A matched comparison of binocular rivalry and depth perception with fMRI

Athena Buckthought
McGill Vision Research Unit, Department of Ophthalmology, McGill University, Montreal, Quebec, Canada

Janine D. Mendola
McGill Vision Research Unit, Department of Ophthalmology, McGill University, Montreal, Quebec, Canada

Psychophysical experiments have demonstrated that it is possible to simultaneously perceive binocular depth and rivalry in plaids (A. Buckthought & H. R. Wilson, 2007). Here, we used fMRI at 3T to image activity in the visual cortex while human subjects perceived depth and rivalry from plaids. Six subjects performed either a rivalry or depth task. The spatial frequencies of the near-vertical and diagonal components were, respectively: 2.5, 6.4 cpd; 6.4, 2.5 cpd; or 6.4, 6.4 cpd. The network of activated cortical areas was very similar for the depth compared to the rivalry task. Nevertheless, regions of superior and inferior parietal cortices (including intraparietal sulcus) were activated more during the depth than the rivalry task, independent of spatial frequency, whereas a bias toward rivalry was seen in a lateral occipital region, superior temporal sulcus, and retrosplenial and ventral temporal cortices. Several retinotopic areas in the visual cortex showed a preference for the task with the higher (V1, V2, V3) or lower spatial frequency component (MT+), regardless of the depth or rivalry condition. Our results indicate that depth and rivalry are processed in a similar network of cortical areas and are perceived simultaneously by coexisting in different spatial channels. These results place constraints on binocular vision models.

Keywords: binocular vision, functional imaging, visual cortex, depth, attention, 3D surface and shape perception, binocular rivalry, stereopsis


Introduction

The relationship between stereoscopic depth and binocular rivalry has been a longstanding issue with implications for models of binocular vision. A number of studies have shown interactions that imply that depth and rivalry are part of the same neural process. For example, the presence of rivalry can disrupt stereopsis; it has been shown that inhibition of one stereo half-image prevents a match to the other image (Harrad, McKee, Blake, & Yang, 1994). Some investigators have also proposed that rivalry ensues when stereopsis fails (Blake, 1989, 2001). In one particular model, rivalry occurs with unmatched components in the images, following binocular matching (Hayashi, Maeda, Shimojo, & Tachi, 2004). Another proposal is that rivalry generally reflects the inhibition of stereoscopic false matches (Blake, Yang, & Wilson, 1991; Rohaly & Wilson, 1993; Wilson, Blake, & Halpern, 1991).

It is also relevant that distinct spatial frequency channels are proposed to be involved in stereopsis and binocular rivalry, as shown by the effects of masking noise on the perception of depth in isotropic filtered noise patterns (Julesz & Miller, 1975). These channels may operate independently, in that one set of channels can be engaged in rivalry while other channels simultaneously support stereopsis. Supporting this idea, it is possible, with isotropic noise, to perceive depth in the presence of rivalry, as long as the components are in different spatial frequency bands (Blake et al., 1991). A recent psychophysical study further showed that it is possible to perceive depth and rivalry simultaneously in one spatial location, as long as the components lie in different orientation or spatial frequency bands (Buckthought & Wilson, 2007). Specifically, this may be observed with binocular plaid patterns in which depth is perceived from the near-vertical components and rivalry from the oblique components. Subjects report that the percept of a rivalrous pattern is spatially superimposed on the tilted surface. This observation suggests that the neural substrates of depth and rivalry might be (partly) different. Alternatively, in addition to the processing involved in solving the correspondence problem for depth or rivalry, there may be a representation at the surface level that would facilitate the grouping of features and allow for more than one feature to be coded at a spatial location.

Both binocular depth and rivalry have been extensively studied in functional neuroimaging paradigms but never with comparable methods. An early neuroimaging study of depth analyzed a restricted set of cortical areas and found that area V3A showed the largest response to stereoscopic depth compared with areas V1, V2, V3, and MT+ (Backus, Fleet, Parker, & Heeger, 2001). A subsequent study reported that the greatest activation to stereoscopic depth occurred in occipito-parietal areas including V3A, V4d-topo,
V7 and caudal intraparietal sulcus (caudal parietal disparity region, CPDR), with lesser activation in areas V1, V2, V3, VP (Tsao et al., 2003). Neri, Bridge, and Heeger (2004) found that while dorsal areas (V3A, MT+, V7) showed a large activation to absolute disparity, ventral areas (hV4, V8/V4a) were activated by both absolute and relative disparities. Another study also found activation of ventral areas (i.e., lateral occipital complex) to disparity-defined objects (Gilaie-Dotan, Ullman, Kushnir, & Malach, 2002). Rutschmann and Greenlee (2004) found that areas V2, V3, and V3A but not V1 responded to stimuli with binocular disparity, and the BOLD signal increased with disparity only in dorsal areas of the occipito-parietal cortex. Finally, Bridge and Parker (2007) found that the largest activation for correlated disparity stimuli (contrasted with anticorrelated controls without depth) occurred in the lateral occipital lobe, including MT+ and areas LO-1 and LO-2 (areas within “kinetic occipital” area KO), as well as V3 and hV4. Hence, the most consistent sites of activation for depth across previous studies were occipito-parietal areas, including V3A, V7, V4d-topo, or other lateral occipital areas, and intraparietal sulcus (Bridge & Parker, 2007; Rutschmann & Greenlee, 2004; Tsao et al., 2003). Some studies also reported superior parietal lobe activation (Iwami et al., 2002; Nishida et al., 2001; Rutschmann & Greenlee, 2004).

