Neural Evidence for Competition-Mediated Suppression in the Perception of a Single Object

Laura Cacciamani[1], Paige E. Scalf[2], & Mary A. Peterson[1,3]

[1]Department of Psychology, University of Arizona

[2]Department of Psychology, Durham University

[3]Cognitive Science Program, University of Arizona

Corresponding Author:

Laura Cacciamani

Department of Psychology

1503 E. University Blvd.

University of Arizona

Tucson, AZ 85721

Email: lcacciamani@gmail.com

Phone: 520-621-5543

Abstract

Multiple objects compete for representation in visual cortex. Competition may also underlie the perception of a single object. Computational models implement object perception as competition between units on opposite sides of a border. The border is assigned to the winning side, which is perceived as an object (or "figure"), whereas the other side is perceived as a shapeless ground. Behavioral experiments suggest that the ground is inhibited to a degree that depends on the extent to which it competed for object status, and that this inhibition is relayed to low-level brain areas. Here, we used fMRI to assess activation for ground regions of task-irrelevant novel silhouettes presented in the left or right visual field (LVF or RVF) while participants performed a difficult task at fixation. Silhouettes were designed so that the insides would win the competition for object status. The outsides (grounds) suggested portions of familiar objects in half of the silhouettes and novel objects in the other half. Because matches to object memories affect the competition, these two types of silhouettes operationalized, respectively, *high competition and low competition* from the grounds. The results showed that activation corresponding to ground regions was reduced for *high-* vs. *low-competition* silhouettes in V4, where receptive fields (RFs) are large enough to encompass the familiar objects in the grounds, and in V1/V2, where RFs are much smaller. These results support a theory of object perception involving competition-mediated ground suppression and feedback from higher to lower levels. This pattern of results was observed in the left hemisphere (RVF), but not in the right hemisphere (LVF). One explanation of the lateralized findings is that task-irrelevant silhouettes in the RVF captured attention, allowing us to observe these effects, whereas those in the LVF did not. Experiment 2 provided preliminary behavioral evidence consistent with this possibility.

Keywords: object perception, figure-ground segregation, competition, ground suppression, feedback

1. Introduction

The ability to perceive objects is essential to our interaction with the world. Although it seems effortless and immediate, object perception is a computationally demanding task that requires extensive neural processing. Prior research has focused on understanding this neural processing, yet object perception is still not entirely understood. The current study aims to shed light on the neural mechanisms involved in the perception of a single object.

Previous work indicates that competitive suppressive interactions occur among representations of multiple objects shown simultaneously such that the firing rate in stimulated brain areas is lower than the sum of the responses to each item presented individually (Luck, Chelazzi, Hillyard, & Desimone, 1997; MacEvoy & Epstein, 2009; Miller, Gochin, & Gross, 1993; Reddy & Kanwisher, 2007; Reynolds, Chelazzi, & Desimone, 1999; Rolls & Tovee, 1995; Snowden, Treue, Erickson, & Andersen, 1991; Zoccolan, Cox, & DiCarlo, 2005). The reduced response is thought to result because multiple simultaneously presented items engage in inhibitory competition for representation (Moran & Desimone, 1985; Reynolds et al., 1999). Competition-mediated suppression between multiple objects has been observed in visual areas as low as V1/V2 and as high as the inferior temporal cortex (IT) via single-cell recording in monkeys (Miller et al., 1993; Reynolds et al., 1999) and functional magnetic resonance imaging (fMRI) in humans (Beck & Kastner, 2005; 2007; Kastner, De Weerd, Desimone, & Ungerleider, 1998; Scalf & Beck, 2010; see Beck & Kastner, 2009, for a review).

It has been proposed that suppressive competition underlies the perception of a single object when a border shared by two visual field regions is perceived as a boundary for the region on one side but not the other. The side to which the border belongs is perceived as a shaped object (or *figure)*; the other side is perceived as a locally shapeless *ground* continuing behind the object. Computational theories of figure-ground perception implement suppressive competition between low-level edge-units, feature-units and/or image-based properties such as convexity, symmetry, and small area detected on opposite sides of a shared border (e.g., Craft, Schutze, Niebur, & von der Heydt, 2007; Grossberg, 1994; Kienker, Sejnowski, & Hinton, 1986; Kogo, Strecha, Van Gool, & Wagemans, 2010; Roelfsema, Lamme, Spekreijse, & Bosch, 2002; Sejnowski & Hinton, 1987; Vecera & O'Reilly, 1998). In these models, units on the winning side (the object or "figure" side) are ultimately enhanced relative to units on the losing side (the groundside). And indeed, neural evidence shows that responses to figures are enhanced relative to responses to grounds (Roelfsema et al., 2002).

