image-specific behavioral difficulty from ensemble neural activity

We investigated the relationship between neural activity and behavioral responses on an image-by-image basis. If the decisions are greatly determined by the content of the images, then we might expect that certain images would tend to elicit specific behavioral responses and associated brain activity patterns. Images with difficult search targets might elicit “target-absent” responses from a majority of observers and also neural activity that is consistent with target-absent experimental condition. If so, then it might be possible to predict the behavioral difficulty of images using aggregates of neural activity across multiple brains. Our results suggest that this is indeed the case (Fig. 4). Those images for which observers (or MVPA) rarely agreed on the decision can be subjectively identified as hard images. We found that several ROIs exhibited a significant relationship between image-specific decisions and neural activity with the highest correlation in the frontoparietal network (Fig. 4). The parietal distribution of the correlation is consistent with studies that have shown parietal representation of task difficulty (Philiastides et al., 2006; Kiani and Shadlen, 2009).

Conclusion

We provided evidence using fMRI and MVPA that regions in the frontoparietal network can code the presence of a variety of target objects in natural scenes arguing for a feature-independent coding mechanism. The activity in these regions is tightly linked with choice, decision confidence, and task difficulty.

References

Alkadhi H, Crelier GR, Boendermaker SH, Golay X, Hepp-Reymond MC, Kollias SS (2002) Reproducibility of primary motor cortex somatotopy under controlled conditions. *AJNR Am J Neuroradiol* 23:1524–1532.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Stat Soc Ser B* 57:289–300.

Bisley JW (2011) The neural basis of visual attention. *J Physiol* 589:49–57.

Bisley JW, Goldberg ME (2003) Neuronal activity in the lateral intraparietal area and spatial attention. *Science* 299:81–86.

Bisley JW, Goldberg ME (2006) Neural correlates of attention and distractibility in the lateral intraparietal area. *J Neurophysiol* 95:1696–1717.

Bisley JW, Goldberg ME (2010) Attention, intention, and priority in the parietal lobe. *Annu Rev Neurosci* 33:1–21.

Borra E, Belmalih A, Calzavara R, Gerbella M, Murata A, Rozzi S, Luppino G (2008) Cortical connections of the macaque anterior intraparietal (AIP) area. *Cereb Cortex* 18:1094–1111.

Boussaoud D, Tanné-Gariépy J, Wannier T, Rouiller EM (2005) Callosal connections of dorsal versus ventral premotor areas in the macaque monkey: a multiple retrograde tracing study. *BMC Neuroscience* 6:67.

Buracas GT, Boynton GM (2002) Efficient design of event-related fMRI experiments using M-sequences. *Neuroimage* 16:801–813.

Buschman TJ, Miller EK (2007) Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* 315:1860–1862.

Codispoti M, Ferrari V, Junghöfer M, Schupp HT (2006) The categorization of natural scenes: brain attention networks revealed by dense sensor ERPs. *Neuroimage* 32:583–591.

Connolly JD, Goodale MA, Menon RS, Munoz DP (2002) Human fMRI evidence for the neural correlates of preparatory set. *Nat Neurosci* 5:1345–1352.

Corbetta M (1998) Frontoparietal cortical networks for directing attention and the eye to visual locations: Identical, independent, or overlapping neural systems? *Proc Natl Acad Sci U S A* 95:831–838.

Corbetta M, Kincade JM, Ollinger JM, McAvooy MP, Shulman GL (2000) Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nat Neurosci* 3:292–297.

Craig AD (2009) How do you feel—now? The anterior insula and human awareness. *Nat Rev Neurosci* 10:59–70.

DeLong ER, DeLong DM, Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 44:837–845.

Devue C, Collette F, Baiteau E, Degueldre C, Luxen A, Maquet P, Brédart S (2007) Here I am: the cortical correlates of visual self-recognition. *Brain Res* 1143:169–182.

Duda RO, Hart PE, Stork DG (2000) *Pattern classification*, ed 2. New York: Wiley-Interscience.

Epstein R, Kanwisher N (1998) A cortical representation of the local visual environment. *Nature* 392:598–601.