In the case of binocular rivalry studies, analysis was often restricted to particular visual areas, making it difficult to draw comparisons to studies of depth. One study that performed a whole brain analysis used a bistable slant rivalry stimulus (in which bistability arises from incongruence between binocular disparity and monocular perspective cues; Brouwer, van Ee, & Schwarzbach, 2005). Transient activation relating to the perceptual alternations was found in caudal and anterior intraparietal sulcus, V4d-topo, and inferior to area MT+. The activation that correlated with alternations toward the depth percept was found in areas V3A, V7, V4d-topo, MT+, and lateral occipital complex. Another binocular rivalry study that performed a whole brain analysis (and used a more conventional binocular rivalry stimulus) found that a widespread network of frontoparietal areas was activated, which included superior and inferior parietal areas in common with depth studies, as well as fusiform gyrus and extrastriate visual cortex (Lumer, Friston, & Rees, 1998). A study that used diverse categories of objects reported activation in intraparietal sulcus, LOC, V3A, and V7 (Fang & He, 2005). One other whole brain study reported activation in middle frontal gyrus, superior parietal lobe, precentral gyrus, medial cingulate and paracingulate gyri, precuneus, and superior occipital gyrus (likely V2/V3 or V3A; Wilcke, O’Shea, & Watts, 2009). In binocular rivalry studies restricted to early visual areas, activation in V1, V2, V3, and V4 (Lee & Blake, 2002; Polonsky, Blake, Braun, & Heeger, 2000) or V1, V2, and V3 (Moradi & Heeger, 2009) was found to be approximately equal in magnitude.

It is difficult to compare the results for depth and rivalry across studies because of substantial stimulus and task differences. Most studies of depth used dynamic random dot stereograms (RDS; Backus et al., 2001; Bridge & Parker, 2007; Gilaie-Dotan et al., 2002; Neri et al., 2004; Rutschmann & Greenlee, 2004; Tsao et al., 2003), or gratings or line drawings (Gilaie-Dotan et al., 2002). Rivalry studies have used a variety of different stimuli, for example, gratings (Lee & Blake, 2002; Wilcke et al., 2009), high- and low-contrast gratings (moving or counter-phasing; Polonsky et al., 2000), faces versus houses (Lee & Blake, 2002), slant rivalry (Brouwer et al., 2005), and gratings versus faces (Lumer et al., 1998). The tasks used in depth and rivalry studies were also not comparable in terms of either the attentional demands or frequency of motor responses. To date, there has been no direct comparison of depth and rivalry with fMRI.

Here, we will employ binocular plaid stimuli in an fMRI paradigm, making it possible to compare conditions in which subjects perform either a depth or rivalry report task while viewing identical plaid patterns, precisely matched for retinal stimulation. Moreover, previous physiological studies provide clear evidence for a spatial frequency preference in certain visual areas, including, for example, high spatial frequencies in V1, V2, and V3 and low spatial frequencies in MT+ (Singh, Smith, & Greenlee, 2000; Tootell et al., 1998). Hence, viewing higher or lower spatial frequency components might be expected to modulate activation in different areas of the visual cortex. We hypothesized that placing the vertical and oblique components in different spatial frequency bands might increase the BOLD signal for both depth and rivalry and make it possible to find greater differences between these two conditions. We thus used spatial frequency conditions in which the near-vertical components were at a lower spatial frequency than the oblique components, or the reverse. A third condition in which the near-vertical and oblique components were at the same spatial frequency was used for comparison. We hypothesized that while depth and rivalry may be processed by similar cortical networks, it is possible for depth and rivalry to be perceived simultaneously because they are processed in different spatial frequency or orientation channels. For comparison, we included a condition in which subjects passively viewed depth alone or rivalry alone in sinusoidal grating stimuli. Finally, to further isolate the stages of binocular combination, a control task that required only the fusion of binocularly matched plaids but contained neither depth nor rivalry was used.

**Methods**

**Subjects**

An author (AB) and six other subjects who were naive as to the hypotheses of the study participated. The subjects
(which included three women) were university students or postdoctoral fellows. Six of the subjects participated in Session 1 including depth and rivalry scan type 1. Six subjects (of which five overlap with Session 1) participated in a second session including depth and rivalry scan type 2 and passive viewing scans. All subjects participated in psychophysical experiments and retinotopic mapping scans. All were right-handed and had normal or corrected-to-normal acuity and stereoacuity thresholds better than 30 s arc, measured using the Titmus stereo test (Stereo Optical, Chicago, IL). The subjects provided written informed consent and were remunerated for their time. The experiments were approved by the Research Ethics Board (REB) of McGill University (Protocol NEU-08-03).

Stimuli

In the present study, plaids in which it is possible to perceive depth and rivalry simultaneously were used (Figure 1). In the figure, the depth alone condition with near-vertical components only is shown in the first row. The rivalry alone condition with oblique components only is shown in the second row (± 30 deg). Adding these components together, the simultaneous depth and rivalry condition is shown in the third row. In three conditions, the spatial frequencies of the near-vertical and diagonal components were, respectively: 6.4 and 6.4 cpd (DHRH; depth high, rivalry high), 6.4 and 2.5 cpd (DHRL; depth high, rivalry low), or 2.5 and 6.4 cpd (DLRH; depth low, rivalry high).

Display

All stimuli were presented on a MacBook Pro Laptop (Intel Core 2 Duo) Macintosh computer with 1024 × 768 resolution and 120-Hz refresh rate with 8 bit/pixel grayscale, which was gamma-corrected using a color look-up table. Stimuli were generated and displayed using Matlab (2007b) and Psychtoolbox Version 3 (PTB-3) software. A Matrox (Dual Head 2Go Analogue Edition) splitter graphics card was used to create two channels for dichoptic displays. Dual LCD (InFocus LP 540) projectors and linear polarizers were used for dichoptic projection (Thompson, Farivar, Hansen, & Hess, 2008). The subjects wore linear polarizers with complementary polarization on their eyepieces. The stimuli were back-projected from the LCD projectors onto a polarization preserving screen at a viewing distance of 134 cm. The same display apparatus were used both for fMRI scan sessions and psychophysical sessions, including the same viewing distances. In the case of fMRI, the screen was placed at the rear end of the MR scanner bore and subjects viewed stimuli through a front-surfaced mirror attached to the head coil. Throughout the experiments, each stimulus was projected through an opaque circular aperture (3.8-deg diameter), which minimized any edge disparities. Viewed through the polarizers, the stimulus had a mean luminance of 30 cd/m² and peak luminance of 60 cd/m². All stimuli were displayed at 80% Michelson contrast. Prior to the main psychophysical tasks, upper fusion limits were measured for all subjects, in order to choose appropriate disparities.

