

Preparatory patterns of neural activity predict visual category search speed

Chun Siong Soon^{*}, Praneeth Namburi, Michael W.L. Chee^{**}

Cognitive Neuroscience Laboratory, Program in Neuroscience and Behavioral Disorders, Duke-NUS Graduate Medical School, Singapore

ARTICLE INFO

Article history:

Accepted 20 October 2012

Available online 26 October 2012

Keywords:

fMRI

LOC

Preparatory attention

MVPA

Visual semantics

ABSTRACT

Rapidly detecting target object categories when objects are embedded in naturalistic scenes is facilitated by preparatory baseline signal changes. However, it is unclear as to what information most strongly predicts perceptual speed in terms of the minimal exposure duration required for accurate detection. Using novel surface-based spatiotemporal pattern classification, we found that while category-specific biases resulting from merely providing a category name can be detected in multiple cortical areas, only biases in lateral occipital complex predicted perceptual speed. These biases likely carry visual semantic information regarding multiple object categories placed in familiar scene contexts. Discriminatory voxels during the preparatory period showed congruent category-selectivity during visual stimulation.

© 2012 Elsevier Inc. All rights reserved.

Introduction

The human visual system faces tremendous challenges when trying to make sense of the constant influx of complex sensory information. An enduring question remains how we can often detect familiar objects embedded in background scenes with sub-second speed and accuracy (Intraub, 1981; Li et al., 2002; Potter and Faulconer, 1975; Potter and Levy, 1969; Thorpe et al., 1996). Biasing the processing of sensory information prior to target appearance is one enabling mechanism, but details of how this might assist categorical target detection in natural scenes continue to evolve (Wolfe et al., 2011).

Early experiments involving abstract visual stimuli revealed preparatory changes in baseline neural activity after one is oriented to search at specified locations or for particular features (Chelazzi et al., 1998; Giesbrecht et al., 2003; Hopfinger et al., 2000; Kastner et al., 1999; Luck et al., 1997; Ress et al., 2000). Baseline signal changes bias sensory processing in functionally specialized visual cortical areas by increasing neuronal activation, and may enhance target detection or discrimination (Giesbrecht et al., 2006; Ress et al., 2000; Stokes et al., 2009; Sylvester et al., 2009).

Subsequent studies involving natural objects such as houses/place scenes and faces showed that orienting to isolated exemplars of these categories results in dissociable baseline changes within specialized face and place areas respectively (Esterman and Yantis, 2010; Puri et

al., 2009; Reddy et al., 2010). Additional advances in our understanding of the visual processing of objects involved extending studies of visual search to other objects embedded in natural scenes. Such explorations are crucial in light of behavioral studies that indicate differences between natural scene searches and those involving abstract or isolated stimuli typically used in experimental settings (Wolfe et al., 2011). In real life, we search for diverse types of objects, but disambiguating imaging signals for object categories that do not activate spatially distinct visual areas, as is the case for faces and places, is problematic with conventional image analysis methods that examine signal changes at single-voxel level.

Multi-voxel pattern analyses (MVPA) utilize differences in activation magnitude between neighboring fMRI voxels to uncover informative shifts in neural activity concealed from traditional univariate approaches (Haynes and Rees, 2006; Kriegeskorte et al., 2006; Norman et al., 2006). In addition to overcoming aforesaid limitations of conventional univariate approaches to MR signal analysis, MVPA may be especially useful when informative patterns do not significantly alter baseline MR signal across a cortical patch (Bode and Haynes, 2009; Bode et al., 2012; Soon et al., 2008). Recently, MVPA was used to distinguish preparatory patterns between a pair of object categories (body parts versus cars) embedded in natural scenes, despite their activating the same cortical patch (Peelen and Kastner, 2011). Verifying that this remarkable finding generalizes to the search for other object categories is of interest in the present study. In particular, testing with four different categories would strengthen the proposition that preparatory signal changes originate from the categorization of varied object categories as opposed to being driven by responses to single category – more likely when only a pair of object classes is tested.

To derive informative patterns attributable to different target categories, we applied surface-based MVPA to repeated pair-wise comparisons across four object categories, including two that do not have distinct ventral visual cortex representations.

^{*} Correspondence to: C.S. Soon, Brain and Consciousness Laboratory, Neuroscience and Behavioral Disorders Program, Duke-NUS Graduate Medical School, 8 College Rd, #06-18, Singapore 169857, Singapore. Fax: +65 62218625.

^{**} Correspondence to: M.W.L. Chee, Cognitive Neuroscience Laboratory, Neuroscience and Behavioral Disorders Program, Duke-NUS Graduate Medical School, 8 College Rd, #06-18, Singapore 169857, Singapore. Fax: +65 62218625.

E-mail addresses: chunsiong.soon@duke-nus.edu.sg (C.S. Soon), michael.chee@duke-nus.edu.sg (M.W.L. Chee).

Detecting a category when presented with only a verbal label as opposed to predefined exemplars (Cichy et al., 2012; Reddy et al., 2010) requires broad prior knowledge of the visual characteristics of the category and its associated contexts (Castelhano and Heaven, 2010). In contrast to keeping in mind a specific reference object, access to such an abstract representation would be expected to involve higher visual areas (Esterman and Yantis, 2010; Peelen and Kastner, 2011; Puri et al., 2009) rather than early visual cortex (Kosslyn et al., 1999).

A recent seminal study suggested that more pronounced category-specificity of preparatory patterns within higher visual cortex improves categorization speed and accuracy (Peelen and Kastner, 2011). An intriguing complement to this finding would be to investigate whether such preparatory patterns predict the *stimulus exposure duration* required for detecting familiar categories in natural scenes. Being able to utilize the most discriminatory elements that differentiate visual object classes would contribute to explaining why for some persons, *exposure* as brief as 20 ms suffices for detecting a target category within a complex scene (Thorpe et al., 1996). Crucially there is a distinction between *stimulus presentation time* and *processing time* (Vanrullen, 2011), the latter being reflected in reaction time (Peelen and Kastner, 2011).

Lastly, preparing to attend to a particular stimulus may resemble actually viewing it (Cichy et al., 2012; Driver and Frith, 2000; Reddy et al., 2010). For isolated objects and well-defined exemplars, patterns during visualization and perception have been remarkably congruent (Cichy et al., 2012; Reddy et al., 2010). We thus determined whether the most category selective voxels during the preparatory period would continue to show congruent category-specificity during the stimulus-evoked phase of the experiment.

