

A Bimodal Tuning Curve for Spatial Frequency Across Left and Right Human Orbital Frontal Cortex During Object Recognition

Anat R. Fintzi¹ and Bradford Z. Mahon^{1,2,3}

¹Department of Brain and Cognitive Sciences, University of Rochester, Rochester NY 14627, USA, ²Department of Neurosurgery, University of Rochester, 601 Elmwood Avenue, Rochester NY 14627, USA and ³Center for Visual Science, University of Rochester, 252, Elmwood Avenue, Rochester NY 14627, USA

Address correspondence to Bradford Z. Mahon, Meliora Hall, University of Rochester, Rochester, NY 14627-0268, USA. Email: mahon@rcbi.rochester.edu

Orbital frontal cortex (OFC) is known to play a role in object recognition by generating “first-pass” hypotheses about the identity of naturalistic images based on low spatial frequency (SF) information. These hypotheses are evaluated by more detailed (and slower) ventral visual pathway processes. While it has been suggested on theoretical grounds, it remains unknown whether OFC also receives postrecognition feedback about stimulus identity. We used a novel paradigm in the context of functional magnetic resonance imaging that permits the first few hundred milliseconds of object recognition to be spread out over 120 s. OFC shows a robust response to low and relatively high SFs, whereas ventral stream regions display unimodal response distributions shifted toward high SFs. These findings in OFC were modulated by hemisphere, with right OFC differentially responding to low SFs and left OFC differentially responding to high SFs. Psychophysical experiments confirmed that the same ranges of SFs preferred by ventral stream regions are critical for determining the accuracy and speed of object recognition. Our findings indicate that OFC accesses global form (low SF information, right OFC) and object identity (high SF information, left OFC), and suggest that OFC receives feedback about the accuracy of its initial hypothesis regarding stimulus identity.

Keywords: fMRI, object recognition, psychophysics, spatial frequency, spatial vision

Introduction

It is a remarkable feat of the primate visual system that it is capable of visually identifying an indefinite number of objects within a few hundreds of milliseconds (Potter 1976; Tanaka and Curran 2001; VanRullen and Thorpe 2001; Rossion et al. 2003; Hauk et al. 2006). An important and unresolved issue concerns the neurocognitive organization of high-level visual information that allows for such fast and efficient visual recognition to occur. One mechanism that has been proposed that would aid in the speed and efficiency of high-level visual recognition is a system for generating predictions about the identity of an object on the basis of coarse and global visual information (Schyns and Oliva 1994; Bullier 2001; Bar 2003; Fenske et al. 2006; Kveraga, Boshyan et al. 2007; Kveraga, Ghuman, et al. 2007; Peyrin et al. 2010; see Riesenhuber and Poggio 2002 for an alternative). Within that framework, fast magnocellular systems (Maunsell et al. 1999) extract coarse-grained (i.e., low spatial frequency, LSF) information (Derrington and Lennie 1984; Tootell et al. 1988) about global form (Livingstone and Hubel 1988; Lamb and Yund 1993; Schyns and Oliva 1994, 1999; Hughes et al. 1996; Parker et al. 1996; Olds and Engel 1998) that constrains the range of

possible identities of a stimulus, with the stimulus being decisively identified on the basis of mid-to-high spatial frequencies (SFs), presumably mediated by parvocellular systems and localized to ventral temporal-occipital regions.

Orbital frontal cortex (OFC) has been hypothesized to be a central hub in this process, receiving LSF information about global form via magnocellular pathways and passing it back to the ventral object-processing stream (Bar 2003; Bar et al. 2006; Fenske et al. 2006; Peyrin et al. 2010). At a theoretical level, it has been proposed that OFC both mediates the transmission of global shape information, and also receives feedback on the outcome of the identification process, perhaps in the form of error (Bar 2003). Independently of the particular neuroanatomical details, this general architecture in which a process or region (e.g., OFC) relays an initial “first guess” of the object and then receives feedback on the outcome of recognition, is an important component of top-down models of object recognition and statistical learning (Behrmann et al. 1998; Bullier 2001; Foxe and Simpson 2002; Bar 2003; Juan and Walsh 2003; Baker et al. 2004; Bar et al. 2006; Fenske et al. 2006; Yuille and Kersten 2006; Kveraga, Boshyan et al. 2007; Kveraga, Ghuman, et al. 2007; Peyrin et al. 2010).

The hypothesis that OFC computes a first-pass global analysis on the basis of LSF information and also receives feedback about object identity (mediated by HSF information) generates the prediction that OFC should exhibit a bimodal peak in neuronal activity during the course of object recognition (Bar 2003). The first peak would correspond to the analysis of global shape, and the second would be contingent on feedback from ventral stream structures identifying the object and would thus correspond to information about object identity. Although a range of empirical findings support the role of OFC (Bar et al. 2006; Kveraga, Boshyan et al. 2007; Peyrin et al. 2010) in computing magnocellular-based information about global form, the critical hypothesis that OFC exhibits a bimodal distribution in response properties corresponding to initial and final states of object recognition remains untested. Thus, it is unclear whether OFC’s involvement in object recognition is limited to sending information about global form to ventral stream structures, or whether it also receives input about the outcome of object identification processes.

Because the blood-oxygen-level-dependent (BOLD) response is too sluggish to resolve a bimodal distribution in neuronal activity, we developed a new method in which the first few hundred milliseconds are “drawn out” over the course of 120 s. This is accomplished by presenting a single image (e.g., a grayscale picture of a cat) for 120 s while slowly moving a bandpass window through frequency space. Thus, while subjects are presented with a large range of SFs over the entire

2-min stimulus presentation, they are exposed to only a narrow frequency window at any given point in time. In this way, time (at the resolution of the BOLD response) can be used as a proxy measure of SF. Phase-lag encoded analyses were then used to analyze the time series. This analysis approach is widely employed within the context of retinotopic mapping (Serenó et al., 1995). In the case of retinotopic mapping, a stimulus moves continuously through space, while in the current functional magnetic resonance imaging (fMRI) paradigm, the stimulus remains in the same spatial location but changes continuously along the dimension of SF. In order to prevent our conclusions about possible effects from being contaminated by saturation in BOLD contrast across the 2-min duration of each stimulus, stimuli were presented with the bandpass window moving from high to low in some subjects and from low to high in other subjects. In other words, the bandpass window could begin at the high SF range or at the low SF range. In this way, it is possible to ensure that any preferences for SF ranges exhibited by a region are not confounded with time within the 2 min of stimulation for each stimulus.