Behavioral experiments using stimuli like those in Figure 1 support the hypothesis that perceiving an object on one side of a border entails suppressive competition. These stimuli were designed so that figural/object status would be highly likely to be assigned to the inside of the bounded region (e.g., the insides were closed, symmetric, and small in area). These figures were novel objects. The amount of competition for object status was varied across two subsets of these stimuli: the borders of half the stimuli suggested a portion of a familiar, real-world object on the outside—the side that was ultimately perceived as the ground (Figure 1B)—whereas the other half did not (Figure 1A). Given that familiar configuration plays a role in figure assignment (see Peterson and Skow-Grant, 2003, for review), these were *high-competition* and l*ow-competition* stimuli, respectively. Consistent with the hypothesis that suppressive competition is one mechanism of object perception, Sanguinetti, Trujillo, Schnyer, Allen, and Peterson (2014) found greater activity in the alpha band of the electroencephalogram (EEG) recordings, indicative of greater inhibition, when participants viewed *high-* rather than *low-competition* stimuli and reported perceiving them as novel objects. Additionally, Peterson and Skow (2008) showed that after participants perceived the insides of *high-competition* stimuli as figures, responses to familiar objects similar to those suggested on the outside were inhibited, indicating that properties of objects suggested on the side of a border that loses the competition are suppressed. Moreover, Salvagio, Cacciamani, and Peterson (2012) showed that orientation discrimination performance was worse for targets presented on the on the groundside of h*igh-competition* than *low-competition* stimuli. These findings go beyond the models of suppressive competition in showing that the location of the losing competitor for object status is suppressed, and that greater suppression is applied to grounds that compete more for object status—a phenomenon we refer to as *competition-mediated ground suppression.*

< Insert Figure 1 about here >

There is some neural evidence for ground suppression. For instance, Likova and Tyler (2008) showed that neural activity is suppressed for regions perceived as grounds compared to regions without figure-ground structure (cf., Strother, Lavell, & Vilis, 2012). Here, we extend those previous studies by using two types of stimuli that are matched for figure-ground structure (and for low-level features – see Section 2.1.2) to investigate whether more suppression is applied to grounds that compete more for object status in virtue of suggesting a familiar configuration (*high-competition* silhouettes) than to grounds that compete less for object status (*low-competition* silhouettes). Evidence that the groundside of a border is suppressed more when it competes more for object status would support the hypothesis that suppressive competition underlies the perception of a single object.

We sought evidence of greater competition-mediated suppression applied to the grounds of *high-competition* than l*ow-competition* silhouettes in visual areas V4, V2, and V1. There are multiple reasons to expect these areas to be implicated in competition-mediated ground suppression. First, these are areas in which evidence of multiple-object suppressive competition has been found (see Beck & Kastner, 2009 for a review). Second, V4 receptive fields (RFs) are large enough to encompass the full vertical extent of the silhouettes (4^{o}), and hence, the portion of the familiar configuration suggested on the outside of the *high-competition* silhouettes. We expected that the differential competition for object status in the two types of silhouettes and therefore any differences in ground suppression would be high in V4. Third, although RFs in V2 and V1 are much smaller (1-2^{o}; Bles, Schwarzbach, De Weerd, Goebel, & Jansma, 2006) and cannot encompass more than 1-2 parts of the familiar object suggested on the groundside of the *high-competition* silhouettes, others have observed evidence for ground suppression in V1/V2 (Likova & Tyler, 2008; Strother et al., 2012). These authors attributed their V1/V2 results to feedback, because the pattern of activity in V1/V2 mimicked that in higher-level regions where stimulus differences between their conditions were represented (MT and LO for Likova & Tyler, 2008, and Strother et al., 2012, respectively). We applied a similar logic here investigating whether evidence of competition-mediated ground suppression observed in V4 is also observed in V2 and V1. If it is, that will be consistent with the hypothesis that feedback from suppressive competition at higher levels modulates responses in lower-level brain regions. Consistent with this feedback hypothesis, Salvagio et al. (2012) reported behavioral evidence of competition-mediated suppression for targets shown on grounds but not for targets shown on figures, even though the distance between targets on figures vs. grounds was less than 1^{o} of visual angle.

We note that our interest in ground suppression requires that we precisely isolate the groundside of the silhouette borders. To do so, we conducted high-resolution scans (1 mm voxels), which restricted our acquisition such that we were unable to assess activation beyond area V4 (see Section 2.1.4). Hence, the highest level at which we are able to assess competition-mediated suppression is V4.