Materials and methods

Participants

21 healthy right-handed participants (mean age 20.8, range 19–24 years; 9 males) gave informed consent prior to undergoing brain imaging with fMRI. Six participants were excluded from analyses due to excessive sudden head motion (see Supplementary Methods for details).

Experimental stimuli

In the MR scanner, a cued category detection task was performed. For each of four possible target categories, 'House', 'Face', 'Car' or 'Bird', 304 complex colored images were obtained from publically available sources (see Supplementary Fig. 1). Each image depicted an object from the specific category as the main subject. This object was embedded in a complex natural scene that did not contain any object from the other three categories (each 'House' image contained the complete external view of a building). All images were randomly assigned to training and experimental trials for each participant. Each object or image was used only once across trials to avoid learning or repetition effects.

Visual stimuli were back-projected (Epson EMP1715, 800×600 pixels, 60 Hz) onto a screen at the rear of the scanner bore. Participants responded with two custom-made button boxes, one in each hand. Auditory stimuli were presented through a MR-compatible headset (Siemens, Erlangen, Germany).

Experimental task

Participants maintained fixation on a cross in the center of the display throughout the experiment. At the beginning of every trial, an auditory cue, 'House', 'Face', 'Car' or 'Bird', specified the target object category (Fig. 1). 6000 ms after the auditory cue onset, four images were presented, one from each category. Each display quadrant contained a randomly assigned image (280×210 pixels; visual angle: closest corner,

3°, furthest corner, 24°). Each quartet of natural scene images was shown for between 16.7 ms and 183.3 ms. Exposure duration was titrated for each participant such that performance accuracy was approximately 85% (see below). The image quartets were immediately followed by 3 different sets of identically configured masks comprising scrambled images, each appearing for 250 ms.

At the same time that the natural scene quartet was presented, the white fixation cross was replaced by a blue fixation cross for 2350 ms. The participant then indicated the location of the image containing the target category by pressing one of four buttons corresponding to quadrants the target image was located. A variable fixation period followed, such that trial onsets were separated by 16–20 s to minimize overlap in the hemodynamic response.

Study protocol

Participants underwent 4 training runs before performing 10 experimental fMRI runs. Each run consisted of 16 trials, 4 from each target category. The first 2 training runs were conducted outside the scanner using a liquid crystal display monitor (800×600 pixels, 60 Hz) to ensure that participants understood and could perform the task. The last 2 training runs were conducted inside the scanner to ensure that the participants could perform the task under identical visual stimulation conditions as the actual fMRI experiment.

The duration of each image quartet was titrated such that each participant achieved target identification accuracy of approximately 85%. In the first training run, the quartets appeared for 100 ms (6 frames). At the end of each run (including training runs), if the participant was able to identify the image containing the target category with >90% accuracy, the stimulus exposure duration was reduced by 1 frame (16.7 ms) for the next run. If accuracy was <80%, the exposure duration was increased by 1 frame for the next run. Otherwise, exposure duration was kept the same. Based on this procedure, exposure duration provided an indicator of each participant's perceptual speed for target category detection that was not affected by differences in motor response speed.

Imaging acquisition

MR images were acquired on a 3-Tesla Tim Trio system (Siemens, Erlangen, Germany). Ten runs comprising 155 functional MRI volumes each were acquired for each participant using a gradient echo-planar imaging (EPI) sequence with the following parameters: repetition time (TR) 2000 ms; echo time 30 ms; flip angle 90 degrees; field of view 180×180 mm; 72×72 pixel matrix per slice; 2.5×2.5 mm in-plane resolution; 36 oblique axial slices; slice thickness 2.50 mm; inter-slice gap 0.25 mm. In order to reconstruct the cortical surface for surface-based MVPA, a high-resolution T1-weighted anatomical volume with 1 mm isotropic voxels was also obtained using a 3D-MPRAGE sequence.

General linear model

All functional images were realigned to the first image of the first functional run using rigid-body transformation (SPM2, <http://www.fil.ion.ucl.ac.uk/spm>). No further smoothing or spatial normalization was performed on the functional data in order to maximally preserve the fine-grained activation patterns.

The first two images in each run were discarded to allow for magnetic saturation effects. A voxel-wise general linear model (GLM) was then created to capture signal variance that was common across all conditions. All correct trials from the four target categories were modeled together using a set of 11 finite impulse response (FIR) predictors, covering 22 s from cue onset. Error trials were modeled using a separate set of 11 FIR predictors. Motion correction parameters and their first-order derivatives were included as covariates.

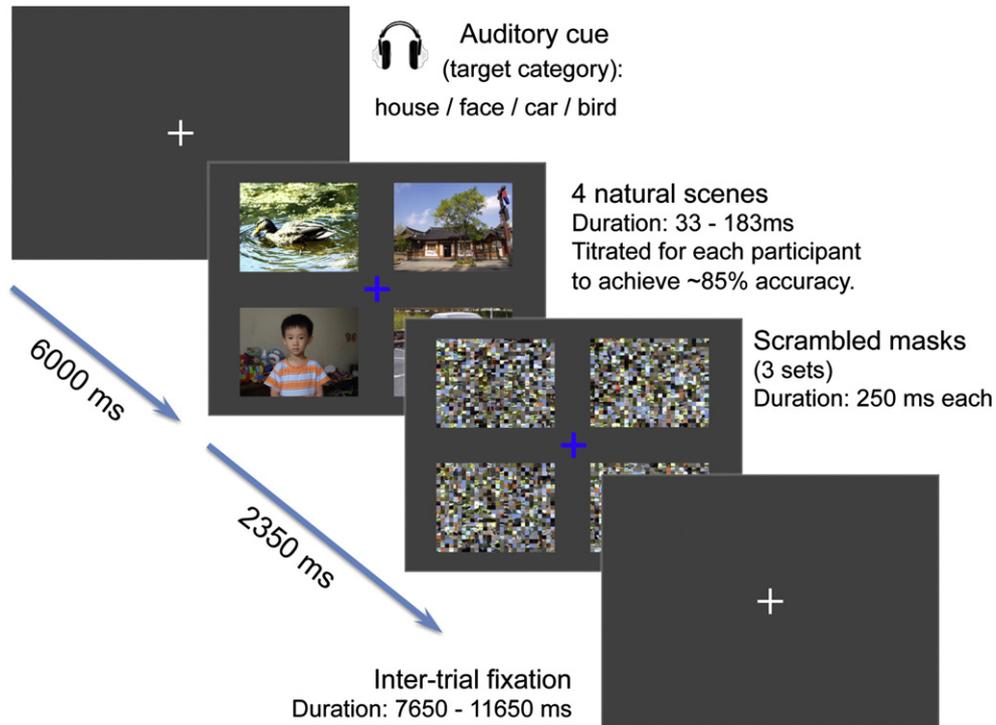


Fig. 1. Schematic of target category detection task. At the beginning of each trial, an auditory word cue specified the target category: 'House', 'Face', 'Car' or 'Bird'. 6000 ms later, four natural scene images were briefly presented before being masked by scrambled images. Participants identified the image containing the target category. The duration images were shown was individually calibrated to achieve correct detection of around 85%.