Fabre-Thorpe M, Richard G, Thorpe SJ (1998) Rapid categorization of natural images by rhesus monkeys. *Neuroreport* 9:303–308.

Ghosh S, Gattera R (1995) A comparison of the ipsilateral cortical projections to the dorsal and ventral subdivisions of the macaque premotor cortex. *Somatosens Mot Res* 12:359–378.

Giesbrecht B, Mangun GR (2005) Identifying the neural systems of top-down attentional control: a meta-analytic approach. In: *Neurobiology of attention* (Itti L, Rees G, Tsotsos JK), pp 63–68. Burlington: Academic.

Giesbrecht B, Woldorff MG, Song AW, Mangun GR (2003) Neural mechanisms of top-down control during spatial and feature attention. *Neuroimage* 19:496–512.

Gitelman DR, Parrish TB, LaBar KS, Mesulam MM (2000) Real-time monitoring of eye movements using infrared video-oculography during functional magnetic resonance imaging of the frontal eye fields. *Neuroimage* 11:58–65.

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.

Heekeren HR, Marrett S, Bandettini PA, Ungerleider LG (2004) A general mechanism for perceptual decision-making in the human brain. *Nature* 431:859–862.

Heekeren HR, Marrett S, Ungerleider LG (2008) The neural systems that mediate human perceptual decision making. *Nat Rev Neurosci* 9:467–479.

Hopfinger JB, Buonocore MH, Mangun GR (2000) The neural mechanisms of top-down attentional control. *Nat Neurosci* 3:284–291.

Ipata AE, Gee AL, Goldberg ME, Bisley JW (2006) Activity in the lateral intraparietal area predicts the goal and latency of saccades in a free-viewing visual search task. *J Neurosci* 26:3656–3661.

Ipata AE, Gee AL, Bisley JW, Goldberg ME (2009) Neurons in the lateral intraparietal area create a priority map by the combination of disparate signals. *Exp Brain Res* 192:479–488.