To assay competition in the perception of a single object, we adapted an fMRI design similar to that used by others to measure competitive suppression between objects (e.g., Beck & Kastner, 2005; 2007; Scalf, Basak & Beck, 2011). Participants performed a demanding rapid serial visual presentation (RSVP) task at fixation while the silhouettes appeared individually in either the upper right visual field (RVF) or the upper left visual field (LVF). Using lateralized presentations allowed us to accurately locate representations in visual cortex of the contralateral hemisphere. Previously published fMRI studies have presented competing stimuli to the RVF only and assessed left hemisphere (LH) activation only (e.g., Beck & Kastner, 2005; 2007; Scalf et al., 2011). Thus, we expected to observe evidence of competition-mediated ground suppression in the LH when we presented our silhouettes in the RVF, but we were uncertain whether we would observe evidence of competition-mediated ground suppression in the right hemisphere (RH) when we presented our silhouettes in the LVF, especially given that our stimuli are so different from those previously used. Moreover, prior studies have observed different patterns of activation in the LH and RH visual cortex when participants made familiarity judgments regarding figure-ground stimuli (Peterson, Cacciamani, Barense, & Scalf, 2012), suggesting that neural responses are not consistent across hemispheres. Accordingly, we plan to test activation in each hemisphere separately.

If neural responses show evidence of competition-mediated suppression of the losing ground regions, we expected to observe reduced blood oxygen level-dependent (BOLD) responses to ground regions of *high-competition* compared to *low-competition* silhouettes. If we find the patterns of BOLD responses predicted here, these results would elucidate how suppressive competition generates object perception, and would provide neural evidence that object memories enter into the competition.

1.1. Précis

In Experiment 1, LH BOLD activation was significantly lower in the cortical representation of the grounds of *high-competition* than *low-competition* silhouettes, supporting the hypothesis that, in the perception of a single object, the degree of ground suppression varies with the magnitude of cross-border competition. The same pattern of activation was evident in all visual areas (V1, V2, and V4), consistent with the hypothesis that suppression, initiated in higher-level regions where the competition occurs, is fed back to lower-level brain regions. These differences were observed for single objects shown in the periphery—objects that were irrelevant to participants' primary RSVP task at fixation. Evidence of competition-mediated ground suppression was observed only in the LH with RVF silhouette presentation, however, not in the RH for LVF presentation. One explanation for these lateralized findings is that attention is more likely to be captured by task-irrelevant stimuli in the RVF than LVF (e.g., Newman, O'Connell, & Bellgrove, 2013; Reuter-Lorenz, Kinsbourne, & Moscovitch, 1990; Takio, Koivisto, Tuominen, Laukka, & Hämäläinen, 2013). Using performance on an RSVP task at fixation to assay attentional capture. Experiment 2 showed that task-irrelevant silhouettes presented in the RVF captured attention whereas those shown in the LVF did not. The results of Experiment 2 suggest that attention may be necessary to amplify differential figure-ground responses indexed by fMRI (Poort, Raudies, Wannig, Lamme, Neumann, & Roelfsema, 2012) thereby allowing us to observe differences in competition-mediated ground suppression.

2. Experiment 1

2.1. Experiment 1 Methods

2.1.1. Participants

Six right-handed volunteers (4 females; ages 22-39) were recruited for this experiment and were compensated for their participation. All participants reported normal or corrected-to-normal visual acuity. Before the experiment, participants gave written informed consent to participate, which was approved by the Institutional Review Board of the University of Arizona.

2.1.2. Stimuli

The stimuli were 80 enclosed, symmetric, silhouettes (cf., Peterson & Kim 2001; Peterson & Skow, 2008; Salvagio et al., 2012; Sanguinetti, Allen, & Peterson, 2014; Trujillo, Allen, Schnyer, & Peterson, 2010) that were smaller in area than the backdrops on which they were presented. These properties favored that percept that the inside was the object. The borders of all of the silhouettes depicted novel shapes in the inside region. The vertical borders of *low-competition* silhouettes (N=40) suggested novel shapes on the outside (groundside) as well. The vertical borders of *high-competition* silhouettes (N=40) suggested portions of real-world, familiar objects on the outside (ground) (see Appendix A for a list of the familiar objects suggested on the outside of the *high-competition* silhouettes). Crucially, participants were not aware of the familiar objects suggested on the outside/groundside of the silhouette borders; the inside region was always perceived as the object (as determined by extensive post-experiment questioning; see below). The *high*- and *low-competition* silhouettes were equated on low-level stimulus attributes, including luminance, contour length, area, curvature, spatial frequency, horizontal span, symmetry, enclosure, and convexity (see Trujillo et al., 2010). Therefore, we are reasonably confident that the only difference between the *high*- and *low-competition* silhouettes was the suggestion of a familiar object on the groundside of the borders of the former but not the latter.

Silhouettes subtended 4^{o} in height and an average of 5.2^{o} in width. They were presented in either the upper LVF or the upper RVF and positioned such that the nearest bottom corner was 4^{o} from the center of the screen. The color of the silhouette (blue, green, red, yellow, or white) was randomly chosen, with the constraint that 5 colors were displayed equally often for the *high*- and *low-competition* silhouettes. The screen backdrop was black throughout the experiment.