Materials and Methods

Participants

A total of 12 adults (mean age 20.9 years, 6 females, all right-handed) participated in the fMRI experiment. A total of 36 adults (mean age 21.3 years, 16 females) participated across the 2 behavioral studies (10 in Behavioral experiment 1, 9 females; 26 in Behavioral experiment 2; 6 participants were excluded from Behavioral experiment 2 for failing to keep their heads in the chinrest; of the 20 remaining in Experiment 2, 10 were female.) There was no overlap in participants between any of the experiments. Participants were recruited from the University of Rochester undergraduate and graduate communities. All participants had normal or corrected-to-normal vision. Participants gave informed consent for the study according to the University of Rochester Institutional Review Board.

Stimuli and Procedure

fMRI Experiment

Stimuli were 24 unique images (12 tools, 12 animals), each of which had been SF filtered at 300 central SFs using the MATLAB image processing toolbox (scripts available upon request). Frequency-filtered images were generated by multiplying the Fourier transform with a 1.2-octave bandpass log-of-Gaussian filter. Central SFs were evenly distributed with 0.03 octaves between each peak from 0.17 to 9.14 cycles per degree visual angle (cpd). Images were 400 × 400 pixels and subtended 8° visual angle horizontally and vertically. The root mean square contrast was equalized across all images after they were bandpass filtered.

Each scanning session consisted of 3 runs. Each run contained 4 unique images (2 tools, 2 animals; session duration = ~40 min). Participants 1 to 5 completed 2 sessions, and saw each of the 12 unique tools and animals used in Experiment 1. Participants 6 to 12 completed 1 session. Stimuli were displayed using the MATLAB Psychophysics Toolbox (Brainard 1997; Pelli 1997) and back projected onto a screen that was visible through a mirror mounted on the MRI head coil. Item order was semirandomized within run with the restriction that no 2 items from the same category could be presented sequentially. Participants were instructed to free-view 2-min long “movies” of the stimulus items moving through frequency space (see Fig. 1A). Stimuli gradually transitioned from LSF to HSF (subjects 1, 3, and 4), or the reverse (subjects 2, 5, 6, and 7–12). Each subject saw only one version of the stimuli (i.e., either all movies went from low SF to high SF or the reverse). Critically, only a narrow range of SFs (1.2 octaves)

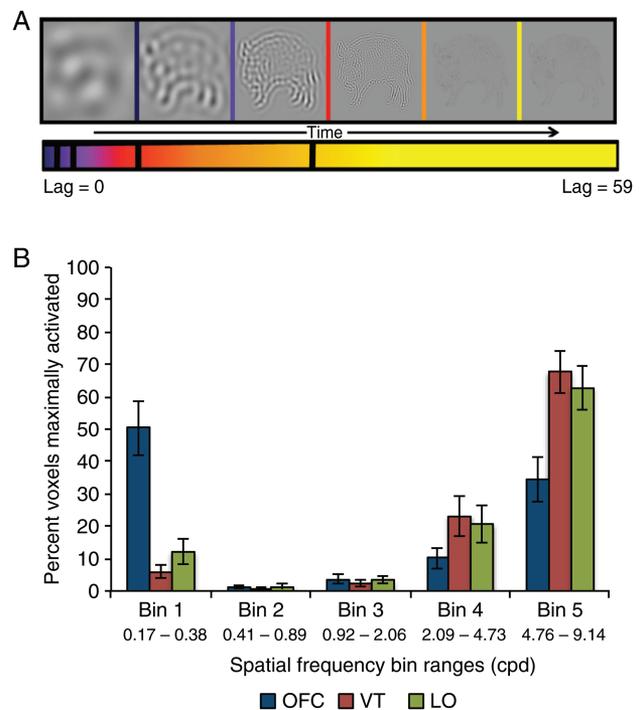


Figure 1. (A) Schematic of the stimulus “movies” that subjects were shown during functional magnetic resonance imaging (fMRI). The duration of each movie was 120 s and depicted a single object in which a bandpass window of SF information was changed continuously frame-by-frame. (B) Proportion of voxels maximally activated by each SF range in each cortical location. Error bars (standard error of the mean) are calculated over hemispheres studied ($n = 24$). The plot shows that OFC exhibits a bimodal distribution of SF preferences, whereas LO and VT exhibit a unimodal distribution biased toward mid-to-high SFs. OFC: orbital frontal cortex; LO: lateral occipital; VT: ventral temporal cortex.

was present for any given frame (frame duration = 400 ms). All images were presented in the center of the visual field. Runs began and ended with 16 s of a fixation cross. The interval between the movie presentations was jittered between 2 and 8 s drawn from a distribution with hyperbolic density.

Experiment 1b

In a subset of subjects (7–12), VT and lateral occipital (LO) cortices were functionally defined with an objects versus scrambled objects localizer (see Table 1 for Talairach coordinates). The localizer used a standard procedure (e.g., Fang and He 2005; Mahon et al. 2009) and included faces, places, animals, and tools (12 items per category, 8 exemplars for each item, yielding 384 stimuli). Stimuli were presented in mini-blocks of duration 6 s (500-ms duration, 0-ms inter-stimulus interval) interspersed by 12-s fixation periods. Within each run, 8 mini-blocks of intact stimuli and 4 mini-blocks of phase-scrambled versions of the same stimuli were presented. Order was randomized within blocks/mini-blocks such that every cell of the design was replicated twice after 8 block presentations. Each of the 6 participants who had also completed the main experiment completed 8 runs of the object localizer in a separate scanning session. Regions of interest (ROIs) were defined as cortical locations showing greater activation to intact compared with scrambled images, thresholded at false discovery rate (FDR) $q < 0.05$ (or stricter) for each subject.

MR Acquisition and Analysis

Whole-brain BOLD imaging was conducted on a 3-Tesla Siemens MAGNETOM Trio scanner with a 32-channel head coil located at the Rochester Center for Brain Imaging. High-resolution structural T_1 contrast images were acquired using a magnetization prepared rapid gradient echo pulse sequence at the start of each session (TR = 2530 ms,

Table 1

Regions of interest for group-level analyses were drawn as spheres (10 mm diameter) centered on Talairach coordinates for the orbital frontal cortex (OFC), ventral temporal cortex (VT), and lateral occipital cortex (LO), as derived from the previous literature

	Left hemisphere			Right hemisphere			Citations					
	x	y	z	x	y	z						
Defined: previous literature												
OFC	-30	-28	-11	10	23	-19	Kveraga, Boshyan et al. (2007)					
VT	-31	-50	-8	40	-48	-9	Gauthier et al. (2000)					
LO	-44	-67	-3	42	-67	-4	Grill-Spector et al. (2000)					
Defined: functionally												
VT	-31	±1.8	-50	±2.8	-12	±1.3	35	±1.8	-56	±2.8	-15	±1.5
LO	-37	±0.6	-71	±2.1	-6	±1.7	35	±1.3	-74	±3.1	-9	±2.8

Note: A subset of participants ($n = 6$) participated in a category localizer to functionally define VT and LO. These regions were defined as those areas that showed greater activation to intact compared with scrambled images ("intact > scrambled"; FDR $q < 0.05$).