Critically, the residuals from this model can be expected to contain information regarding *differences* in activation across the four different target conditions. For each trial, the corresponding GLM residuals (0 to 22 s from trial onset) were extracted using trial-by-trial surface searchlight classification.

Defining standardized surface searchlights

MVPA using a 'searchlight' approach can detect information encoded in local spatial activation patterns in a sensitive and unbiased manner (Haynes et al., 2007; Kriegeskorte et al., 2006; Soon et al., 2008). Using searchlights defined along the cortical surface further improves the anatomical specificity of the resultant information map compared to

volume-based searchlights (Chen et al., 2011). Here, we developed a novel standardized cortical searchlight approach, whereby each searchlight captured a patch of cortex anatomically matched across participants (Fig. 2). This allowed us to directly assess whether the activation patterns within the searchlight captured functionally relevant information consistently across different individuals (even though the actual patterns may differ across individuals). This highly sensitive method allowed us to identify cortical regions that encoded the target category *before* and *after* visual stimulation. The details of this approach are described below.

For each participant, the cortical surface (gray-white boundary) was reconstructed from the high resolution T1 image. The mesh of each hemisphere was then inflated into a sphere of 100 mm radius and normalized to a standard template ('fsaverage') based on cortical

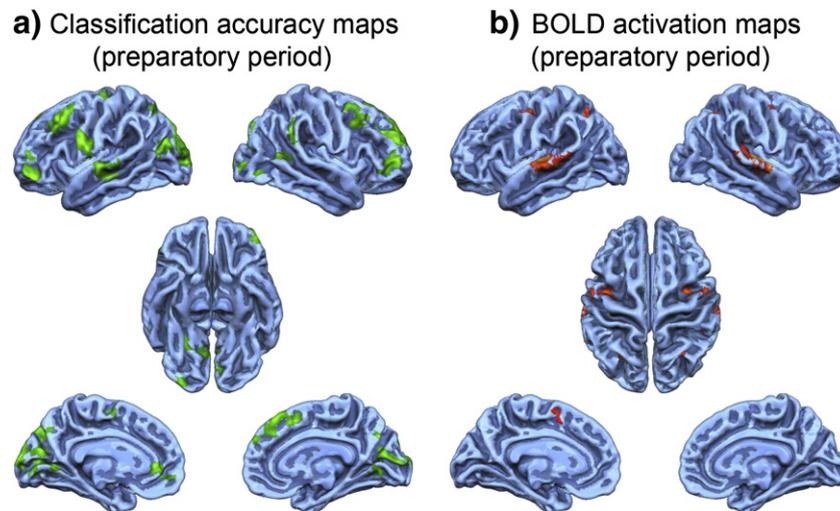


Fig. 2. Cortical activity during the preparatory period, prior to visual stimulation. a) Centers of searchlights encoding the target category. Once the target category was specified, it could be decoded with above chance accuracy from local activation patterns in various regions ($p < .05$, Bonferroni corrected; 25 mm² cluster threshold). b) Cortical regions showing changes in BOLD signal magnitude during the preparatory period, averaged across all categories ($p < .001$, uncorrected; 25 mm² cluster threshold).

folding patterns (FreeSurfer, <http://surfer.nmr.mgh.harvard.edu>). Next, the functional voxels corresponding to the gray matter surface were identified as follows: a surface in the middle of the gray matter (“graymid”) was created by inflating the gray-white boundary surface by 50% towards the pial surface (Chen et al., 2011). This graymid surface was then coregistered to the first image of the first functional run. For each vertex on the graymid surface, the corresponding EPI voxel containing the vertex was indexed. FreeSurfer maintains a one-to-one correspondence of vertices when morphing the original gray-white boundary mesh into the graymid and normalized spherical meshes. This facilitates the identification of the corresponding graymid vertex and EPI voxel for a given vertex on the normalized spherical mesh.

There are two typical approaches to normalize individual participant data to a standard template space for group-level comparisons: 1) normalizing the original EPI images *before* performing MVPA, and 2) normalizing the resultant accuracy maps *after* performing MVPA. The problem with the first approach is that interpolation during image reslicing may reduce information-rich differences between neighboring voxels before MVPA. The second approach avoids this problem. However, while classification accuracy is usually assigned to the central voxel or vertex of the searchlight (but see Bjornsdotter et al., 2011), the information is actually encoded within the spatial activation patterns of the *whole* searchlight, not just the central voxel or vertex. Given individual differences in anatomy, the classification accuracy in a given voxel or vertex may not arise from the identical patch of cortex across different participants if the normalization were to be performed after MVPA.

To circumvent these issues, we developed a new approach utilizing standardized searchlights that captured anatomically matched cortical patches across different individuals, without having to spatially normalize the functional data itself. We first selected a reference structural dataset, ‘fsaverage6’ (<ftp://surfer.nmr.mgh.harvard.edu/pub/data/fsaverages.tgz>), which is a downsampled version of ‘fsaverage’, with ~40 K vertices instead of ~160 K vertices per hemisphere. This was more than sufficient given that there are ~15 K gray matter voxels per hemisphere at the EPI image resolution used here. For each vertex on the normalized spherical mesh of this reference dataset, a circular searchlight centered on the vertex was defined (geodesic radius 20 mm). Each searchlight was then ‘projected’ onto every participant’s normalized spherical mesh, using a nearest-neighbor approach. This analysis was performed separately for each hemisphere. Since the spherical meshes of all participants and the reference dataset were normalized to the same template for each hemisphere, each searchlight effectively captured anatomically matched cortical patches across participants. The EPI voxels that corresponded to the vertices within the searchlight were then identified as described above, and the functional data extracted for MVPA.

The meshes of each hemisphere contained white matter vertices due to the cut through the medial wall connecting both hemispheres, and searchlights centered on such vertices were excluded, leaving ~35 k searchlights for each hemisphere.