2.1.3. Experimental Design & Equipment

During the experiment, participants performed a difficult RSVP task at fixation. Specifically, they searched for lowercase letters (*a*, *b*, or *c*) embedded in a 4 Hz stream of digits (0-9) and ASCII symbols (#, %, %, *) presented centrally in white 40-point font (subtending 0.5^{o} of visual angle). The RSVP stream was continuous throughout the entire experiment, and (unbeknownst to the participants) a target appeared every 4 seconds on average. Participants' task was to press a button every time a lowercase letter appeared. This task was based on an RSVP task used by others who have argued that its demanding nature ensures that participants' attention is focused on fixation (cf. Beck & Kastner, 2005; 2007; Scalf, Basak & Beck, 2011).[1] We used it for the same purpose.

Figure 2 depicts the block design employed in this experiment. Each block consisted of 10 silhouettes presented individually in random order for 250 ms each, separated by a jittered inter-stimulus interval (ISI; 750, 1000, 1250, 1500, or 1750 ms). Each run consisted of four 15-second blocks, with one block of each condition (LVF *low-competition*, LVF *high-competition*, RVF *low-competition*, RVF *high-competition*). All 80 stimuli were presented once in each visual field, resulting in 8 runs. A 12-second baseline "off" period (in which the RSVP task continued but no silhouettes were presented) started and ended each run and separated each block.

< Insert Figure 2 about here >

RSVP responses were recorded using a Lumina (Cedrus Corp) response pad and controller. Participants viewed the stimuli on a ThinkVision 1920 x 1200 LED monitor via rear projection onto a mirror positioned above the head coil. Stimuli were presented using Vision Egg (Straw, 2008) software running under Windows 7 Professional.

Post-experiment questions. After the scanning session, all participants were asked a series of questions designed to ascertain whether they were aware of any of the familiar objects suggested on the groundside of the *high-competition* silhouettes. When they were asked about this, they were shown a sample *high-competition* silhouette that was not used in the experiment while the experimenter traced the portion of the familiar object suggested on the groundside of the border, and made certain the participant saw the object and could identify it. The

---

[1] Indeed, in post-experiment questioning, participants reported they had difficulty with the task and that as a result, they could not actively pay attention to the silhouettes in the periphery.

experimenter then directly asked whether the participant ever noticed anything suggested on the groundsides of the borders of the silhouettes shown during the experiment. None of the participants reported having seen portions of familiar objects in the grounds of the silhouettes during the experiment. Participants generally reported that they barely noticed the silhouettes; if they noticed anything, they reported it was simply that something colorful was appearing in their periphery. In fact, participants were quite surprised to hear that these colorful silhouettes suggested familiar objects on their groundsides. We therefore are reasonably confident that the colored, inside regions of the silhouettes were perceived as the objects and that the outsides of the both the *high*- and the *low-competition* silhouettes were perceived as shapeless grounds.

2.1.4. Data Acquisition and Analysis

Images were collected on a 3.0 T Siemens Skyra whole-body scanner using a 32-channel head coil. We used an EPI pulse sequence (TR = 3, TE = 30, FOV = 240 x 240 mm) to acquire 27 coronal slices starting from the occipital pole. Slices were 1 mm thick in order to finely localize the ground regions (see below); due to this high resolution, slices extended only as far forward as V4. Each of the 8 experimental runs consisted of 40 repetitions. To assist in image registration, a T2 image coplanar with the functional slices was also collected.

FMRIB (Oxford University Centre for Functional MRI of the Brain) Software Library (FSL) was used to analyze the functional data. Preprocessing steps included brain extraction, high-pass filtering (60 seconds), and motion correction [FSL 4.1.9 (Jenkinson et al., 2002; Smith et al., 2004)]. Data were concatenated across all eight runs and submitted to a GLM analysis using FSL's FMRI Expert Analysis Tool (FEAT) v. 5.98 [FSL 4.1.9 (Smith et al., 2004; Woolrich et al., 2001)]. We modeled 4 regressors of interest corresponding to each of our conditions in both hemispheres (LVF *low-competition*, LVF *high-competition*, RVF *low-competition*, RVF *high-competition*). These regressors were convolved with a double-gamma model of the HRF (Phase 0s). The resulting statistical maps were registered into each participant's individual anatomical space.

2.1.5. Retinotopy Procedures and Analysis

All participants underwent a phase-encoded retinotopic mapping scan. The procedure for this scan was derived from Sereno et al. (1995). Participants viewed a flickering black-and-white checkerboard wedge extending outwards from a central fixation cross to the edge of the screen. The wedge rotated clockwise at a rate of one rotation per minute, stimulating early visual areas in a predictable pattern. Participants were asked to fixate on the central cross and press a button to indicate the appearance of a gray circle in the checkerboard pattern on the wedge; this ensured that participants attended to the wedge. During this session, an EPI pulse sequence (TR = 3.05, TE = 30, FOV = 240 x 240 mm) was used to acquire one run of 30 coronal slices starting from the occipital pole (1 mm thickness). A high-resolution (1 mm) T1-weighted MPRAGE scan was also collected to submit to Freesurfer (Dale et al., 1999; Fischl et al., 1999; 2001) for averaging and segmentation.