- Juan CH, Shorter-Jacobi SM, Schall JD (2004) Dissociation of spatial attention and saccade preparation. *Proc Natl Acad Sci U S A* 101:15541–15544.
- Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci* 17:4302–4311.
- Kastner S, Pinsk MA, De Weerd P, Desimone R, Ungerleider LG (1999) Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron* 22:751–761.
- Kiani R, Shadlen MN (2009) Representation of confidence associated with a decision by neurons in the parietal cortex. *Science* 324:759–764.
- Kincade JM, Abrams RA, Astafiev SV, Shulman GL, Corbetta M (2005) An event-related functional magnetic resonance imaging study of voluntary and stimulus-driven orienting of attention. *J Neurosci* 25:4593–4604.
- Kourtzi Z, Kanwisher N (2001) Representation of perceived object shape by the human lateral occipital complex. *Science* 293:1506–1509.
- Kravitz DJ, Peng CS, Baker CI (2011) Real-world scene representations in high-level visual cortex: it's the spaces more than the places. *J Neurosci* 31:7322–7333.
- Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping. *Proc Natl Acad Sci U S A* 103:3863–3868.
- Lotze M, Erb M, Flor H, Huelsmann E, Godde B, Grodd W (2000) fMRI evaluation of somatotopic representation in human primary motor cortex. *Neuroimage* 11:473–481.
- Markowitsch HJ, Irl E, Emmans D (1987) Cortical and subcortical afferent connections of the squirrel monkey's (lateral) premotor cortex: evidence for visual cortical afferents. *Int J Neurosci* 37:127–148.
- Norman KA, Polyn SM, Detre GJ, Haxby JV (2006) Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn Sci* 10:424–430.
- Ossandón JP, Helo AV, Montefusco-Siegmund R, Maldonado PE (2010) Superposition model predicts EEG occipital activity during free viewing of natural scenes. *J Neurosci* 30:4787–4795.
- Park S, Brady TF, Greene MR, Oliva A (2011) Disentangling scene content from spatial boundary: complementary roles for the parahippocampal place area and lateral occipital complex in representing real-world scenes. *J Neurosci* 31:1333–1340.
- Peelen MV, Fei-Fei L, Kastner S (2009) Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature* 460:94–97.
- Philiastides MG, Ratcliff R, Sajda P (2006) Neural representation of task difficulty and decision making during perceptual categorization: a timing diagram. *J Neurosci* 26:8965–8975.
- Phillips AN, Segraves MA (2010) Predictive activity in macaque frontal eye field neurons during natural scene searching. *J Neurophysiol* 103:1238–1252.
- Ploghaus A, Tracey I, Gati JS, Clare S, Menon RS, Matthews PM, Rawlins JN (1999) Dissociating Pain from Its Anticipation in the Human Brain. *Science* 284:1979–1981.
- Ptak R (2011) The frontoparietal attention network of the human brain: action, saliency, and a priority map of the environment. *The Neuroscientist*. Advance online publication. Retrieved August 5, 2011. doi: 10.1177/1073858411409051.
- Schall JD (2002) The neural selection and control of saccades by the frontal eye field. *Philos Trans R Soc Lond B Biol Sci* 357:1073–1082.
- Serences JT, Saproo S (2012) Computational advances towards linking BOLD and behavior. *Neuropsychologia* 50:435–446.
- Sereno AB, Maunsell JH (1998) Shape selectivity in primate lateral intraparietal cortex. *Nature* 395:500–503.
- Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RB (1995) Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268:889–893.
- Shulman GL, d'Avossa G, Tansy AP, Corbetta M (2002) Two attentional processes in the parietal lobe. *Cereb Cortex* 12:1124–1131.
- Slagter HA, Giesbrecht B, Kok A, Weissman DH, Kenemans JL, Woldorff MG, Mangun GR (2007) fMRI evidence for both generalized and specialized components of attentional control. *Brain Res* 1177:90–102.
- Thomas NW, Paré M (2007) Temporal processing of saccade targets in parietal cortex area LIP during visual search. *J Neurophysiol* 97:942–947.
- Thompson KG, Bichot NP, Schall JD (1997) Dissociation of visual discrimination from saccade programming in macaque frontal eye field. *J Neurophysiol* 77:1046–1050.
- Thompson KG, Biscoe KL, Sato TR (2005) Neuronal basis of covert spatial attention in the frontal eye field. *J Neurosci* 25:9479–9487.
- Thorpe S, Fize D, Marlot C (1996) Speed of processing in the human visual system. *Nature* 381:520–522.
- Tootell RB, Hadjikhani N (2001) Where is “dorsal V4” in human visual cortex? Retinotopic, topographic and functional evidence. *Cereb Cortex* 11:298–311.
- Tyler CW, Likova LT, Chen CC, Kontsevich LL, Schira MM, Wade AR (2005) Extended concepts of occipital retinotopy. *Curr Med Imag Rev* 1:319–329.
- VanRullen R, Thorpe SJ (2001) The time course of visual processing: from early perception to decision-making. *J Cogn Neurosci* 13:454–461.
- Walther DB, Caddigan E, Fei-Fei L, Beck DM (2009) Natural scene categories revealed in distributed patterns of activity in the human brain. *J Neurosci* 29:10573–10581.
- Weil RS, Rees G (2010) Decoding the neural correlates of consciousness. *Curr Opin Neurol* 23:649–655.
- Weissman DH, Roberts KC, Visscher KM, Woldorff MG (2006) The neural bases of momentary lapses in attention. *Nat Neurosci* 9:971–978.
- Wolfe JM (1998) Visual search. In: *Attention* (H. Pashler, ed), pp 13–73. London: University College London.
- Yarkoni T, Barch DM, Gray JR, Conturo TE, Braver TS (2009) BOLD correlates of trial-by-trial reaction time variability in gray and white matter: a multi-study fMRI analysis. *PLoS ONE* 4:e4257.
- Zeki S, Watson JD, Lueck CJ, Friston KJ, Kennard C, Frackowiak RS (1991) A direct demonstration of functional specialization in human visual cortex. *J Neurosci* 11:641–649.