Surface searchlight MVPA

For each searchlight, support vector machines (SVM; LIBSVM implementation, <http://www.csie.ntu.edu.tw/~cjlin/libsvm>) were used to assess whether the target category was encoded in the spatiotemporal pattern of activation during the preparatory period. First, spatiotemporal vectors were created for each correct trial by concatenating the GLM residuals from all voxels within the searchlight, and the 4 timepoints from 0 to 8 s after cue onset (Mourao-Miranda et al., 2007). The visual stimulus was presented at 6 s, but given the delay in the hemodynamic response, visual-evoked BOLD changes were not expected until after 8 s. To avoid biasing against any given condition, an identical number of trials were selected for each of the 4 target category conditions (‘House’, ‘Face’, ‘Car’ and ‘Bird’) by excluding the last few trials from conditions with more correct trials. Next, SVM classifiers were trained

to distinguish the spatiotemporal patterns of activation between pairs of conditions (‘House’ vs ‘Face’, ‘House’ vs ‘Car’, ‘House’ vs ‘Bird’, etc.; chance level 50%). Trials from each condition were divided into 5 separate sets. Four trial sets were combined to train an SVM model, which was then tested using the independent 5th trial set. This training and testing cycle was repeated 5 times (5-fold cross-validation), such that each trial was tested once. The classification accuracy averaged across all 6 pair-wise classifications and 5-fold cross validation (~360 classification tests for each participant) was then assigned to the central vertex of the searchlight on the downsampled reference mesh. For each participant, this process was repeated independently for each searchlight to create a map of classification accuracies for the whole cortex.

The accuracy map for each participant then underwent heat kernel smoothing (Chung et al., 2005). Next, a stringent approach was adopted to evaluate whether the classification accuracies for a searchlight were significant at the group level. For each participant and searchlight, considering each of 360 classification tests as a Bernoulli trial with 50% probability of success, a mean accuracy value $\geq 55\%$ would be significant (Pereira et al., 2009). Hence, for each searchlight, we performed a group-level t-test of classification accuracy against 55% (instead of 50%). A significant result ($p < .05$, Bonferroni corrected) would imply that the spatiotemporal activation patterns within a searchlight contained information about the target category. We postulated that category-specific biases in activation would be evidenced in higher order visual areas in the preparatory period prior to target exposure.

To establish the behavioral relevance of such preparatory biases, we correlated the mean searchlight classification accuracy (averaged across all 6 pair-wise classifications) with the stimulus exposure duration for that individual. We expected a functionally relevant region to show a significant negative correlation. To correct for multiple comparisons, permutation testing (1000 Monte Carlo simulations) implemented in BrainVoyager QX version 2.12 (Brain Innovation) was used to estimate the cluster-size threshold corresponding to a global error probability of $p < .05$ (single-vertex threshold: $r < -0.64$, $p < .005$, one-tailed) (Forman et al., 1995; Hayasaka and Nichols, 2003; Nichols and Holmes, 2002).

Additional classification analyses were conducted to locate areas that encoded the target category *after* the natural scene quartets were shown (4 timepoints from 10 to 16 s after cue onset). Control analyses were also conducted to ensure that our classification procedures were not prone to false positives (see Supplementary Fig. 4).

Results

At the criterion level of target detection accuracy (mean = 84.8%, S.E. = 0.01%, Supplementary Table 1), exposure duration ranged from 53 to 182 ms across the 15 participants whose imaging data contributed to the analyses (mean = 105 ms, median = 95 ms). This provided the critical measure of perceptual speed.

Following the auditory word cue, patterns in several cortical regions became informative of the cued target category (Fig. 2 and Table 1; Supplementary Tables 2 to 4 and Supplementary Movie 1 contain results for stimulus-evoked periods and univariate GLM). These included the left superior temporal gyrus and supramarginal gyrus, reflecting perceptual processing of the auditory cue. Medial and lateral frontal patterns could be involved in establishing task goals and voluntary control of attention. Category-specific biases in visual cortices were also evident in early visual cortex as well as higher visual cortex — parahippocampal gyrus and the lateral occipital complex (LOC).

In accordance with the hypothesis that preparatory biases facilitate target detection, we predicted that individuals with stronger biases would require shorter stimulus exposure time to detect targets. The strength of preparatory biases was indexed by classification accuracies averaged across all four categories.

The only cortical region that showed such a correlation was an anterior lateral portion of the right LOC, stretching from occipito-temporal

sulcus to inferior temporal sulcus (MNI coordinates: 55, –56, –4; peak $r = -0.91$; cluster-size corrected $p < .005$; Figs. 3a and b). No significant positive correlations were found.

To ensure that this finding was not spurious (Vul et al., 2009), we repeated the whole classification analyses four times with different randomization of trials for classifier training and testing (see Supplementary Fig. 2). In all analyses the right LOC area was the only cortical region showing significant negative correlations, with peak values of $r < -0.85$.

Similar correlation analyses were performed for the classification accuracies of each individual category (averaged across 3 pair-wise classifications involving the specific category) and category-pair. Significant negative correlations with calibrated exposure duration were seen in right anterior LOC for 'House' (peak $r = -0.84$; cluster-size corrected $p < .05$), 'Face' (peak $r = -0.89$; cluster-size corrected $p < .05$), and 'Bird' (peak $r = -0.83$; cluster-size corrected $p < .01$; Fig. 3c). A cluster of vertices within right anterior LOC also showed strong negative correlations for 'Car' (peak $r = -0.85$), but did not survive cluster-size correction ($p > .05$). Similarly, although there were small clusters in right anterior LOC with negative correlations ($r < -0.64$) for all category-pairs except 'House' vs 'Car', none survived cluster-size correction (Supplementary Fig. 3). No other significant regions were found across all correlation analyses.

Next, we determined whether category selectivity would be preserved between preparatory and stimulus-evoked periods at the voxel level. From a separate GLM (similar to that described above, but with a separate set of FIR predictors for each category), parameter estimates

were obtained for the preparatory period (averaged across 3 timepoints covering 2 to 8 s after auditory cue onset) and stimulus-evoked period (averaged across 3 timepoints covering 2 to 8 s after visual stimulus onset). For each category the 10 most positively selective (i.e., higher BOLD signal than the mean of other categories) and 10 most negatively selective voxels (i.e., lower BOLD signal than the mean of other categories) during the preparatory period were identified from the right anterior LOC for each participant. The mean category selectivity of these voxels was preserved in the stimulus-evoked period (positively selective voxels: $t_{14} = 7.10$, $p < 10^{-5}$; negatively selective voxels: $t_{14} = -7.90$, $p < 10^{-6}$, one-tailed tests). The results from a representative participant are shown in Fig. 4.