A region of interest (ROI) analysis was employed to interrogate visual cortex using the retinotopy data. Each participant's V1, V2, and V4 were delineated for each visual field and projected into their individual Freesurfer space.[2]

---

[2] The delineation of visual area VP was difficult to determine because the representations of its borders were often overlapping—an issue that others have encountered previously (Shipp, Watson, Frackowiak, & Zeri, 1995). Thus, VP was not included in any analyses.

2.1.6. Functional Localizer

*Ground ROIs***.** In order to identify the cortical area corresponding to the grounds of our silhouettes, a separate series of functional localizer scans was conducted after the experimental runs. Given that the width of the silhouettes varied, only ground regions delineated by the vertical border closest to fixation in each visual field were localized; these locations were the same for every silhouette. Specifically, the groundside of the silhouettes was defined as a rectangular patch extending 2° inward (towards fixation) from an imaginary vertical line drawn on the edge of the silhouette border closest to fixation (see Figure 3A; all silhouettes were aligned with the same imaginary vertical line). Defining the ground in this manner was necessary because the vertical borders of the silhouettes were articulated; thus, the precise location of the ground immediately adjacent to the object varied.

< Insert Figure 3 about here >

During these localizer scans, participants passively viewed vertically elongated dynamic (cycling at a rate of 8 Hz), colored Gabor stimuli presented in this location (in both the RVF and LVF) while maintaining fixation on a central white cross. Gabors were presented for 250 ms and repeated 40 times in each of four 20-second blocks per run. A 20-second off period separated each block. For each participant, four localizer runs were collected. Data were concatenated across all four runs and were overlaid on the participant's retinotopic map. Within each visual area (V1, V2, and V4), the peak activation for the ground region localization was found, and thresholding was set at 30% of that peak value. All suprathreshold voxels that were contiguous with the peak activation were assigned to that region's ground ROI. Figure 3B shows an example of the ground ROIs in each visual area in the left hemisphere in one of the participants. In addition, for each subject in each hemisphere, a larger “combined” ground ROI collapsed across visual area was created by adding together the V1, V2, and V4 ground ROIs in that hemisphere. Featquery (Smith et al., 2004) was used to extract the mean parameter estimates for the *high-competition* and *low-competition* conditions from this combined ground ROI, as well as from the separate ground ROIs for each visual region. The region extending 2° on the other side of the imaginary vertical line was also localized (see Figure 3A). Because the borders of the silhouettes were articulated, this ROI necessarily contained substantial portions of both figure and ground. Any voxels that were active during this localization were removed from the ground ROI in order to produce a more precise representation of the ground region. Therefore, our localization methods and analyses were highly conservative in that we ensured that no voxels activated for the figure were included in our definition of the ground.

Figure ROIs. In addition to these ground ROIs, we also created figure ROIs in each visual area. These figure ROIs were not created via a localization scan as localizing the figure using a Gabor patch would have been difficult because in anchoring our silhouettes such that the ground region was in the same location for every silhouette, the figure location necessarily varied, given the different articulated borders and widths of the silhouettes. Therefore, figure ROIs were defined based on peak activation (in each visual area) resulting from presentation of both *high-* and *low-competition* silhouettes. Thresholding was set at 30% of that peak value, and all contiguous suprathreshold voxels in a given brain region were assigned to that region’s figure ROI. As with the ground ROI, a larger “combined” figure ROI was created by adding together the V1, V2, and V4 figure ROIs. Using Featquery (Smith et al., 2004), parameter estimates for both *high-* and *low-competition* silhouettes were extracted from the figure ROI combined across visual areas within each hemisphere, as well as from the separate ROIs in each visual area. This figure ROI was created to allow us to test whether any differential suppression we observed for *high-* and *low-competition* silhouettes was restricted to grounds as expected, or whether it extended to figures as well. The latter finding would be consistent with the hypothesis that there is greater inhibitory competition for object status in *high-competition* than *low-competition* silhouettes but would not provide evidence for competition-mediated ground suppression.

2.2. Experiment 1 Results

2.2.1. Behavioral results

An analysis of behavioral performance on the RSVP letter-detection task showed that accuracy was low (mean accuracy = 14.5%). The task was designed to be difficult in order to keep participants focused on fixation; the RSVP stream was fast (4 Hz), continuous throughout the entire experiment, and small in size (0.5^o of visual angle). Thus, we expected low accuracy, and our participants' accuracy was even lower than in previous studies using this task (e.g., Beck & Kastner, 2005; 2007), as we did not provide training or practice trials beforehand as others did. The low accuracy scores suggest that participants adopted a very conservative response criterion; importantly, this means that although misses were high, correct rejections were also high. To account for correct responses in the form of hits as well as correct rejections, a d' analysis of sensitivity was conducted, where d' = z(Proportion Hits) – z(Proportion False Alarms). The resulting d' scores provide a more precise measure of performance than accuracy percentages alone, with higher scores indicating better performance. This analysis revealed a mean d' of 2.36, which falls within the range of normal d' scores that others have observed on this task (e.g., Scalf et al., 2011; Scalf & Beck, 2010). Therefore, although accuracy was at floor, the d' analysis suggests that participants were indeed focused on their difficult task at fixation. No differences in RSVP task performance (as measured by both accuracy scores and d' scores) were observed between baseline blocks where no silhouette was present and blocks where silhouettes were presented in either the LVF or RVF (*p*s > .30).