![Figure 1](https://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/932793/ on 10/23/2018)
Psychophysical tests

Depth matching (method of adjustment)

In order to characterize the stimuli used in the fMRI scans, the depth perceived in the plaid patterns was compared to depth perceived in sinusoidal grating patterns. The subjects adjusted the depth of a sinusoidal grating patch (which will be referred to henceforth as the depth match stimulus) until it matched the apparent depth in the plaid pattern. The disparity in the plaid pattern was held constant at 5 deg (orientation disparity). The depth match stimulus had depth produced in the same manner as the plaid, using an orientation disparity. The plaid pattern was displayed in a circular patch (3.8-deg diameter) on the left side of the screen while the depth match stimulus was displayed in a circular patch on the right (3.8-deg diameter), with a central nonius fixation cross (0.3 deg × 0.25 deg). Each circular patch was 2.8 deg from the screen center. The stimulus patches were surrounded by a rectangle (9-deg width × 5-deg height) to aid fixation. The remainder of the screen was at mean luminance. Subjects had an unlimited amount of time to make the depth adjustment. Subjects were tested twice at each condition.

Depth perception (method of constant stimuli)

In the previous depth matching task, unlimited stimulus durations were used, which made it impossible to exclude eye movements. Hence, each subject was also tested with a depth perception test in which it was possible to use shorter stimulus durations. The display was similar to the depth matching task. As in the depth matching task, a sinusoidal grating and plaid stimulus were displayed in (3.8 deg) circular patches to the left and right of a central nonius fixation cross. As before, the plaid stimulus had an orientation disparity of 5 deg. However, in this case, the disparity of the sinusoidal grating was varied (randomly) from trial to trial (with disparities ranging from 3 to 10 deg in 0.1 increments), and the subject reported with a key press whether the sinusoidal grating had more or less depth (i.e., tilt) than the plaid stimulus. The stimulus duration was 180 ms. The psychometric function obtained for each subject was fit to a Weibull function (Weibull, 1951) to obtain the inflection point, which provided a measure of the perceived depth in the plaid patterns. Subjects were tested twice at each condition.

Alternation rates

Alternation rates were measured during binocular rivalry with the plaid or with left/right oblique grating stimuli. Subjects reported perceptual alternations continuously over 90-s trials. Subjects pressed one key when the left image predominated (over at least two-thirds of the stimulus) or another key when the other image predominated, using the same criterion. Subjects were tested twice at each condition.

Functional magnetic resonance imaging

Acquisition of fMRI data

All images were acquired using a 3T MR scanner (Siemens, Trio, Germany) at the Montreal Neurological Institute, with a 32-channel head coil (20 channels for retinotopic mapping). Functional whole brain images were acquired using a T2*-weighted gradient echo, echo-planar imaging sequence (38 slices, repetition time (TR) 2500 ms, echo time (TE) 30 ms, FOV 192, voxel size 3 × 3 × 3 mm). Functional images for retinotopic mapping were acquired with a T2*-weighted sequence, with slices oriented perpendicular to the calcarine sulcus (28 slices, repetition time (TR) 2000 ms, echo time (TE) 30 ms, FOV 128, voxel size 4 × 4 × 4 mm). Anatomical images were acquired by using a T1-weighted magnetization-prepared rapid gradient-echo (MP-RAGE) sequence optimized for contrast between gray and white matter (176 slices, repetition time (TR) 2300 ms, echo time (TE) 2.98 ms, FOV 256, voxel size 1 × 1 × 1 mm).

Binocular rivalry and depth scan type 1 (active condition)

A block design was used, with a block length of 20 s. There were separate depth scans and rivalry scans each with three repetitions of five conditions for a total length of 5 min. The rivalry scan conditions included a baseline blank condition, a rivalry task with plaids or sinusoidal grating patterns, as well as two other control versions that are not reported here. For the rivalry task, subjects used a button box to report when their dominant percept switched to that of a left oblique or right oblique component (following the procedure described above).

The depth scan conditions included a baseline blank condition, a depth task with plaids or sinusoidal grating patterns, as well as two other control versions. For the depth task, subjects reported whether the top or bottom of the stimulus pattern (plaid or sinusoidal grating) appeared to be tilted forward. The orientation disparity was 5 deg. The depth changed every 3 s during each block. In order to prevent the task from being too easy, there was one catch trial in every 20-s block in which the disparity did not switch to the opposite percept but was changed to a slightly different orientation disparity (7 deg). Each rivalry or depth scan included three repetitions of each type of block. Each subject participated in four depth scans and four rivalry scans per session. Each subject participated in three scan sessions on different days for each of the three spatial frequency conditions.

Binocular rivalry and depth scan type 2 (active condition)

A block design was used, with a block length of 20 s. In total, each scan included four types of blocks (rivalry task with plaids, depth task with plaids, orientation change
task with plaids, and the blank baseline condition). There were four repetitions of each of these stimulus conditions, for a total of 16 blocks. A small letter at the fixation point for 2 s was used to cue subjects to a change in condition and task. For the rivalry task, subjects used a button box to report when their dominant percept switched to that of a left or right component (following the procedure described above). For the depth task, subjects again reported whether the top or bottom of the plaid pattern appeared to be tilted forward, which changed every 3 s during each block (with catch trials as before). The orientation change task was a new control condition that required binocular fusion but evoked neither depth nor rivalry. For this task, subjects reported the orientation of the plaid pattern, in which the most horizontal component of the plaid could be oriented up toward the right or down. The plaids were matched in the two eyes (Figure 1, row 4). The orientation also changed every 3 s during a block, with an orientation change of 8 deg (with catch trials, as in the depth task). There were three types of these scans corresponding to the three spatial frequency conditions. Each subject participated in two runs for each of the three spatial frequency conditions, for a total of six runs in one session.