Discussion

We observed category discriminant BOLD signal patterns in multiple cortical areas as participants prepared to locate targets belonging to four familiar object categories. While these baseline patterns were informative regarding the target category, only biases within the anterior LOC predicted how much exposure time one needed to identify the target object category. Informative voxels in LOC preparatory patterns, both positively and negatively selective for a target category, continued to show the same category selectivity when viewing target objects.

Functional relevance of preparatory activity

The magnitude of preparatory shifts in BOLD signal in visual cortex has been shown to predict behavioral performance in several studies (Giesbrecht et al., 2006; Peelen and Kastner, 2011; Puri et al., 2009; Ress et al., 2000). However, this association is not obligatory (Fannon, 2008). The loci and conditions under which preparatory activity are predictive of behavior remain to be clarified. For example, elevated activity can reflect non-target related information such as temporal expectation (Esterman and Yantis, 2010), task difficulty (Ress et al., 2000) or spatial location (McMains et al., 2007) that may not materially contribute to behavioral performance.

Here, whole brain searchlights were used to evaluate preparatory patterns prior to focusing on specific visual cortical regions (Peelen and Kastner, 2011; Peelen et al., 2009; Reddy et al., 2010). Additionally, using four different categories allowed us to verify that our results were representative of familiar object categories in general (Fig. 3c and Supplementary Fig. 3), as opposed to being driven by a single category – more likely when only comparing a pair of categories. Of the various correlation analyses conducted, the mean accuracy across all 6 pair-wise classifications was most predictive of the minimal exposure duration required for correct target identification. While the correlation results of individual categories and category-pairs were not completely equivalent, none of them was more robust than the mean accuracy results, both in terms of peak correlation and spatial extent. In other words, the mean accuracy correlation results could not be fully explained by any single category or category-pair. These findings are indicative that preparatory activity in anterior LOC can facilitate detection for various familiar categories.

Discriminatory patterns and their relevance to perceptual speed

Within ventral visual areas, category-specific biases that differentiated the target category from three other categories but did not correlate with perceptual speed were detected in early visual areas, posterior portions of LOC, and parahippocampal gyrus.

Early visual cortex has been implicated in the visualization of simple and complex objects (Cui et al., 2007; Kosslyn et al., 1995; Kosslyn et al., 1999). Baseline shifts have even been correlated with the subjective vividness of mental images (Cui et al., 2007). However, other studies have not shown baseline changes in early visual cortex signal (D'Esposito et al., 1997; Formisano et al., 2002). Critically, while valid

Table 1
Cortical regions showing category-specific activity during the preparatory period ($p < .05$, Bonferroni corrected; 25 mm² cluster threshold on cortical surface).

Cortical region	Hemisphere	BA	MNI			Classification Accuracy (%)
			X	Y	Z	
Visual						
Calcarine	L	17	-5	-81	-3	63.1
Calcarine	R	17	8	-71	20	63.2
Lingual gyrus	R	17	9	-61	-2	63.4
Middle occipital gyrus	R	17	30	-87	0	63.6
Superior occipital gyrus	L	18	-29	-93	14	65.1
Superior occipital gyrus	L	19	-22	-83	16	63.6
Superior occipital gyrus	R	19	25	-81	36	63.3
Lateral occipital sulcus	L	19	-39	-79	11	64.2
Lateral parieto-occipital sulcus	L	19	-12	-77	47	63.3
Inferior occipital gyrus	R	18	27	-89	-12	61.6
Collateral sulcus	R	19	23	-52	-9	63.2
Inferior occipital gyrus	R	18	44	-77	-6	62.3
Posterior middle temporal gyrus	R	19	58	-57	7	62.9
Frontal						
Precentral gyrus	L	6	-47	5	16	63.4
Precentral gyrus	L	4	-35	-22	65	61.6
Middle frontal gyrus	L	9	-40	17	32	63.8
Middle frontal gyrus	R	9	40	15	54	63.6
Lateral frontopolar	L	45/46	-30	49	6	64.1
Lateral frontopolar	R	45/46	53	33	-4	64.3
Superior frontal gyrus	R	10	23	55	24	65.6
Medial superior frontal gyrus	L	6	-5	-3	52	61.5
Medial superior frontal gyrus	R	6/8	6	22	52	64.8
Rostromedial prefrontal cortex	L	11	-12	45	-7	63.3
Parietal						
Intraparietal sulcus	L	7	-18	-67	45	63.8
Precuneus	R	31	8	-70	34	62.0
Auditory						
Superior temporal gyrus	L	22	-63	-12	-2	64.2
Supramarginal gyrus	R	39	63	-41	24	63.1