TE = 3.44 ms, flip angle = 7°, FOV = 256 mm, matrix = 256 × 256, 1 × 1 × 1-mm sagittal left-to-right slices). An echo-planar imaging pulse sequence was used for T_2^* contrast (TR = 2000 ms, TE = 30 ms, flip angle = 90°, FOV = 256 mm, matrix 64 × 64, 30 sagittal left-to-right slices, voxel size = 4 × 4 × 4 mm). The first 6 volumes of each run were discarded to allow for signal equilibration.

fMRI data were analyzed with the "Brain Voyager" software package (Version 2.1) and in-house scripts drawing on the BVQX toolbox written in MATLAB (wiki2.brainvoyager.com/BVQXtools). Preprocessing of the functional data included, in the following order: slice scan time correction (sinc interpolation), motion correction with respect to the first volume of the first functional run, and linear trend removal in the temporal domain (cutoff: 2 cycles within the run). Functional data were registered (after contrast inversion of the first volume) to high-resolution deskulled anatomy on a participant-by-participant basis in native space. For each individual participant, echo-planar and anatomical volumes were transformed into standardized (Talairach and Tournoux 1988) space. Functional data were smoothed at 6 mm (1.5 voxels) at full-width at half-maximum, and interpolated to 3 × 3 × 3 mm voxels.

Behavioral Experiment 1

Stimuli were identical to those of the fMRI experiment. Images were filtered with a 1.2-octave bandpass filter; the central SF was shifted 0.03 octaves between each image. Subjects saw a subset of 8 randomly chosen items (4 tools, 4 animals) from the image set used in the fMRI experiment. Image order was randomized. Each image was seen once at each SF level, yielding 8 images at each SF level. Each trial consisted of a central fixation cross presented for 300 ms, followed by an image presented for 50 ms, and then a 75-ms high contrast noise mask (See Fig. 5A). Images were 400 × 400 pixels and subtended 12.2° visual angle horizontally and vertically. Participants were seated 0.5 m away from the monitor. Distance was kept constant by instructing participants to keep their heads in a chinrest (UHCO standard Head Spot chinrest). Stimulus presentation and behavioral response recording were controlled using E-Prime software (as well as for Behavioral experiment 2, see below). Responses were recorded with a button box with ms precision (Psychology Software Tools, model 200A, 0-ms debounce period).

Behavioral Experiment 2

Stimuli were 48 broadband target images (12 phase-aligned images from each of 4 categories; animals, faces, tools, places). To facilitate comparison between Behavioral experiments 1 and 2 and the fMRI experiment, which used only tools and animals, only data from the tool and animals conditions are reported herein. Six SF-filtered prime images were generated for each target image using the MATLAB image processing toolbox (Mathworks, Inc., Sherbon, MA). The central SF of primes was evenly distributed in Fourier space, ranging from very low SF (0.17 cpd) to very high SF (11.1 cpd), following the same approach for filtering described above (1.32-octaves bandwidth

around the central SF, central SFs: 0.17, 0.40, 0.92, 2.12, 4.87, and 11.1 cpd). Twelve phase-randomized broadband primes were also created for each category to use as a baseline prime condition. The root mean square contrast was equalized across all images after they were bandpass filtered. Images were 400 × 400 pixels and subtended 12.2° visual angle horizontally and vertically. Participants' distance from the monitor was the same as in Behavioral experiment 1.

Each experimental trial began with a central fixation cross (500 ms) followed by a high contrast noise image (forward mask; 148 ms), followed by the prime (33 ms), followed by a different high contrast noise image (backward mask; 148 ms) followed by the broadband target image. Primes, when they were not scrambled images, were always identical to the target images. Targets were presented for 3000 ms or until a response was made (Fig. 6A). Trials with phase-randomized primes were identical to experimental trials, and phase-randomized images were always paired with the broadband nonscrambled image used to generate the prime. Participants were instructed to categorize the target images as "living" or "nonliving" as quickly and accurately as possible (button response). Blocks consisted of 384 trials. Blocks were counterbalanced, and image order within each block was randomized. Subject 2 completed 5 blocks; all other subjects completed 6 blocks.

After completing the priming experiment, participants completed a prime awareness task. Each prime was presented twice, once for 33 ms and once for 50 ms, flanked by forward and backward high contrast noise masks (duration = 148 ms). The order of the images was randomized. Only data from primes presented for 33 ms were analyzed; the longer duration prime presentation was to provide participants with some information about what they should be looking for in the prime discrimination task.

Results

fMRI

To answer the question of whether SFs implicated in the final stages of object recognition-driven BOLD responses in OFC, we compared the distribution of SF preferences in OFC to the distributions in LO cortex and (VT), 2 ventral stream regions known to be critical for object identification (Malach et al. 1995; Chao et al. 1999; Grill-Spector et al. 2001; Haxby et al. 2001; Kourtzi et al. 2003; O'Toole et al. 2005). The continuous SF stimuli (duration = 120 s) were binned into 5 SF ranges (0.17–0.38, 0.41–0.89, 0.92–2.06, 2.09–4.73, and 4.76–9.14 cpd). Each bin corresponded to ~1.2 octaves of frequency space. In the first-level analysis, all ROIs were defined independently of the current dataset, as spheres centered at Talairach coordinates obtained from previous work on object processing (Gauthier et al. 2000; Grill-Spector et al. 2000; Kveraga, Boshyan et al. 2007; See Table 1, Fig. 2). A subset of