Passive viewing scans

In the passive viewing scans, a block design was used, with a block length of 20 s. Subjects passively viewed depth alone or rivalry alone in sinusoidal grating stimuli. In total, each scan included seven types of blocks (rivalry alone, depth alone, a blank baseline condition, and four other conditions that are not reported here). In the depth alone condition, either the top or bottom of the sinusoidal grating pattern appeared to be tilted forward, which changed every 3 s during each block. Three repetitions of each of these types of blocks were shown, for a total of 420 s.

Practice sessions

Before each scan session, each subject had a short practice session (20 min) outside the scanner but using the same equipment and stimulus display to become familiarized with the tasks.

Retinotopic mapping and the localization of MT+

Retinotopic mapping was carried out in a separate session. The stimuli for retinotopic mapping consisted of high-contrast, chromatic, flickering checkerboard patterns of two specific types. The rotating wedge stimulus swept through polar angles, and the expanding/contracting ring stimulus mapped eccentricity. There were four acquisition runs for each subject: eccentricity mapping (fovea to periphery and vice versa) and polar mapping (clockwise and counterclockwise). The polar mapping runs consisted of 8 cycles (full hemifield rotation of both wedges), lasting a total of 512 s. The eccentricity mapping runs consisted of 8 cycles of expanding or contracting rings, lasting a total of 512 s. Both stimuli compensated for the cortical magnification factor by increasing in size as they approached the periphery. The eccentricity stimuli traversed space with a logarithmic transformation. A central fixation marker was present at all times, and subjects were required to perform a task monitoring the orientation of the fixation marker to aid fixation stability. These retinotopic mapping runs were used to define foveal regions of interest for V1, V2, and V3, defined as the region of occipital pole activated in the central 2.9 deg of visual angle. Area V3A was also defined using these scans. In addition, subjects performed two runs of MT+ localization (256 s) consisting of eight 16-s epochs of low-contrast stationary rings compared to eight 16-s epochs of moving rings (Tootell et al., 1995).

Data analysis

We used the BrainVoyager QX analysis package, version 1.10.4.1250 (Brain Innovations, Maastricht, The Netherlands) for most functional data analyses as well as for the creation of inflated and flattened cortical representations. The freely available Freesurfer analysis package, version v4.5.0 (http://surfer.nmr.mgh.harvard.edu/), was found to be better for retinotopic mapping data analysis on the reconstructed inflated brain, and the identified retinotopic areas were transferred to BrainVoyager using anatomical landmarks.

The anatomical and functional scans were analyzed in BrainVoyager using a standard processing sequence, described as follows. The anatomical scans were used to create surface reconstructions of each subject’s cerebral cortex. The computed cortical surface representation was inflated and then flattened. Each subject’s reconstructed folded cortical representation was normalized to spherical coordinate space and aligned to a target brain (chosen as an individual subject) using cortex-based alignment. The cortex-based alignment is part of the standard processing sequence and is performed in order to obtain a good match between corresponding brain regions for the group-level statistical data analysis. Before analysis of the functional scans, the first two volumes of every scan were discarded. All functional images were subjected to a standard set of preprocessing steps: (1) motion correction; (2) slice timing correction; (3) linear trend removal using a high-pass filter; (4) transformation of the functional data into Talairach coordinate space (Talairach & Tournoux, 1988); and (5) coregistration to anatomical images. A voxel-by-voxel, fixed effects general linear model (GLM) was used for analysis, with all of the stimulus conditions as predictors (i.e., in scan type 1, plaids and sine-wave patterns with either rivalry or depth task and baseline; in scan type 2, rivalry task, depth task, orientation task, and baseline; in passive viewing scans, rivalry alone, depth alone, and baseline). The statistical analysis was carried out by
performing relevant subtractions between these different conditions. The functional results were then viewed on an individual’s cortical surface, producing maps of statistical significance (t-tests with a false discovery rate of $p < 0.05$). In addition, we separately analyzed the BOLD signal changes within regions of interest (retinotopic areas, MT+), using a fixed effects GLM analysis.

**Results: Psychophysics**

Figure 2 shows the results of psychophysical studies carried out before the fMRI sessions. The results of the depth adjustment task show that it was possible to perceive depth in the plaid patterns in all three spatial frequency conditions, although the perceived depth was reduced relative to the depth in sinusoidal gratings (Figure 2A). This difference was only significant in the D$_{H}$R$_{H}$ condition ($t(df = 5) = -4.49, p < 0.05$ comparing plaids and gratings). The results of the depth perception test using the method of constant stimuli at short durations (180 ms) were very similar (Figure 2B). The perceived depth did not depend upon stimulus duration, making it unlikely that the effect of vergence eye movements was an important factor in the overall pattern of results. It should also be noted that subjects consistently made depth judgments with appropriate disparity signs, confirming that depth aliasing did not occur.

The results of the rivalry task demonstrate that it was also possible to perceive rivalry with plaids in all three of the spatial frequency conditions, although the dominance durations with the plaids were, in all cases, significantly slower than those with the sinusoidal grating patterns (D$_{L}$R$_{H}$, $t(df = 5) = 5.12, p < 0.05$; D$_{H}$R$_{H}$, $t(df = 5) = 2.83, p < 0.05$; D$_{H}$R$_{L}$, $t(df = 5) = 4.30, p < 0.05$). In particular, we note that the alternation rates for the plaids in the three spatial frequency conditions did not differ (one-way repeated measures ANOVA, $F(1, 5) = 3.11, p > 0.05$). The results from the key presses data obtained during the depth and rivalry scans were comparable, indicating the alternation rates for the plaids in the three spatial frequency conditions did not differ (one-way repeated measures ANOVA, $F(1, 5) = 2.05, p > 0.05$), and the depth judgments were at least 75% correct for each scan run. These results validate our use of the plaid stimuli, in which it was possible to perceive both depth and rivalry, in the fMRI paradigm.

**Results: fMRI comparison of depth and rivalry**

Figure 3 shows brain activation maps averaged across the six subjects, on inflated brains, with lateral, medial, and ventral views (left and right hemispheres). Figure 3A shows the results for passive viewing of depth or rivalry for sinusoidal gratings at the high spatial frequency compared to baseline. The activation for depth alone is shown in red. The activation for rivalry alone is shown in yellow. The area of overlap between depth and rivalry is shown in orange. The network of activated cortical areas was remarkably similar for rivalry compared to depth, as indicated by the large area of overlap in orange. The areas included occipital pole, lateral occipital cortex and ventral temporal areas, superior and inferior parietal cortices, including intraparietal sulcus and areas close to temporo-parietal junction, as well as middle and inferior frontal cortex, supplementary motor area, and premotor cortex.