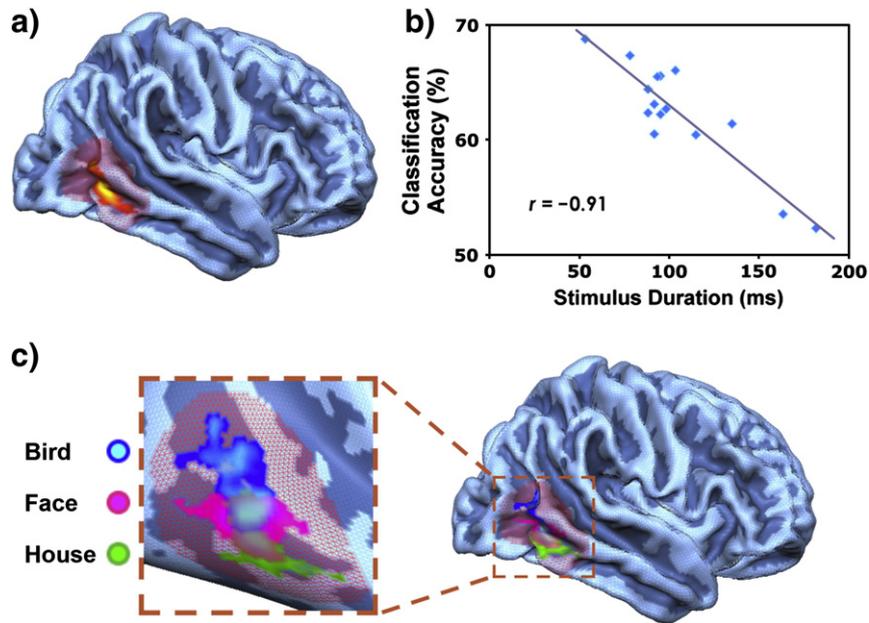


Fig. 3. Significant correlation between preparatory category-specific biases and behavioral performance. a) Cortical map depicting searchlights in which classification accuracy (averaged across all categories) showed significant correlation with individually calibrated stimulus duration (yellow: center of searchlights with Pearson's $r < -0.64$; cluster-size corrected $p < .05$; pink mesh: full extent of significant searchlights). b) Correlation of classification accuracy (averaged across all categories) and stimulus duration, from peak searchlight shown in Fig. 3a. Participants showing stronger category-specific preparatory biases in right anterior LOC required shorter exposure to a stimulus to locate the target category. c) When the classification accuracies were considered separately for each of the 4 categories, significant correlations with stimulus duration could still be found in right anterior LOC for 'House', 'Face' and 'Bird' ($r < -0.64$; cluster-size corrected $p < .05$; a suprathreshold cluster for 'Car' did not survive cluster size correction). The cortical surface has been inflated in the inset to more clearly illustrate the overlap between the different categories (pink mesh: full extent of significant searchlights from Fig. 3a).

expectations may facilitate the detection of complex objects, invalid expectations could hinder performance (Puri and Wojciulik, 2008). As such, expectation related neural activity related to the visualization of exemplars in early visual cortex (Kosslyn et al., 1999) may hinder speeded detection of target objects specified only by category (Peelen and Kastner, 2011). This could be compounded if objects are embedded in unspecified scenes (Peelen and Kastner, 2011). Hence, changes in early visual cortex neural activity may contribute to the expectation of a visual target but not correlate with recognition speed, perhaps reflecting the occasional use of visualization to focus attention on the target category.

While the parahippocampal area is best known for sensitivity to houses and place scenes, it also appears to provide contextual information regarding the how an object relates to its spatial context (Bar et al., 2008). It follows that this area may also generate informative patterns during scene-based target searches (Peelen and Kastner, 2011; Peelen et al., 2009).

Of the higher visual areas, only discriminatory voxels in the anterior LOC predicted perceptual speed. It would not be advantageous for participants to visualize target objects as only a category label was provided and no object was presented more than once (Peelen and

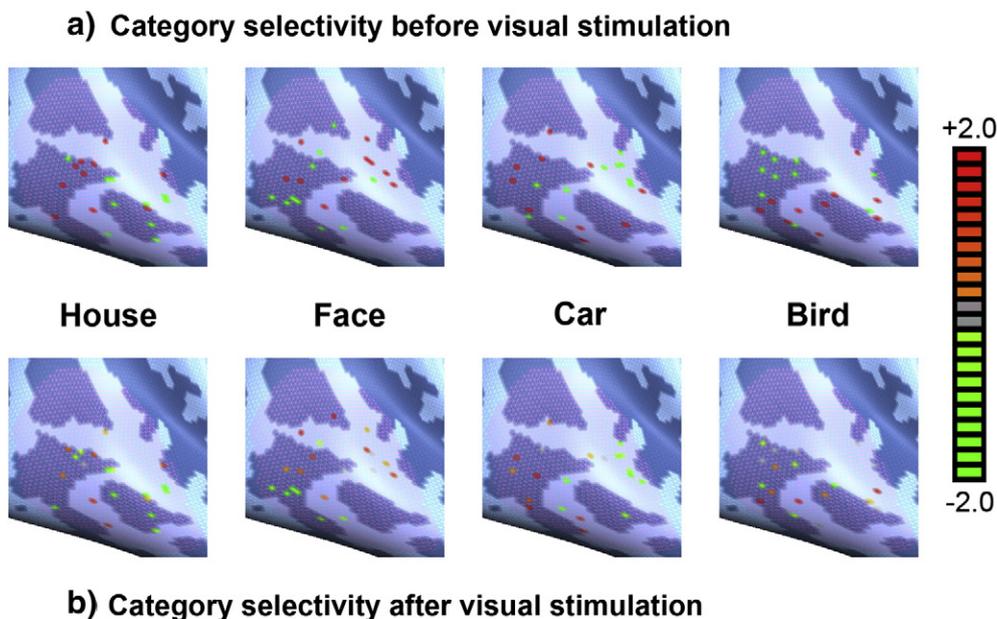


Fig. 4. Voxel-level category selectivity for a representative participant. a) For each category, the 10 most positively selective (i.e., greater magnitude than other categories) and 10 most negatively selective (i.e., lower magnitude than other categories) voxels in the right anterior LOC during the preparatory period are shown on an inflated cortical surface. b) Most of the voxels from 4a continued to show the same direction of category selectivity after visual stimulation.

Kastner, 2011). Participants also did not know which quadrant to expect the target. Attention to other low-level features like orientation and color would not have been helpful as these were varied across different pictures. Collectively, these points suggest that the preparatory LOC patterns uncovered here denote visual characteristics that generalize across members of a category.

Recently, preparatory activity in the anterior LOC was found to correlate with accuracy and response time when participants had to determine whether an image contained an object from a target category (Peelen and Kastner, 2011). Here, we add to this finding by demonstrating that more distinct preparatory patterns in anterior LOC reduced the *exposure duration* needed for target detection. Hence, anterior LOC preparatory activity can modulate perceptual speed for visual search both in terms of the minimum stimulus exposure time and overall perceptual processing time required.

Stimulus presentation times affect the quantity of sensory information available to the earliest levels of visual representation (Vanrullen, 2011). Thus, our findings suggest that preparatory activity in the anterior LOC may modulate the sensory information needed for successful target categorization. For example, being able to anticipate salient visual characteristics specific to one category may hasten target detection. This may partly account for the findings in Peelen and Kastner's (2011) study in which stimulus presentation times were equated across participants (100 ms).

It is remarkable that searching for a target category from four complex natural scenes compared to one central scene (in Peelen and Kastner, 2011), across wider visual angles (24° vs 10°), did not lead to poorer detection accuracy (85% vs 82%) when mean stimulus exposure times were comparable (105 ms vs 100 ms). This reinforces the idea that category-specific biasing mechanisms in object-selective cortex can operate in parallel across the visual field (Peelen et al., 2009), even during the preparatory phase. However, there did appear to be additional processing costs when comparing the response times (~1000 ms vs ~700 ms).

The occipito-temporal area identified here (which extends to a posterior middle temporal region) has been shown to participate in visual object categorization (Moore and Price, 1999). The left hemisphere homologue of this area is involved in semantic processing involving concrete visual objects (Whatmough et al., 2002). It also shows greater activation when one is evaluating semantic associations between drawings as opposed to semantic associations related to word stimuli (Chee et al., 2000).