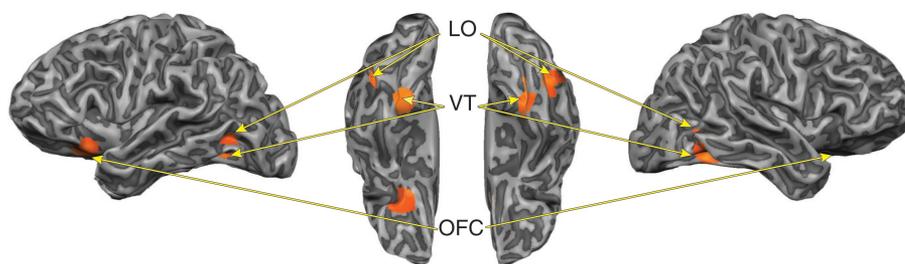


Figure 2. Projection of 10-mm spheres centered on Talairach coordinates from previous literature (see Table 1) onto cortical surface.

the subjects (7 of 12) completed a localizer for VT and LO in a separate scanning session; thus, in a second level analysis, VT and LO was functionally defined on a subject-by-subject basis and the pattern observed for the literature-defined ventral stream ROIs was compared with the pattern observed for the functionally defined ROIs. The distributions of SF preferences were calculated as the proportion of voxels within the ROI that exhibited a preference for each SF range (thresholded at $r \geq 0.13$, FDR $q < 0.05$).

In a first analysis, ROI data were analyzed using a 2-way ANOVA with cortical region (3 levels; OFC, VT, LO) and SF (5 levels). There was a main effect of SF ($F_{4,92} = 34.22$, $MSE = 33\,822.32$, $P < 0.0001$), and an interaction between SF and region ($F_{8,184} = 12.09$, $MSE = 5676.92$, $P < 0.0001$; there can be no main effect of region because the data are expressed as proportions summing to 100). As can be seen in Figure 1B, the distribution of voxels responding maximally to each SF range is bimodal in OFC and skewed toward high SFs in the ventral stream ROIs. Pairwise comparisons in OFC revealed that very low SFs (Bin 1; 0.17–0.38 cpd) and mid-to-high SFs (Bins 4 and 5; 2.09–9.14 cpd) elicit differential BOLD contrast when compared with mid-to-low SFs (Bin 2; 0.41–0.89 cpd; $t(23) = 5.66$, $P < 0.0001$; $t(23) = 2.62$, $P < 0.01$; $t(23) = 4.72$, $P < 0.0001$; Bins 1, 4, and 5 compared with Bin 2, respectively.) There was no difference in activation between Bins 2 and 3 ($t(23) = -1.47$, $P = 0.15$). Pairwise comparisons further revealed that low SFs drive activation significantly more in OFC than in LO ($t(23) = 4.20$, $P < 0.0001$) or in VT ($t(23) = 5.28$, $P < 0.0001$).

The pattern of responses in LO and VT were very similar, suggesting that the 2 regions could be collapsed for simplicity. This was confirmed by a 2-way ANOVA (main effect of SF [$F_{4,92} = 34.69$, $MSE = 33\,874.71$, $P < 0.0001$; no interaction between LO/VT and SF, $F_{4,92} = 1.57$, $MSE = 210.70$, $P > 0.1$]). Having collapsed the data across the 2 ventral stream regions, we reconfirmed the presence of an interaction between region (ventral stream, OFC) and SF level ($F_{4,92} = 13.82$, $MSE = 8357.36$, $P < 0.0001$). We explored whether the dissociation between OFC and the ventral stream was affected by the direction of the stimulus movie. However, there was no interaction between the factors cortical region, SF, and movie direction ($F_{4,88} = 1.167$, $MSE = 700.77$, $P > 0.1$), so movie direction was excluded as a factor for further consideration.

We then asked whether the patterns observed in the ventral stream and OFC differed by hemisphere. In an ANOVA with hemisphere (2 levels; left, right), SF bin (5 levels), and cortical region (2 levels; ventral stream and OFC), there was an interaction ($F_{4,88} = 28.05$, $MSE = 7795.51$, $P < 0.0001$). Figure 3 shows the data broken down by hemisphere. In a 2-way ANOVA for the left hemisphere (cortical region, SF

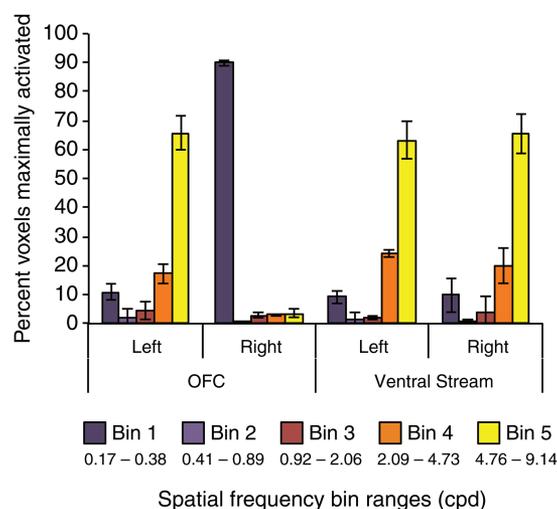


Figure 3. Hemispheric differences in the proportion of voxels maximally activated by each SF range in OFC compared with the ventral stream. Error bars (standard error of the mean) were calculated separately over hemispheres ($n = 12$). The plot shows that right OFC exhibits a low SF preference, driving the bimodal peak shown in Figure 1. Left OFC and the bilateral ventral stream ROIs exhibit mid-to-high SF biases.

level), there was a main effect of SF ($F_{4,44} = 37.2$, $MSE = 16\,382.9$, $P < 0.0001$), and in contrast with the Bar et al. (2006) results, we found no interaction between cortical region and SF ($F_{4,44} = 0.3$, $MSE = 79.4$, $P > 0.5$). This means that in the left hemisphere, OFC and the ventral stream are driven by the same SF ranges. In the right hemisphere, there was both a main effect of SF ($F_{4,44} = 33.4$, $MSE = 109\,65.2$, $P < 0.0001$) and a significant interaction between cortical region and SF ($F_{4,44} = 57.6$, $MSE = 16\,073.4$, $P < 0.0001$). Pairwise comparisons show that right OFC responds more to low SFs than the right ventral stream ($t(11) = 17.39$, $P < 0.0001$). In contrast to the left and right OFC, the left and right ventral streams did not exhibit a SF versus cortical region interaction ($F_{4,44} = 0.6$, $MSE = 32.7$, $P > 0.5$).

Two representative brains and the proportions of voxels responding to each SF range are shown in Figure 4. The group-level analysis holds at the single-subject level, as evidenced by the histograms in the figure. For both subjects, Left OFC behaves like the ventral object-processing stream, responding maximally to higher SFs, whereas right OFC is driven more strongly by LSFs.