Figure 3B shows the results for the active depth task and rivalry task for plaids compared to baseline (the results for the active task for gratings are not shown because of their similarity to those shown here). In this condition, the spatial frequencies for the near-vertical and oblique components were the same (D$_{H}$R$_{H}$). Analogous to Figure 3A, the network of activated cortical areas was
remarkably similar for the rivalry task compared to the depth task, as indicated by the large area of overlap in orange. The overall pattern was similar to that for the passive condition and included similar areas of occipital, parietal, and ventral temporal cortices. However, it can be seen that the activation was generally higher for the active condition, particularly in frontal areas. In Figure 3B, the peaks of activation from the binocular rivalry study by Lumer et al. (1998) have been superimposed with blue dots on the brain activation maps. The network of activated cortical areas in that study including frontoparietal activation was very similar. However, we found activation in these areas for both depth and rivalry. The areas of activation included superior (SP) and inferior parietal (IP) cortices, middle (MF) and inferior frontal (IF) cortices, supplementary motor area (SMA), premotor cortex (PM), occipital pole (OP), lateral occipital (LO), and ventral temporal (VT) areas. The peaks of activation from the binocular rivalry study by Lumer et al. (1998) have been superimposed with blue dots on the brain activation maps.

Thus far, a number of similarities have been highlighted between the areas activated by the depth and rivalry tasks with the plaid patterns. In contrast, Figure 4A shows explicitly the differences between the two, with the subtraction of the depth task minus rivalry task for identical plaid patterns. Again, this is for plaid with oblique and near-vertical components at the same spatial frequency (DHRH). The brain activation maps, averaged across the six subjects, are shown on flattened representations of the posterior left and right brains. There was greater activation for rivalry in lateral occipital cortex (LO), retrosplenial cortex (RS), posterior superior temporal sulcus (STS), and ventral temporal cortex (right hemisphere only), as well as a calcaine area (foveal V1; left hemisphere only). We note in passing that the results are consistent with both alternatives in the debate over whether lower or higher level areas are involved in competition in rivalry, as there was activation in calcarine as well as ventral temporal areas (Lee & Blake, 1999; Leopold & Logothetis, 1996, 1999). It is also worth noting again that the rate of button presses did not differ between depth and rivalry conditions. Moreover, even the distributions of the time intervals between key presses for the
rivalry and depth tasks are not significantly different (Kolmogorov–Smirnov, \( Z = 1.23, p < 0.05 \)), so this is unlikely to be a confound.

**Results: fMRI spatial frequency dependence**

The next issue we investigated was the spatial frequency dependence in retinotopic visual areas for depth or rivalry. We hypothesized that when the depth and rivalry components are in different spatial frequency bands, this might increase the fMRI activation, making it possible to find greater differences for depth and rivalry. Furthermore, previous studies provided an expectation for a spatial frequency bias, including, for example, a bias for high spatial frequencies in areas V1, V2, and V3 and low spatial frequencies in MT+ (Singh et al., 2000; Tootell et al., 1998).

Figure 4B shows the region-of-interest analysis for the depth task minus rivalry task, for plaid patterns. The results are shown for two conditions: \( \text{D}_H \text{R}_L \) (depth high, rivalry low) and \( \text{D}_L \text{R}_H \) (depth low, rivalry high). In the \( \text{D}_H \text{R}_L \) condition, early visual areas V1, V2, and V3 showed a preference for depth when it had the higher spatial frequency. However, in the \( \text{D}_L \text{R}_H \) condition, V1, V2, and V3 showed a preference for rivalry when it had the higher spatial frequency. This was statistically significant in all cases except for V2 in the right hemisphere, for \( \text{D}_L \text{R}_H \). That is, V1, V2, and V3 showed a preference for the higher spatial frequency component, regardless of task. MT+ showed the opposite pattern, showing a preference for the lower spatial frequency component, regardless of whether it was a depth or rivalry task. In the third condition where the components were at the same spatial frequency (\( \text{D}_H \text{R}_H \)), the differences between depth and rivalry were generally not statistically significant (the only exceptions were a bias for rivalry in left hemisphere V1 and a lateral occipital region, LOR; see definition below in the next section and Supplementary Figure 1). This provided strong support for the hypothesis that it was possible to observe greater differences between the activation for depth and rivalry in these retinotopic visual areas when the components were in different spatial frequency bands. Finally, the lateral occipital region (LOR) showed a different pattern of results, showing a preference for rivalry over depth at all spatial frequencies but only in the left hemisphere.

**Results: fMRI comparison to control orientation task**

In the next manipulation, we included, as a control, an orientation change task, which had similar stimulus features to the depth and rivalry tasks. In this case, the
subject had to indicate with a key press which way the image was rotated. The orientation change condition required binocular fusion of matched features but evoked neither depth nor rivalry, serving to isolate those stages of binocular combination. This task was also matched to the depth and rivalry tasks in terms of the number of stimulus changes (which occurred every 3 s) and key presses. In Figure 5A, the activation for the depth task minus the orientation change task is shown in red. The activation for the rivalry task minus the orientation change task is shown in yellow. The overlap between these areas is shown in orange. The results are shown for identical plaid patterns in which the components were at the matched high spatial frequency (DhRm). Color scale ranges from $t = 2.77$ to $t = 8.00$ for either yellow or red (FDR < 0.05). The activation is shown on the flattened brain of an individual subject, with superimposed contours showing the borders of retinotopic visual areas (V1, V2, V3, V3A). Note that MT+ is labeled only with a circle. A lateral occipital region (LOR) is highlighted in these subtractions. Superior (SP) parietal cortex, retinotopic visual areas, and retrosplenial (RS) areas were also activated but only in the left hemisphere. (B) Corresponding region-of-interest analysis, shown for left and right hemispheres. Results show the subtractions, depth task minus orientation change task, and rivalry task minus orientation change task, in percentage signal change. Both V3A and LOR showed significant activation to either depth or rivalry, particularly in the right hemisphere. In early visual areas, the activation was generally much lower but was significant bilaterally in V3 and in V1 and V2 in the left hemisphere only.