Preparatory activity in frontal control regions

Although not predictive of perceptual performance, category-specific information could also be decoded from several medial and lateral frontal control regions. These likely serve as sources of top-down control signals that bias visual cortex (Peelen and Kastner, 2011).

Different frontal regions may exercise different control functions. The frontal region most directly relevant to object processing is likely to be rostromedial prefrontal cortex. Category-specific cue effects for scene discrimination have recently been observed here (Peelen and Kastner, 2011). This region may serve to maintain updated representations of potential scene contexts that in turn modulate object processing in visual cortex (Bar, 2004). Medial and lateral rostral prefrontal cortex are involved in maintaining specific intentions over delays (Burgess et al., 2003; Gilbert, 2011; Haynes et al., 2007) and may use such information to guide future responses (Karnath et al., 1991; Lepage and Richer, 2000).

Correspondence between preparatory and stimulus evoked activity

Concurrent enhancement of target-related preparatory activity and suppression of distractor-related preparatory activity in visual cortex occurs when respective categories activate spatially separable

areas (Puri et al., 2009). We found that voxels that were most selective for the target category during the preparatory period continued to show the same category selectivity after the visual stimulus. Correspondingly, voxels that were negatively selective for the target category, i.e., lower activation magnitude for the target than non-targets, remained negatively biased. Notably, this selectivity was evident even when exemplars of all four categories were presented concurrently within natural scenes, and participants had no prior knowledge of where the target would appear. As such, the present results generalize prior findings concerning concurrent enhancement and suppression of target and non-target neural activity to instances where category selectivity is evident at the voxel level – a finer spatial scale.

Conclusion

We found that preparatory patterns in the anterior LOC that distinguish target objects embedded in natural scenes from non-target objects predict perceptual performance. Patterns in the preparatory and stimulus-evoked phases within this cortical region reflecting increased as well as decreased signal changes relative to baseline were strikingly congruent. These signals likely correspond to visual semantic information related to the object category and its likely surroundings.

Acknowledgments

We would like to thank Vanessa Chen Wen-Xi and Koh Shuwei for contributing to stimulus preparation. Vanessa Chen Wen-Xi, Koh Shuwei, Annette Chen contributed to collecting fMRI data. This work was supported by grants awarded to Dr Michael Chee from the Defense Science and Technology Agency Singapore (POD0713897) and National Medical Research Council Singapore (StaR/0004/2008).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.neuroimage.2012.10.036>.

References

- Bar, M., 2004. Visual objects in context. *Nat. Rev. Neurosci.* 5, 617–629.
- Bar, M., Aminoff, E., Schacter, D.L., 2008. Scenes unseen: the parahippocampal cortex intrinsically subserves contextual associations, not scenes or places per se. *J. Neurosci.* 28, 8539–8544.
- Bjornsdotter, M., Rylander, K., Wessberg, J., 2011. A Monte Carlo method for locally multivariate brain mapping. *NeuroImage* 56, 508–516.
- Bode, S., Haynes, J.D., 2009. Decoding sequential stages of task preparation in the human brain. *NeuroImage* 45, 606–613.
- Bode, S., Bogler, C., Soon, C.S., Haynes, J.D., 2012. The neural encoding of guesses in the human brain. *NeuroImage* 59, 1924–1931.
- Burgess, P.W., Scott, S.K., Frith, C.D., 2003. The role of the rostral frontal cortex (area 10) in prospective memory: a lateral versus medial dissociation. *Neuropsychologia* 41, 906–918.
- Castelhano, M.S., Heaven, C., 2010. The relative contribution of scene context and target features to visual search in scenes. *Atten. Percept. Psychophys.* 72, 1283–1297.
- Chee, M.W., Weekes, B., Lee, K.M., Soon, C.S., Schreiber, A., Hoon, J.J., Chee, M.W., 2000. Overlap and dissociation of semantic processing of Chinese characters, English words, and pictures: evidence from fMRI. *NeuroImage* 12, 392–403.
- Chelazzi, L., Duncan, J., Miller, E.K., Desimone, R., 1998. Responses of neurons in inferior temporal cortex during memory-guided visual search. *J. Neurophysiol.* 80, 2918–2940.
- Chen, Y., Namburi, P., Elliott, L.T., Heinzle, J., Soon, C.S., Chee, M.W., Haynes, J.D., 2011. Cortical surface-based searchlight decoding. *NeuroImage* 56, 582–592.
- Chung, M.K., Robbins, S.M., Dalton, K.M., Davidson, R.J., Alexander, A.L., Evans, A.C., 2005. Cortical thickness analysis in autism with heat kernel smoothing. *NeuroImage* 25, 1256–1265.
- Cichy, R.M., Heinzle, J., Haynes, J.D., 2012. Imagery and perception share cortical representations of content and location. *Cereb. Cortex* 22, 372–380.
- Cui, X., Jeter, C.B., Yang, D., Montague, P.R., Eagleman, D.M., 2007. Vividness of mental imagery: individual variability can be measured objectively. *Vision Res.* 47, 474–478.
- D'Esposito, M., Detre, J.A., Aguirre, G.K., Stallcup, M., Alsop, D.C., Tippet, L.J., Farah, M.J., 1997. A functional MRI study of mental image generation. *Neuropsychologia* 35, 725–730.