In a second series of analyses, VT and LO were functionally defined (in a subset, 7 of 12, subjects, as described above). To confirm that there were no differences in SF preferences between the ventral stream functionally defined and literature-based ROIs, we conducted a mixed ANOVA with cortical

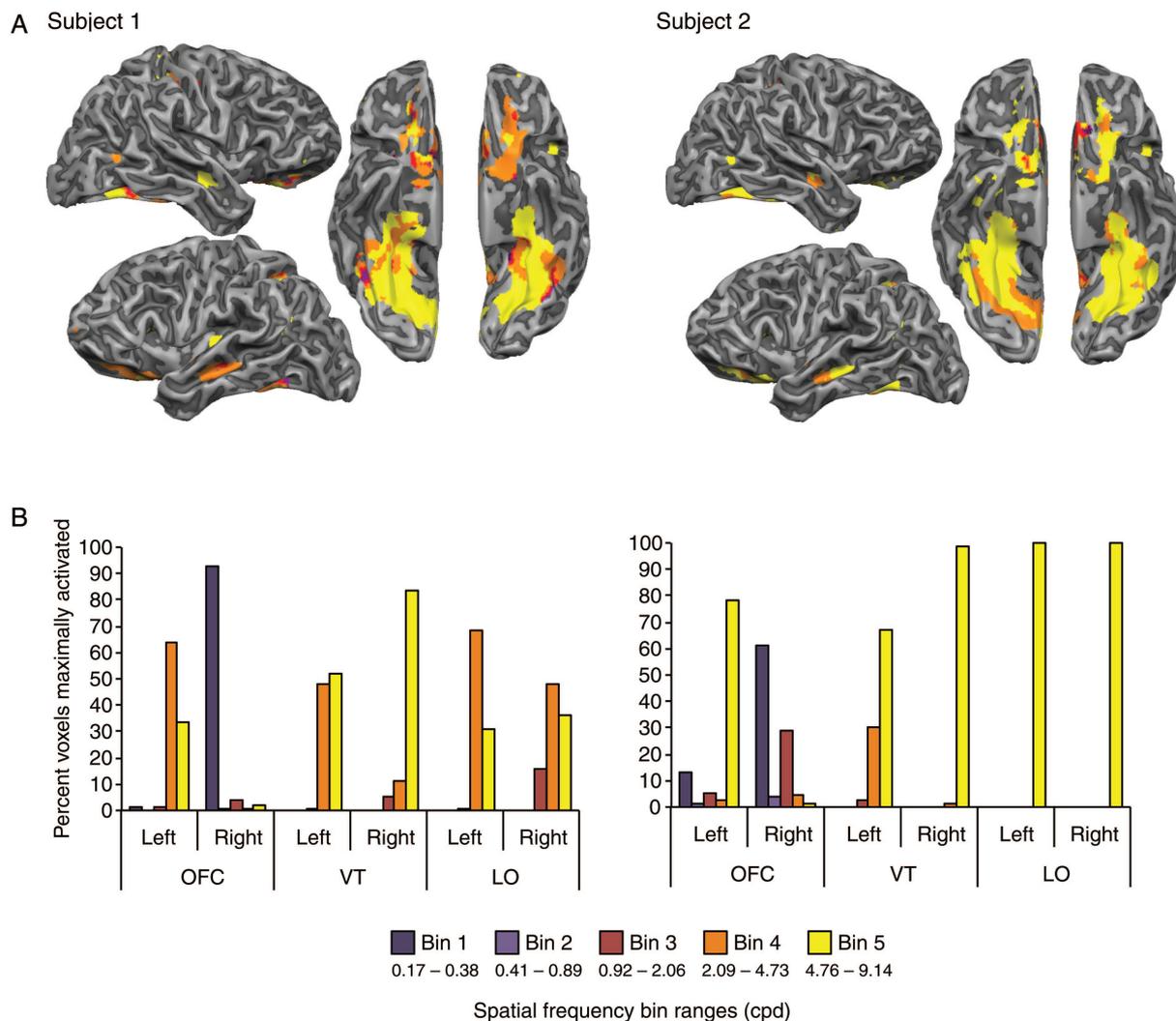


Figure 4. Representative single-subject analyses. (A) Phase-lag maps showing low SF peaks in the right OFC and a bias toward mid-to-high SF ranges in the left OFC and ventral stream. The analysis was restricted to voxels (whole brain) that showed activation to all regressors (“all on”; phase-lag analysis, 60 lags, 2 s TR). Activated voxels are depicted in the color corresponding to the SF range to which they responded most strongly (see legend). (B) Proportion of voxels maximally responsive to each SF range in the same ROIs used for the group-level analysis. Subjects are the same as in (A). Values in the histogram are separated by hemisphere. OFC: orbital frontal cortex; LO: lateral occipital; VT: ventral temporal cortex.

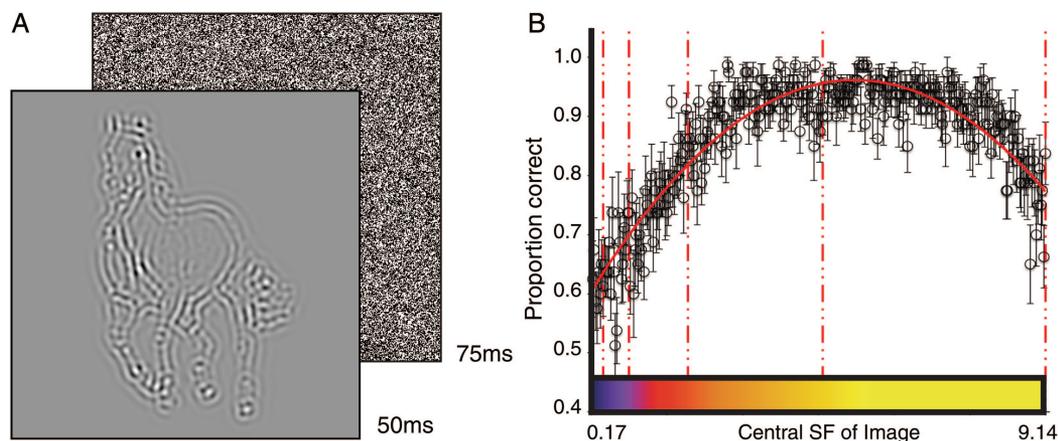


Figure 5. (A) Example stimulus and schematic of trial structure for Behavioral experiment 1. Trials begin with a 300-ms fixation cross (omitted in the schematic), followed by a backward masked bandpass filtered target image. (B) Categorization accuracy as a function of the central SF of the target image. The vertical bins correspond to the SF bins used in the analysis of the fMRI experiment. Bins 4 and the first half of Bin 5 are at or near ceiling, demonstrating the SF range at which SF-filtered images are most easily categorized. Error bars (standard errors of the mean) are calculated over subjects.