These results are not shown because of the similarity to the ones shown for the intersection. Since the lateral occipital area is activated by either depth or rivalry compared to the orientation change task, it is consistent with the idea of a role in binocular grouping or segmentation common to both forms of perceptual processing. However, in the left hemisphere, the significantly higher activation to rivalry than depth does not satisfy a strictly cue-invariant model. In comparison, there were lower activation levels overall in the early visual areas, reaching statistical significance in V3 (bilaterally), as well as V1 and V2, in the left hemisphere only.

**Discussion**

In conclusion, the whole brain network of activated cortical areas was remarkably similar for the rivalry task compared to the depth task when subjects viewed identical plaid patterns. However, when subjects viewed plaids, some areas were activated more during the depth than the rivalry task independent of spatial frequency, such as superior and inferior parietal cortices (including TPJ and intraparietal sulcus). In contrast, the posterior STS, retrosplenial, calcarine, lateral occipital, and ventral temporal areas were activated more for the rivalry task independent of spatial frequency. In other cortical regions,
the bias was for the higher spatial frequency component (e.g., early visual areas V1, V2, or V3) or for the lower spatial frequency component (e.g., MT+), regardless of task. A lateral occipital region (LOR) and area V3A were active for either depth or rivalry compared to a control task that required fusion but no surface grouping/segmentation.

The areas of activation selective for depth (Figure 4) included superior and inferior parietal lobes, which included intraparietal sulcus. The superior parietal lobe has already been associated with stereoscopic depth processing in previous fMRI studies (Iwami et al., 2002; Nishida et al., 2001; Rutschmann & Greenlee, 2004). A large number of previous fMRI studies have found that the intraparietal sulcus is selective for stereoscopic depth (Chandrasekaran, Canon, Dahmen, Kourtzi, & Welchman, 2007; Durand et al., 2007; Durand, Peeters, Norman, Todd, & Orban, 2009; Fang & He, 2005; Georgieva, Peeters, Kolster, Todd, & Orban, 2009; Orban et al., 2006; Shikata et al., 2001, 2003, 2008; Tsao et al., 2003). Based upon the Talairach coordinates and anatomical landmarks, it is clear that the superior parietal area of activation (Figure 4) in the current study includes several areas in the intraparietal sulcus (IPS) involved in 3D shape perception from disparity, which have been described in previous studies, dorsal IPS anterior (DIPSA) and dorsal IPS medial (DIPSM) (e.g., DIPSA: \(x = -30, y = -46, z = 45; x = 33, y = -45, z = 46\); DIPSM: \(x = -21, y = -58, z = 48; x = 21, y = -56, z = 45\); Chandrasekaran et al., 2007), as well as caudal IPS (\(x = -21, y = -64, z = 42; x = 23, y = -69, z = 40\); Rutschmann & Greenlee, 2004). These parietal regions extract 3D shape representations that can support motor functions, such as grasping hand movements or saccadic eye movements toward objects (Shikata et al., 2008). Regions DIPSM and DIPSA likely correspond to LIP and AIP in the monkey and process depth information in order to make eye or hand movements, respectively. Areas LIP and AIP are sensitive to depth structure (i.e., spatial variations in depth along surfaces arising from disparity) but not position in depth (Durand et al., 2009). The caudal IPS likely corresponds to the CIP in the monkey, in which neurons are sensitive to disparity and selective for orientation in depth of surfaces (Orban et al., 2006). It has been proposed based on single neuron studies that a processing pathway from caudal IPS to DIPSA progressively refines the 3D shape representation necessary for hand grasping movements (Orban et al., 2006). In human fMRI studies, both DIPSA and caudal IPS have shown activation related to spatial discrimination of surface orientation, but caudal IPS activation was more tightly linked to the discrimination task and did not change during grasping movements (Shikata et al., 2001, 2003). In contrast, DIPSA was activated during grasping hand movements toward the surface orientation. DIPSA (or AIP) is involved in coding object shape and orientation in three-dimensional space to match hand movements with 3D objects (Shikata et al., 2001, 2003). These parietal areas (DIPSA, DIPSM, and caudal IPS) are also more strongly activated by curved surfaces than tilted surfaces, showing a full representation of a range of different 3D shapes from disparity (Georgieva et al., 2009). Furthermore, these parietal areas appear to be involved in cue-invariant processing of 3D shape, including processing of monocular cues to depth (e.g., texture gradients, perspective, motion, shading; Durand et al., 2009; Orban et al., 2006; Shikata et al., 2001, 2003).

Areas selective for rivalry included the posterior superior temporal sulcus (STS) and retrosplenial cortex. The STS has not generally been found to be activated during rivalry alternations in previous fMRI studies. Single neuron responses in the macaque STS have been found to be strongly modulated by rivalry, although it is not clear that this area is a direct homologue to human STS (Leopold & Logothetis, 1999; Logothetis & Schall, 1989; Sheinberg & Logothetis, 1997). We know that neurons in area TE near rostral STS in monkeys do show depth selectivity (Orban et al., 2006). In the case of retrosplenial cortex, it is considered part of a network of brain regions subserving different cognitive functions, including episodic memory, spatial navigation, and a number of other tasks involving spatial transformations (Vann, Aggleton, & Maguire, 2009). It has been shown to be involved in tasks in which it was necessary to (endogenously) switch between different sets of conflicting spatial cues and pay attention to just one cue. In the current study, it may have been selective for rivalry since the rivalry task involved the monitoring of conflict between two oriented gratings. It is interesting to note that responses in the neighboring precuneus region are affected by visual information accumulated over long time scales (~12 s) and the predictability of a sequence of visual stimuli, suggesting a role in the accumulation of visual input over time (Hasson, Yang, Vallines, Heeger, & Rubin, 2008). In addition, activation in this region was reported during rivalry alternations in one previous study (Wicke et al., 2009).