- Driver, J., Frith, C., 2000. Shifting baselines in attention research. *Nat. Rev. Neurosci.* 1, 147–148.
- Esterman, M., Yantis, S., 2010. Perceptual expectation evokes category-selective cortical activity. *Cereb. Cortex* 20, 1245–1253.
- Fannon, S.P., 2008. Baseline shifts do not predict attentional modulation of target processing during feature-based visual attention. *Front. Hum. Neurosci.* 1.
- Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintun, M.A., Noll, D.C., 1995. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn. Reson. Med.* 33, 636–647.
- Formisano, E., Linden, D.E., Di Salle, F., Trojano, L., Esposito, F., Sack, A.T., Grossi, D., Zanella, F.E., Goebel, R., 2002. Tracking the mind's image in the brain I: time-resolved fMRI during visuospatial mental imagery. *Neuron* 35, 185–194.
- Giesbrecht, B., Woldorff, M.G., Song, A.W., Mangun, G.R., 2003. Neural mechanisms of top-down control during spatial and feature attention. *NeuroImage* 19, 496–512.
- Giesbrecht, B., Weissman, D.H., Woldorff, M.G., Mangun, G.R., 2006. Pre-target activity in visual cortex predicts behavioral performance on spatial and feature attention tasks. *Brain Res.* 1080, 63–72.
- Gilbert, S.J., 2011. Decoding the content of delayed intentions. *J. Neurosci.* 31, 2888–2894.
- Hayasaka, S., Nichols, T.E., 2003. Validating cluster size inference: random field and permutation methods. *NeuroImage* 20, 2343–2356.
- Haynes, J.-D., Rees, G., 2006. Decoding mental states from brain activity in humans. *Nat. Rev. Neurosci.* 7, 523–534.
- Haynes, J.-D., Sakai, K., Rees, G., Gilbert, S., Frith, C., Passingham, R.E., 2007. Reading hidden intentions in the human brain. *Curr. Biol.* 17, 323–328.
- Hopfinger, J.B., Buonocore, M.H., Mangun, G.R., 2000. The neural mechanisms of top-down attentional control. *Nature* 3, 284–291.
- Intraub, H., 1981. Rapid conceptual identification of sequentially presented pictures. *J. Exp. Psychol. Hum. Percept. Perform.* 7, 604.
- Karnath, H.O., Wallech, C.W., Zimmermann, P., 1991. Mental planning and anticipatory processes with acute and chronic frontal lobe lesions: a comparison of maze performance in routine and non-routine situations. *Neuropsychologia* 29, 271–290.
- Kastner, S., Pinsk, M.A., De Weerd, P., Desimone, R., Ungerleider, L.G., 1999. Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron* 22, 751–761.
- Kosslyn, S.M., Thompson, W.L., Kim, I.J., Alpert, N.M., 1995. Topographical representations of mental images in primary visual cortex. *Nature* 378, 496–498.
- Kosslyn, S.M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J.P., Thompson, W.L., Ganis, G., Sukel, K.E., Alpert, N.M., 1999. The role of area 17 in visual imagery: convergent evidence from PET and rTMS. *Science* 284, 167–170.
- Kriegeskorte, N., Goebel, R., Bandettini, P., 2006. Information-based functional brain mapping. *Proc. Natl. Acad. Sci. U. S. A.* 103, 3863–3868.
- Lepage, M., Richer, F., 2000. Frontal brain lesions affect the use of advance information during response planning. *Behav. Neurosci.* 114, 1034–1040.
- Li, F.F., VanRullen, R., Koch, C., Perona, P., 2002. Rapid natural scene categorization in the near absence of attention. *Proc. Natl. Acad. Sci. U. S. A.* 99, 9596–9601.
- Luck, S.J., Chelazzi, L., Hillyard, S.A., Desimone, R., 1997. Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J. Neurophysiol.* 77, 24–42.
- McMains, S.A., Fehd, H.M., Emmanouil, T.A., Kastner, S., 2007. Mechanisms of feature- and space-based attention: response modulation and baseline increases. *J. Neurophysiol.* 98, 2110–2121.
- Moore, C.J., Price, C.J., 1999. A functional neuroimaging study of the variables that generate category-specific object processing differences. *Brain* 122 (Pt 5), 943–962.
- Mourao-Miranda, J., Friston, K.J., Brammer, M., 2007. Dynamic discrimination analysis: a spatial-temporal SVM. *NeuroImage* 36, 88–99.
- Nichols, T.E., Holmes, A.P., 2002. Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum. Brain Mapp.* 15, 1–25.
- Norman, K.A., Polyn, S.M., Detre, G.J., Haxby, J.V., 2006. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn. Sci.* 10, 424–430.
- Peelen, M.V., Kastner, S., 2011. A neural basis for real-world visual search in human occipitotemporal cortex. *Proc. Natl. Acad. Sci. U. S. A.* 108, 12125–12130.
- Peelen, M.V., Fei-Fei, L., Kastner, S., 2009. Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature* 460, 94–97.
- Pereira, F., Mitchell, T., Botvinick, M., 2009. Machine learning classifiers and fMRI: a tutorial overview. *NeuroImage* 45, S199–S209.
- Potter, M.C., Faulconer, B.A., 1975. Time to understand pictures and words. *Nature* 253, 437–438.
- Potter, M.C., Levy, E.I., 1969. Recognition memory for a rapid sequence of pictures. *J. Exp. Psychol.* 81, 10–15.
- Puri, A.M., Wojciulik, E., 2008. Expectation both helps and hinders object perception. *Vision Res.* 48, 589–597.
- Puri, A.M., Wojciulik, E., Ranganath, C., 2009. Category expectation modulates baseline and stimulus-evoked activity in human inferotemporal cortex. *Brain Res.* 1301, 89–99.
- Reddy, L., Tsuchiya, N., Serre, T., 2010. Reading the mind's eye: Decoding category information during mental imagery. *NeuroImage* 50, 818–825.
- Ress, D., Backus, B.T., Heeger, D.J., 2000. Activity in primary visual cortex predicts performance in a visual detection task. *Nat. Neurosci.* 3, 940–945.
- Soon, C.S., Brass, M., Haynes, J.-D., 2008. Unconscious determinants of free decisions in the human brain. *Nat. Neurosci.* 11, 543–545.
- Stokes, M., Thompson, R., Nobre, A.C., Duncan, J., 2009. Shape-specific preparatory activity mediates attention to targets in human visual cortex. *Proc. Natl. Acad. Sci. U. S. A.* 106, 19569–19574.
- Sylvester, C.M., Shulman, G.L., Jack, A.I., Corbetta, M., 2009. Anticipatory and stimulus-evoked blood oxygenation level-dependent modulations related to spatial attention reflect a common additive signal. *J. Neurosci.* 29, 10671–10682.
- Thorpe, S., Fize, D., Marlot, C., 1996. Speed of processing in the human visual system. *Nature* 381, 520–522.
- Vanrullen, R., 2011. Four common conceptual fallacies in mapping the time course of recognition. *Front. Psychol.* 2, 365.
- Vul, E., Hanus, D., Kanwisher, N., 2009. Attention as inference: selection is probabilistic; responses are all-or-none samples. *J. Exp. Psychol. Gen.* 138, 546–560.
- Whitmough, C., Chertkow, H., Murtha, S., Hanratty, K., 2002. Dissociable brain regions process object meaning and object structure during picture naming. *Neuropsychologia* 40, 174–186.
- Wolfe, J.M., Alvarez, G.A., Rosenholtz, R., Kuzmova, Y.I., Sherman, A.M., 2011. Visual search for arbitrary objects in real scenes. *Atten. Percept. Psychophys.* 73, 1650–1671.