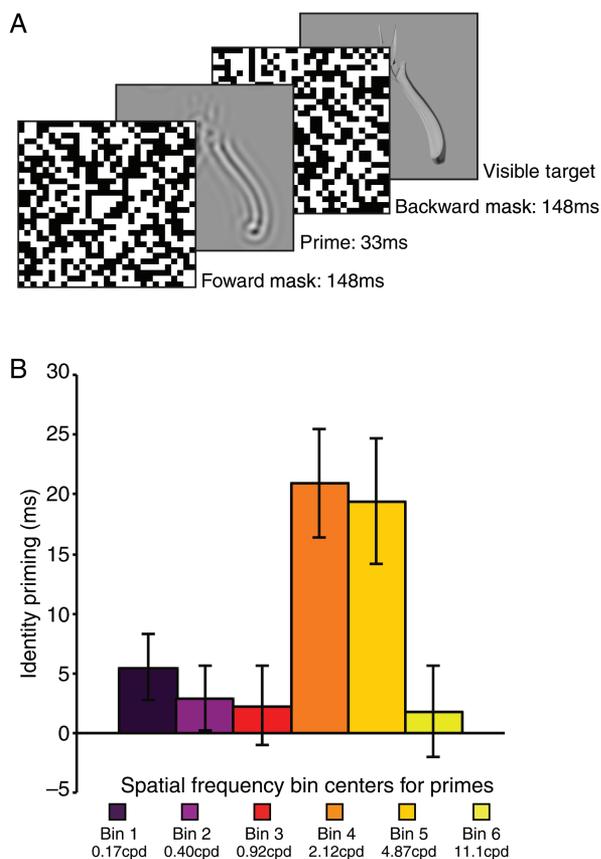


Figure 6. (A) Example stimulus and trial structure for Behavioral experiment 2. Trials began with a 500-ms fixation cross (omitted in the schematic), followed by a forward and backward masked prime image. The backward mask was followed by a broadband target image identical to the prime image and presented for 3000 ms or until a categorization response was made. On one-seventh of the trials, a broadband, phase-scrambled image drawn from the same category as the target replaced the SF-filtered prime, serving as a baseline against which to measure identity priming effects. (B) Identity priming effects from Behavioral experiment 2. The vertical axis plots mean priming effects (error bars are standard errors of the mean across subjects). The ordinate is drawn such that a positive number indicates faster categorization than scrambled baseline. After completing the priming task, all participants completed a prime discrimination task (see Materials and Methods section for details). The pattern of prime discrimination by SF replicated the pattern observed in Behavioral experiment 1 (Fig. 5). One-sample *t*-tests (reference point = chance of 50%) reveal that only primes in SF bins 4 and 5 could be reliably categorized ($t(19) = 4.36, P < 0.001, t(19) = 2.85, P < 0.01$, respectively).

location (2 levels: VT and LO) and SF range (5 levels, see bins above) as within subjects factors, and hemisphere (2 levels; left and right) and ROI definition (2 levels; functionally and literature defined) as between subjects factors. All interactions with the factor ROI definition were nonsignificant (SF range \times ROI definition: $F_{4,80} = 0.14, P > 0.1$; SF range \times hemisphere \times ROI definition: $F_{4,80} = 0.09, P > 0.1$; Cortical location \times SF range \times ROI definition: $F_{4,80} = 0.11, P > 0.1$; Cortical location \times SF range \times hemisphere \times ROI definition: $F_{4,80} = P > 0.1$). This indicates that the pattern of responses reported above using ROIs from the previous literature is not different from the pattern observed when the ROIs are functionally defined.

Psychophysics

We then conducted 2 behavioral experiments to independently confirm 2 assumptions that were made about the experimental materials by the fMRI analyses. The first

assumption is that the SF ranges labeled as mid to high are critical for supporting object identification, a process known to depend on processing in the ventral stream. The second assumption is that the approach of binning SF into ranges defined on the basis of ~ 1 -octave change in frequency space captures the cortically and behaviorally relevant variation in SF across all frames in the stimulus movies.

In the first behavioral experiment, participants were required to categorize (living or nonliving) individual frames from the movies that had been used in the fMRI experiment (see Fig. 5A for an example stimulus and schematic of the trial structure). The images were backward masked in order to bring performance off of ceiling, and to thus be able to describe the specific role of different SF ranges in object recognition. Figure 5B, which plots accuracy as a function of SF, shows that performance is highest for mid-to-high SF ranges (second-order polynomial fit, $r^2 = 0.83$). These data also provide independent confirmation for the procedure of binning by 1.2 octaves for the sake of the phase-lag analysis.

In a second behavioral experiment, we sought additional evidence using the implicit measure of priming. SF-filtered prime images (bandpass centers at 0.17, 0.40, 0.92, 2.12, 4.87, or 11.1 cpd) were briefly presented (duration = 33 ms) and forward and backward masked with high contrast pattern masks (mask duration = 148 ms; see Figure 6A for a schematic; for precedent on this procedure see Breitmeyer and Ganz 1976; Almeida et al. 2008, 2010). The offsets of the backward masks were immediately followed by the onset of a visible target image, which participants categorized as “living” or “nonliving.” A baseline of scrambled primes was used to evaluate the magnitude of priming elicited on the targets. Average response time differences from baseline for each SF range are shown in Figure 6B (primary *y*-axis). A repeated-measures ANOVA revealed a main effect of SF ($F_{5,95} = 8.33, \text{MSE} = 188.92, P < 0.001$) and simple contrasts (using SF range 6 as a reference) showed that this effect is driven by SF ranges 4 ($F_{1,19} = 14.56, \text{MSE} = 498.63, P < 0.001$; 2.12 cpd) and 5 ($F_{1,19} = 14.27, \text{MSE} = 431.17, P < 0.001$; 4.87 cpd). Thus, priming significantly affected reaction time when the primes were filtered to include mid-to-high, but not low (central SF 0.17, 0.40, 0.92) or very high (central SF 11.1) SFs.