Selectivity for rivalry over depth was also found in foveal V1 and ventral temporal cortex, areas likely involved in neural competition. Generally, activation in V1, V2, V3, and ventral temporal cortex has been prominently reported for rivalry, particularly in studies using gratings as stimuli (Lee & Blake, 2002; Moradi & Heeger, 2009; Polonsky et al., 2000). The selectivity for rivalry over depth in V1 leads us to consider whether there could be any role of stimulus orientation per se in this effect. In our experimental design, the depth stimuli were always vertically oriented, while the rivalry stimuli were always obliquely oriented. Some studies have shown greater activation for cardinal orientations in V1 (Furmanski & Engel, 2000). However, others found a bias for oblique orientations (Mannion, McDonald, & Clifford, 2010), so the results are equivocal. If there is, in fact, a reliable bias for vertical orientations in V1, then it may contribute to our result, although it remains conceptually interesting to consider that a bias for vertical orientations may have developed in order to support depth perception.
A final point with regard to the direct comparison of depth and rivalry tasks is that we cannot completely rule out a differential role of visual attention in the two tasks. However, in addition to the matched key press performance reported here, additional psychophysical experiments were conducted in which these two tasks were performed concurrently. The symmetrical effects on performance supported our suggestion that the two tasks did not differ in task difficulty (Braun & Julesz, 1998).

In another manipulation, an orientation change task was used as a control task to further isolate binocular perceptual processing, by requiring fusion but eliciting neither depth nor rivalry (Figure 5). When the orientation change task was subtracted from either the depth or rivalry task, a lateral occipital area was highlighted, as well as an area dorsal to V3A, which likely corresponds to area V7 or ventral intraparietal sulcus (VIPS) based upon the Talairach coordinates, \( x = -22, y = -80, z = 27; x = 26, y = -80, z = 25 \) (Levy, Schluppeck, Heeger, & Glimcher, 2007; Silver, Ress, & Heeger, 2005). The LOR region of interest, which was a smaller area derived from the intersection of the depth minus orientation change and rivalry minus orientation change subtractions, consisted of just the area corresponding to V7/VIPS (Levy et al., 2007; Silver et al., 2005). V7/VIPS is an area sensitive to depth structure, as well as depth position, that is also sensitive to motion, 3D structure from motion, and 2D shape (Durand et al., 2009; Georgieva et al., 2009). In previous fMRI studies, this area has also been described as showing activation strongly correlated with the depth amplitude (defined by disparity) perceived by subjects (Georgieva et al., 2009).

Beyond V7/VIPS, the lateral occipital region highlighted in Figure 5 has been referred to by a wide variety of nomenclatures by different studies. The activation we report also likely included kinetic occipital area (KO) as the Talairach coordinates for the most posterior and ventral part, \( x = 30, y = 85, z = 3; x = -32, y = 90, z = 3 \), are close to published coordinates for this area (Dupont et al., 1997; Kononova et al., 2003; Tyler, Likova, Kontsevich, & Wade, 2006; Van Oostende, Sunaert, Van Hecke, Marchal, & Orban, 1997; Zeke, Perry, & Bartels, 2003). Moreover, the KO activation likely includes both LO-1 and LO-2, as previous studies showed that these areas are included in KO (Larsson & Heeger, 2006). Finally, this same lateral occipital region is centered upon the area also described in previous studies as V4d-topo (\( x = 35, y = -80, z = 8; x = -36, y = -80, z = 8 \); Malach et al., 1995; Tootell & Hadjikhani, 2001; Tsao et al., 2003).

Regardless of terminology, it is clear from a number of fMRI studies that these regions of the lateral occipital cortex have considerable sensitivity to depth and surface structure (Bridge & Parker, 2007; Chandrasekaran et al., 2007; Kourtzi, Erb, Grodd, & Bulthoff, 2003; Mendola, Dale, Fischl, Liu, & Tootell, 1999; Tsao et al., 2003; Tyler et al., 2006; Welchman, Deubelius, Conrad, Bulthoff, & Kourtzi, 2005). In particular, the kinetic occipital (KO) region is responsive to motion-defined borders and illusory contours of the figure–ground type (Mendola et al., 1999; Van Oostende et al., 1997; Zeki et al., 2003), and several studies have found that KO is responsive to depth structure (Mendola et al., 1999; Tsao et al., 2003). Tyler et al. (2006) subsequently confirmed that the KO region is specialized for encoding depth structure, whether derived from motion cues, figure–ground segregation, or binocular disparity. A novel result from the current study is the association of binocular rivalry with this region as well. This is congruent with previous studies that concluded that rivalry occurs not just between individual elements but between grouped features or surfaces (Alais & Blake, 1999; Kovacs, Papathomas, Yang, & Feher, 1996). Thus, these regions are involved in segmentation and grouping computations in common with both depth and rivalry, perhaps distinct from the parietal regions involved in metrical depth for the purpose of action. A caveat to the interpretation of a cue-invariant representation is that the activation in these areas was significantly higher for rivalry than depth, in the left hemisphere only. Nevertheless, these findings are also consistent with a number of cue combination studies showing that other cues are combined with disparity in the perception of surfaces (e.g., Todd, Tittle, & Norman, 1995). In the case of conflict, other cues may easily override disparity, and the percept from disparity alone is not robust. Hence, we would expect that the neural representation of surfaces would include binocular features along with a number of other surface attributes.