2.2.2. Ground ROI analysis

Figure 4 graphs the mean parameter estimates extracted from the ground region in visual cortex for *high*- and *low-competition* silhouettes for contralateral visual field presentation (RVF for LH, LVF for RH) and ipsilateral visual field presentation (LVF for LH, RVF for RH). For each hemisphere, activation when stimuli are presented in the ipsilateral visual field serves as a baseline against which to compare activation when stimuli are presented in the contralateral visual field. We expect that for a given hemisphere, any difference between the activation for *high*- and *low-competition* silhouettes should be observed when the stimuli are in the contralateral visual field and not the ipsilateral visual field.

Separate ANOVAs were conducted on the parameter estimates extracted from the combined ground ROI collapsed across visual area (V1, V2, and V4) for each hemisphere.[3] As can be seen in Figure 4A, for the LH, a significant interaction between visual field (contra/ipsilateral) and silhouette type (*high/low-competition*) was observed, $F(1,5) = 7.19$, $p = .04$. Follow-up t-tests on the LH data revealed that for contralateral visual field presentation, mean activation in the ground region was significantly lower for *high*- than for *low-competition* silhouettes, $t(5) = 2.76$, $p = .04$, as predicted if ground suppression is greater when grounds

---

[3] Recall that separate tests of the activity in each hemisphere were justified a priori (see Section 1). Separate analyses for the two hemispheres are required to avoid double dipping as well since we used activation with ipsilateral presentation as the baseline for activation with contralateral presentation.

compete more for object status. In contrast, for ipsilateral visual field presentation, no significant differences between *high-* and *low-competition* silhouettes were observed in the LH ($p > .50$). Five of our 6 participants showed this overall activation pattern in the LH (see Figure 5).

< Insert Figures 4 and 5 about here >

In contrast to the results obtained in the LH, no significant differences between the activation of grounds of *high*- vs. *low-competition* silhouettes were observed in the RH, regardless of whether the stimuli were presented in the contralateral or the ipsilateral visual field (*p*s > .24; see Figure 4B).

To further compare the patterns of activation between visual areas, parameter estimates were extracted from the ground ROIs in each visual area and submitted to a 3-way ANOVA with factors of visual field (contra/ipsilateral), silhouette type (*high/low-competition*), and visual area (V1/V2/V4) for each hemisphere (see Figure 4C). While we acknowledge that the mean parameter estimates extracted from ROIs in each region of visual cortex are not independent of each other, comparing across them can provide insight into their contributions to the pattern observed in the ANOVA on the combined ROI. This 3-way ANOVA conducted on the LH data revealed a significant interaction between visual field and silhouette type, $F(1,2) = 21.97$, $p < .001$, with significantly less activation in the grounds of *high*- vs. *low-competition* silhouettes when they were in the contralateral visual field but not when they were in the ipsilateral visual field. Importantly, no main effects or interactions involving visual area were observed, *p*s > .42. This lack of significance indicates that the pattern of activity does not differ by visual area—a result consistent with the hypothesis that ground suppression is fed back from high to low levels of the visual cortex.

A similar 3-way ANOVA conducted on the RH data produced no significant main effects or interactions, *p*s > .16 (see Figure 4D). This pattern replicates that found in the ANOVAs performed on the data from the ground ROI combined across visual area.

2.2.3. Figure ROI analysis

Figure 6 graphs the mean parameter estimates extracted from the figure ROI in visual cortex for *high*- and *low-competition* silhouettes for contralateral and ipsilateral visual field presentation. Separate ANOVAs with factors of visual field (contra/ipsilateral) and silhouette type (*high/low-competition*) were conducted on the data extracted from the combined figure ROI in each hemisphere, collapsed across visual area (see Figure 6A&B). In the RH, activation was statistically higher for stimuli presented in the contralateral rather than the ipsilateral visual field, *F*(1,5) = 7.61, *p* = .04; this difference was marginally significant in the LH, *F*(1,5) = 4.53, *p* = .08. Importantly, though, no interactions with or main effects of silhouette type (*high*- vs. *low-competition*) were observed in either the LH or RH (*p*s > .13). To investigate the patterns of activation between visual areas, a 3-way ANOVA was conducted on the data extracted from the figure ROIs in each visual area with factors of visual field (contra/ipsilateral), silhouette type (*high/low-competition*), and visual area (V1/V2/V4) (see Figure 6C&D). A main effect of visual field was observed in both hemispheres, *F*s(1,2) = 5.01 & 8.23, *p*s = .05 & .03 for LH and RH, respectively. No other significant main effects or interactions were observed in either the LH or the RH, *p*s > .10.