Discussion

Previous research indicates that the information about global shape represented and processed by OFC is conveyed through magnocellular channels within the visual system (Bar 2003; Bar et al. 2006; Fenske et al. 2006; Kveraga, Boshyan et al. 2007; Kveraga, Ghuman, et al. 2007; Peyrin et al. 2010). However, whether top-down signals from OFC to ventral stream regions are followed by error-driven feedback about the outcome of object recognition processes has been only a theoretical speculation. We have reported that OFC exhibits a bimodal peak in SF preferences across the 2 hemispheres, with differential responses for both low and mid-to-high SFs. These results stand in contrast to the response distributions in LO and VT, in which there were either no, or significantly dampened, responses to LSF information, but responses equivalent to those of OFC for mid-to-high SFs. When the data were separated by hemisphere, the distribution of responses in left OFC resembled that of the ventral stream, whereas the distribution of right OFC was in sharp contrast to

that of the ventral stream, displaying a preference for low SF information. The modulation in SF preferences in OFC confirms the expectations laid out by Bar (2003). More generally, these data indicate that human OFC processes both an initial “first guess” about global shape information, and receives detailed information about object identity conveyed by the mid-to-high SF range.

The fact that left OFC exhibits a peak of activation late in the process of stimulus processing (i.e., in response to HSF information which would come later than LSF information), places an important constraint on models of object processing in that the “priors” that are generated by OFC can be shaped by the outcome of object recognition. Our findings also indicate a specific neural basis for feedback and feed forward mechanisms that have been posited at a theoretical level (Bar 2003). Furthermore, they establish a specific framework for interhemispheric transfer of information between left and right OFC in developing and refining predictions about object identity that can be pursued with future work.

The 2 psychophysical experiments reported above provide independent validation for the ranges of SF that were used in the analysis of the fMRI data, as well as for the supposition that SFs categorized herein as mid to high are particularly important for object identification. An aspect of the entire pattern of findings is the contrast between the low categorization accuracy for stimuli in the low SF ranges and the strong response to those same low SF ranges in right OFC. This finding may suggest that “activation,” as it is operationalized herein, does not simply track informativeness or saliency—it also tracks the processing characteristics local to each brain region. This finding may additionally indicate that low categorization accuracy may relate to many “initial hypotheses” being generated on the basis of low SF information.

OFC is known to have bidirectional connections with visual areas and has been implicated in reward processing (Rolls and Baylis 1994; Rolls 2000; Kringelbach and Rolls 2004). Thus, in addition to guiding a coarse, magnocellular driven interpretation of stimuli, it is reasonable to speculate that OFC is reactivated once the mid-to-high SF information from a stimulus has driven object identification, thus strengthening the trace associated with the initial coarse guess. This would reinforce the low SF information associated with a given set of mid-to-high SFs, and serve to speed recognition of that set of low SFs in future processing. More broadly, these results suggest that OFC is involved both in generating a number of “initial hypotheses” about stimulus identity, as well as representing exact stimulus identity, thus providing a strong candidate region for integrating initial coarse information with definite information about object identity. While the current report has focused on the response characteristics of OFC because of the clear theoretical motivation regarding this region, an important future direction would be to test whether other regions are also sensitive to both low and high SF information.

A number of issues are framed by our findings that merit further investigation. For instance, it may be asked whether there is overlap in the response preferences of different subpopulations of voxels. The analysis approach adopted within the current report, phase-lag analysis, is not suited to address such questions of overlap as voxels are categorized as exhibiting a preference for a given range of SF information. However, relating the phase-lag approach to other

approaches such as multivariate pattern analyses would provide additional information about potential overlap in the response distributions of subpopulations of voxels within OFC. Additionally, functional connectivity analyses, particularly between left and right OFC may shed new light on the dynamics of information updating throughout the course of visual object recognition.

Another issue that is framed by our findings concerns why OFC would receive feedback about the outcome of object recognition processes. We have suggested, on the basis of the models and evidence reviewed in the introduction, that such feedback may form an important aspect of how the system comes to be more efficient at recognizing visual input. In particular, how the system adapts to handle uncertainty in a more efficient manner is an important function that such a feedback mechanism would be well suited to address. However, whether or not the feedback suggested here is, in fact, integral to object recognition processes, or learning about the identity of visual input over repeated presentations, or is merely an epiphenomenal by-product of recognition processes needs to be directly addressed through future research.

Given the above considerations, and the evidence reported herein, it is somewhat surprising that patients with OFC lesions do not seem to exhibit obvious impairments in object recognition. However, it may be that such patients fail to exhibit an ability to efficiently adapt to uncertain visual information. Studies with patients with lesions to the left or right OFC would be well positioned to address the causal role that OFC plays in object recognition. fMRI in such patients may be able to test the key prediction that the ventral stream together with left and right OFC forms an integrated circuit for generating and refining predictions about object identity based on a fast and coarse “first-pass” analysis of visual input.

Funding

This research was supported by NIH grant R21 NS076176 to B.Z.M. Additional funds supporting this research were contributed by Norman and Arlene Leenhouts.

Notes

We thank Alena Stasenko and Frank Garcea for assistance with running the fMRI experiment, and Brittany Eltman for running Behavioral experiment 2 and for her contribution to the design and analysis of that experiment. *Conflict of Interest:* None declared.

References

- Almeida J, Mahon BZ, Caramazza A. 2010. The role of the dorsal visual processing stream in tool identification. *Psychol Sci.* 21:772–778.
- Almeida J, Mahon BZ, Nakayama K, Caramazza A. 2008. Unconscious processing dissociates along categorical lines. *Proc Natl Acad Sci USA.* 105:15214–15218.
- Baker CI, Olson CR, Behrmann M. 2004. Role of attention and perceptual grouping in visual statistical learning. *Psychol Sci.* 15:460–466.
- Bar M. 2003. A cortical mechanism for triggering top-down facilitation in visual object recognition. *J Cognit Neurosci.* 15:600–609.
- Bar M, Kassam KS, Ghuman AS, Boshyan J, Schmid AM, Dale AM, Hämäläinen MS, Marinkovic K, Schacter DL, Rosen BR et al. 2006. Top-down facilitation of visual recognition. *Proc Natl Acad Sci USA.* 103:449–454.