Overall, our results are comparable to previous fMRI studies of depth and rivalry. The most consistent areas of activation for depth across previous studies are occipito-parietal areas, including V3A, V7, V4d-topo, or other lateral occipital areas, and intraparietal sulcus (Bridge & Parker, 2007; Brouwer et al., 2005; Rutschmann & Greenlee, 2004; Tsao et al., 2003). In addition to these areas, we also had activation in the superior parietal lobe, which has been found in some previous studies (Iwami et al., 2002; Nishida et al., 2001; Rutschmann & Greenlee, 2004) but not others (Brouwer et al., 2005; Tsao et al., 2003). Furthermore, we also had activation in early visual areas (V1, V2 and V3), which has not always been prominently reported in depth studies (Bridge & Parker, 2007; Brouwer et al., 2005; Iwami et al., 2002; Nishida et al., 2001; Rutschmann & Greenlee, 2004; Tsao et al., 2003). A few studies of depth also reported activation in more ventral areas (hV4, V8/V4a, or ventral lateral occipital areas), but this likely occurred because of the different methods used, emphasizing relative over absolute disparity processing (Neri et al., 2004) or object perception (Gilaie-Dotan et al., 2002). Our results are also generally consistent with rivalry studies, although there are few previous studies using comparable methods. One binocular rivalry study reported activation in superior and inferior parietal cortices as well as fusiform gyrus and extrastriate visual cortex (Lumer et al., 1998). Another rivalry study found activation in
caudal and anterior intraparietal sulcus, V4d-topo, and inferior to area MT+ (Brouwer et al., 2005). A third rivalry study reported activation in intraparietal sulcus, lateral occipital complex, V3A, and V7 (Fang & He, 2005). Our study shows activation in all of the areas reported in these studies, but we also had retrosplenial and STS activation. We also additionally noted activation of early visual areas (V1, V2, V3, and V3A), which has been reported in some previous rivalry studies (Lee & Blake, 2002; Moradi & Heeger, 2009; Polonsky et al., 2000). Our study goes beyond these studies in reporting V3A, V7/VIPS and LOR (lateral occipital) activation for rivalry when the orientation task was subtracted (rather than using a replay control). In summary, our results are generally consistent with previous fMRI studies of depth or rivalry and suggest that differences between published results are driven largely by the different methods employed.

A novel aspect of our study was restricting depth and rivalry to specific spatial frequencies. We hypothesized that when the depth and rivalry components are in different spatial frequency bands, this may make it possible to find greater differences in fMRI activation for depth and rivalry. Moreover, the fact that visual areas differ in terms of spatial frequency tuning provides a substrate for selective processing (Singh et al., 2000; Tootell et al., 1998). Indeed, the results confirmed that there was a bias for the low spatial frequency component in area MT+ and high spatial frequency component in areas V1, V2, and V3, regardless of task. Either depth or rivalry could strongly activate these areas, suggesting that the same visual areas could subserve either depth or rivalry perceptual processing. When the depth and rivalry components were at the matched high spatial frequency, the activation for depth or rivalry generally did not differ in retinotopic areas. Both our fMRI and psychophysical results lead us to conclude that depth and rivalry are processed in a similar network of cortical areas and are perceived simultaneously by coexisting in different spatial frequency or orientation channels. Hence, this is one of only a few studies using physiological methods to probe parallel spatial channels in the human visual system, which are tuned to spatial frequency and orientation bands (Burr & Morrone, 1987; Campbell & Maffei, 1970; Regan, 1983; Suter et al., 1994). Furthermore, the consensus of a spatial frequency bandwidth of about one octave in past studies was also found previously for these plaid stimuli (Buckthought & Wilson, 2007). These results should be incorporated into future models of binocular vision.

Current models of binocular vision would not predict that it is possible to perceive simultaneous depth and rivalry, as the binocular false matches and suppression from rivalry would prevent binocular matching for depth perception to occur (Hayashi et al., 2004). One possible explanation is that the neural substrates for the representation of surfaces are distinct from that in which the correspondence problem for depth or rivalry is solved. While early visual areas (V1, V2, V3) are likely involved in solving the correspondence problem, possible candidate areas for surface representation include the lateral occipital areas responsive to either depth or rivalry. Since these areas showed greater activation for either rivalry or depth than the orientation change condition, they showed selectivity for binocular features and could potentially resolve the contradictory surface signals from the rivalrous and depth components (in accordance with subjects’ reports that the rivalry appeared to be superimposed on the surface tilted in depth).

Another important suggestion from a number of investigators is the concept of coarse-to-fine interactions across spatial frequency channels, which would imply that it should not be equally possible to perceive depth and rivalry when the component high and low spatial frequencies have been swapped (e.g., Rohaly & Wilson, 1993). In particular, it should be easier to perceive depth when that component is the lower spatial frequency and more difficult when depth is conveyed by the higher spatial frequency because of the greater interference by rivalry at a lower spatial frequency. Instead, both our psychophysical and fMRI results indicated that (with carefully chosen orientations and spatial frequencies) it was possible to perceive robust depth and rivalry in all of the spatial frequency conditions. Thus, coarse-to-fine interactions across spatial frequency channels seem unlikely to be obligatory across the entire stimulus space.

In conclusion, this is the only study to date to directly compare depth and rivalry with matched stimulus features and task conditions and goes even further in using identical retinal stimulation. We also compared both depth and rivalry to a comparable fusion-only control for the first time. Finally, this study is notable in providing one of the few examples in which the function and existence of parallel spatial frequency channels has been addressed using a physiological technique. These plaid stimuli in which both depth and rivalry could be perceived provided a particularly attractive method for shifting the activity of visual areas toward low or high spatial frequencies while keeping the retinal stimulus constant.

Acknowledgments

This work was supported by NSERC and NIH R01 EY015219 Grants, as well as an LOF Grant from the Canadian Foundation for Innovation. We thank Robert F. Hess for providing MR compatible equipment for dichoptic stimulation and anonymous referees for comments on the manuscript.

Commercial relationships: none.
Corresponding author: Athena Buckthought.
Email: athenabuck1@gmail.com.
Address: McGill Vision Research Unit, Royal Victoria Hospital, 687 Pine Ave. W. Rm H4-14, Montreal, Quebec H3A 1A1, Canada.
References


Kourtzi, Z., Erb, M., Grodd, W., & Bulthoff, H. H. (2003). Representation of the perceived 3-D object shape in
the human lateral occipital complex. *Cerebral Cortex, 13*, 911–920. [PubMed] [Article]


Singh, K. D., Smith, A. T., & Greenlee, M. W. (2000). Spatiotemporal frequency and direction sensitivities...
of human visual areas measured using fMRI. *Neuroimage, 12*, 550–564. [PubMed]