That the figure ROIs do not show the same reduced activation for *high-competition* silhouettes compared to *low-competition* silhouettes observed in the ground ROIs is consistent with the hypothesis that the activation observed in the ground ROIs represents competition-mediated ground suppression and not simply reduced activation in the visual area due to competition.

< Insert Figure 6 about here >

2.3. Experiment 1 Discussion

In Experiment 1, BOLD activation in the LH was significantly lower in the cortical representation of the grounds of *high-competition* than *low-competition* silhouettes. We take these data as neural evidence that suppressive competition underlies the perception of a single object. The results show that greater competition for object status from regions ultimately perceived as grounds produces greater suppression of the ground, supporting previous behavioral data (Salvagio et al., 2012). The same pattern was evident across V1, V2, and V4. In V4, receptive fields are large enough to compass the vertical extent of the silhouettes and therefore to encompass the portion of the familiar objects suggested on the groundside of the borders of the *high-competition* silhouettes. Therefore, we expect that competition for object status occurs in V4 or in higher regions. Whether the differential ground suppression we observed originated in V4 or a higher-level brain area must be investigated in future research, as data collection in the present study did not extend past V4 due to the high-resolution scans (1 mm voxels) required to isolate the groundsides of our silhouettes in visual cortex.

In V1 and V2, receptive fields are too small to encompass more than one or two parts of the familiar configurations suggested on the groundsides of the borders of the *high-competition* silhouettes. Other experiments have shown that familiar parts alone are not sufficient to produce effects of familiarity on object status; the parts must be properly arranged to form a familiar configuration (e.g., Gibson & Peterson, 1994; Peterson, Harvey, & Weidenbacher 1991).

Therefore, we do not expect that there was differential competition in V1 and V2 for *high*- vs. *low-competition* silhouettes. Moreover, the *high*- and *low-competition* silhouettes were equated on low-level stimulus features. Therefore, it is unlikely that differences observed in V1 and V2 are due to stimulus differences. Thus, it is likely that evidence of competition-mediated ground suppression observed in V1 and V2 is due to feedback from higher levels where receptive fields are larger. (For similar arguments involving fMRI data, see Likova & Tyler, 2008; Strother et al., 2012.)

The fMRI data reported here provide neural evidence supporting behavioral data showing that both the properties and the location of the potential object on the losing side of the border of a *high-competition* silhouette are suppressed. Peterson and Skow (2008; Salvagio et al., 2012) hypothesized that such suppression may be partially responsible for the fact that the groundside of a border is perceived as shapeless. Based on these results, models that implement suppressive competition in figure-ground perception should be amended to include competition-mediated ground suppression.

The pattern of results observed in LH figure ROIs differed from that observed in LH ground ROIs. This finding allows us to rule out an alternative interpretation that our ground effects simply show reduced activation for all competitors when competition is high, as has been reported by Kastner and colleagues in paradigms contrasting simultaneous (*high-competition*) to successive (*low-competition*) presentation of multiple objects within V4 RFs (for a review, see Beck & Kastner, 2009; for further discussion, see Section 4). Our results are the first neural evidence that high-level competition-mediated ground suppression underlies the perception of a single object.

Experiment 1 employed lateralized presentation to allow us to target the contralateral visual cortex. The use of lateralized presentations uncovered a previously unobserved hemispheric asymmetry: competition-mediated ground suppression was only evident in the LH with RVF presentation, not in the RH with LVF presentation. Hemispheric differences were not observed previously because stimuli were either presented centrally (e.g., Likova & Tyler, 2008; Peterson & Kim, 2001; Peterson & Skow, 2008; Salvagio et al, 2012; Strother et al., 2012) or were presented in the RVF only and only LH activation was assessed (Kaster & Beck, 2005; 2007; Scalf et al., 2011). Although laterality effects are not uncommon in cognitive neuroscience, this result still raises the question of why our effect was only observed the LH.

One possible reason is that our stimuli were task irrelevant, and task-irrelevant stimuli are more likely to attract attention when they are presented in the RVF rather than the LVF (Newman et al., 2013; Reuter-Lorenz et al., 1990; Takio et al., 2013). Perhaps the stimuli must be attended in order for us to be able to observe evidence of competition-mediated suppression using fMRI. Indeed, using single cell recording methods, Poort et al. (2012) showed that modulation of neural responses by figure-ground status is attenuated when attention is directed elsewhere. FMRI may be less sensitive to attenuated responses. In Experiment 2, we test whether task irrelevant silhouettes presented in the periphery are more likely to capture attention when they appear in the RVF rather than the LVF.