- Behrmann M, Zemel RS, Mozer MC. 1998. Object-based attention and occlusion: evidence from normal participants and a computational model. *J Exp Psychol Hum Percept Perform.* 24:1011–1036.
- Brainard DH. 1997. The Psychophysics Toolbox. *Spatial Vision.* 10:433–436.
- Breitmeyer BG, Ganz L. 1976. Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. *Psychol Rev.* 83:1–36.
- Bullier J. 2001. Integrated model of visual processing. *Brain Res Rev.* 36:96–107.
- Chao LL, Haxby JV, Martin A. 1999. Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nat Neurosci.* 2:913–919.
- Derrington AM, Lennie P. 1984. Spatial and temporal contrast sensitivities of neurones in lateral geniculate nucleus of macaque. *J Physiol.* 357:219–240.
- Fang F, He S. 2005. Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nat Neurosci.* 8:1380–1385.
- Fenske MJ, Aminoff E, Gronau N, Bar M. 2006. Top-down facilitation of visual object recognition: object-based and context-based contributions. *Prog Brain Res.* 155:3–21.
- Foxe JJ, Simpson GV. 2002. Flow of activation from V1 to frontal cortex in humans. A framework for defining “early” visual processing. *Exp Brain Res.* 142:139–150.
- Gauthier I, Skudlarski P, Gore JC, Anderson AW. 2000. Expertise for cars and birds recruits brain areas involved in face recognition. *Nat Neurosci.* 3:191–197.
- Grill-Spector K, Kourtzi Z, Kanwisher N. 2001. The lateral occipital complex and its role in object recognition. *Vision Res.* 41:1409–1422.
- Grill-Spector K, Kushnir T, Hendler T, Malach R. 2000. The dynamics of object-selective activation correlate with recognition performance in humans. *Nat Neurosci.* 3:837–843.
- Hauk O, Davis MH, Ford M, Pulvermüller F, Marslen-Wilson WD. 2006. The time course of visual word recognition as revealed by linear regression analysis of ERP data. *NeuroImage.* 30:1383–1400.
- Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science.* 293:2425–2430.
- Hughes HC, Nozawa G, Kitterle F. 1996. Global precedence, spatial frequency channels, and the statistics of natural images. *J Cognit Neurosci.* 8:197–230.
- Juan CH, Walsh V. 2003. Feedback to V1: a reverse hierarchy in vision. *Exp Brain Res.* 150:259–263.
- Kourtzi Z, Erb M, Grodd W, Bühlhoff HH. 2003. Representation of the perceived 3-D object shape in the human lateral occipital complex. *Cereb Cortex.* 13:911–920.
- Kringelbach ML, Rolls ET. 2004. The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. *Prog Neurobiol.* 72:341–372.
- Kveraga K, Boshyan J, Bar M. 2007. Magnocellular projections as the trigger of top-down facilitation in recognition. *J Neurosci.* 27:13232–13240.
- Kveraga K, Ghuman AS, Bar M. 2007. Top-down predictions in the cognitive brain. *Brain Cogn.* 65:145–168.
- Lamb MR, Yund EW. 1993. The role of spatial frequency in the processing of hierarchically organized stimuli. *Percept Psychophys.* 54:773–784.
- Livingstone M, Hubel D. 1988. Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science.* 240:740–749.
- Mahon BZ, Anzellotti S, Schwarzbach J, Caramazza A. 2009. Category-specific organization in the human brain does not require visual experience. *Neuron.* 63:397–405.
- Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy WA, Ledden PJ, Brady TJ, Rosen BR, Tootell RB. 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc Natl Acad Sci USA.* 92:8135–8139.
- Maunsell JHR, Ghose GM, Assad JA, McAdams CJ, Boudreau CE, Noerager BD. 1999. Visual response latencies of magnocellular and parvocellular LGN neurons in macaque monkeys. *Vis Neurosci.* 16:1–14.
- Olds ES, Engel SA. 1998. Linearity across spatial frequency in object recognition. *Vision Res.* 38:2109–2118.
- O’Toole AJ, Jiang F, Abdi H, Haxby JV. 2005. Partially distributed representations of objects and faces in ventral temporal cortex. *J Cognit Neurosci.* 17:580–590.
- Parker DM, Lishman JR, Hughes J. 1996. Role of coarse and fine spatial information in face and object processing. *J Exp Psychol Hum Percept Perform.* 22:1448–1466.
- Pelli DG. 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision.* 10:437–442.
- Peyrin C, Michel CM, Schwartz S, Thut G, Seghier M, Landis T, Marendaz C, Vuilleumier P. 2010. The neural substrates and timing of top-down processes during coarse-to-fine categorization of visual scenes: a combined fMRI and ERP study. *J Cognit Neurosci.* 22:2768–2780.
- Potter MC. 1976. Short-term conceptual memory for pictures. *J Exp Psychol Hum Learn Mem.* 2(5):509–522. Available from URL <http://www.ncbi.nlm.nih.gov/pubmed/1003124> [date last accessed; May 2012].
- Riesenhuber M, Poggio T. 2002. Neural mechanisms of object recognition. *Curr Opin Neurobiol.* 12:162–168.
- Rolls ET. 2000. The orbitofrontal cortex and reward. *Cereb Cortex.* 10:284–294.
- Rolls ET, Baylis LL. 1994. Gustatory, olfactory, orbitofrontal cortex and visual convergence within the primate. *J Neurosci.* 14:5437–5452.
- Rossion B, Joyce CA, Cottrell GW, Tarr MJ. 2003. Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. *NeuroImage.* 20:1609–1624.
- Schyns PG, Oliva A. 1999. Dr. Angry and Mr. Smile: when categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition.* 69:243–265.
- Schyns PG, Oliva A. 1994. From blobs to boundary edges: evidence for time- and spatial-scale-dependent scene recognition. *Psychol Sci.* 5:195–200.
- Sereno MI, Dale A, Reppas JB, Kwong KK, Belliveau J, Brady TJ, Rosen B, Tootell RB. 1995. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science.* 268:889–893.
- Talairach J, Tournoux P. 1988. Co-planar stereotaxic atlas of the human brain. New York: Thieme Medical Publishing.
- Tanaka JW, Curran T. 2001. A neural basis for expert object recognition. *Psychol Sci.* 12:43–47.
- Tootell RB, Silverman MS, Hamilton SL, Switkes E, De Valois RL. 1988. Functional anatomy of macaque striate cortex. V. Spatial frequency. *J Neurosci.* 8:1610–1624.
- VanRullen R, Thorpe SJ. 2001. The time course of visual processing: from early perception to decision-making. *J Cognit Neurosci.* 13:454–461.
- Yuille A, Kersten D. 2006. Vision as Bayesian inference: analysis by synthesis? *Trends Cogn Sci.* 10:301–308.