3. Experiment 2

It has been shown that the “capture” of attention by the onset of a peripheral stimulus can impair performance on an RSVP task at fixation; worse RSVP task performance indicates greater capture of attention by the peripheral stimulus (Folk, Leber, & Egeth, 2002; see also Folk, Remington, & Johnston, 1992; Remington, Johnston, & Yantis, 1992; Yantis & Jonides, 1990). The RSVP data obtained from Experiment 1 did not show any evidence of differential attentional capture by silhouettes in the RVF vs. the LVF, but accuracy on the RSVP task was at floor, which may have prevented the emergence of differential effects. In Experiment 2, we used the presentation parameters of a typical experiment in the attentional capture literature to assess whether silhouettes are more likely to capture attention when they are shown in the RVF rather than the LVF. As in Experiment 1, participants' primary task was an RSVP task at fixation. Here, their task was to report the identity of the single target letter that appeared among digits and ASCII characters during individual trials. Participants reported the identity of the target letter after each trial. Only one silhouette appeared during each trial. The absence of repeated onsets in the periphery made the RSVP task easier in Experiment 2 than in Experiment 1, as did dividing the RSVP task into short trials and requiring a response after every trial. Therefore, we expected RSVP accuracy to be much better in Experiment 2 than in Experiment 1. The silhouette was always presented two items prior to the target letter—the lag at which Folk et al. (2002) observed the greatest attentional capture effects. Note, however, that the temporal lag between the distractor and the target was variable in Folk et al.'s experiment whereas in Experiment 2, we used a fixed lag between the silhouette and the RSVP target. Therefore, if participants' attention is captured by the silhouettes in our experiment, they may learn to use the silhouette as a cue to when the RSVP target is likely to occur. This would not have happened in Experiment 1.

3.1. Experiment 2 Methods

3.1.1. Participants

Participants were 26 naïve right-handed undergraduates (18 female; ages 18-22) at the University of Arizona who participated in this experiment for partial fulfillment of course requirements. All participants reported normal or corrected-to-normal visual acuity. Before the experiment, participants gave written informed consent to participate, which was approved by the Institutional Review Board of the University of Arizona. An additional 3 participants were removed from the analysis for having less than 20% accuracy across trials.

3.1.2. Stimuli

The silhouettes used in Experiment 2 and the RSVP task items were the same as those used in Experiment 1.

3.1.3. Experimental Design & Equipment

As in Experiment 1, participants in Experiment 2 performed a difficult RSVP task at fixation, but because our interest here is in testing whether our task-irrelevant stimuli are more likely to capture attention when they appear in the RVF rather than the LVF, the procedure was adapted in order to be optimal for measuring attentional effects from a peripheral stimulus on performance (cf., Folk et al., 2002). Each trial began with a central fixation cross; the participant pressed the space bar when ready to begin. A 15-item RSVP stream then appeared: 14 of the items were digits or ASCII symbols, and 1 item was the target letter (a-z). Each item appeared in the center of the screen for 42 ms and was followed by a 42-ms inter-stimulus interval (ISI) before the next item appeared; thus, each trial lasted 1026 ms. The target letter appeared equally often in positions 8-12 (randomly intermixed) of the 15-item stream. One silhouette appeared during each trial; it was always presented in the periphery at the same time as an RSVP item that preceded the target letter by two items (i.e., in positions 6-10 of the 15-item RSVP stream.

Participants' task was to identify the target letter after each trial using the keyboard. Participants received no feedback on their performance. Following their response, the fixation cross for the next trial appeared. Each item in the RSVP stream was presented for a shorter duration in Experiment 2 than in Experiment 1, but the between-item ISI, the shorter trials, and the absence of repeated onsets in the periphery made the task easier.

On baseline trials, no other stimuli appeared on the screen during the RSVP stream; this provided a measure of participants' baseline performance on the RSVP task. On test trials, a silhouette (either *high*- or *low-competition*) appeared in either the upper RVF or LVF two items prior to the appearance of the target letter. As in Experiment 1, the bottom-most corner of the silhouettes appeared 4^{o} from fixation, all silhouettes were 4^{o} high, and each silhouette appeared once in each visual field. There were 200 trials, 40 of each condition (baseline, LVF *high-competition*, LVF *low-competition*, RVF *high-competition*, and RVF *low-competition*), randomly intermixed. Prior to the experimental trials, there were 8 practice trials on which feedback was given. None of the silhouettes shown during practice was used during the experiment.

Stimuli were presented using Experiment Builder software on a 21-in. Sony CRT monitor and a personal computer. Participants used a chin rest to stabilize their head during the experiment. To ensure that participants were maintaining fixation and not looking at the silhouettes in the periphery, eye movements were monitored using Eyelink 1000 eye-tracking software with a desktop mount. One participant who looked at the silhouettes on more than 15% of trials was removed from the analysis (criterion established in advance). For the remaining participants, individual trials on which participants moved their eyes were selectively removed from the analysis, as poor performance on these trials would be attributed to those eye movements away from their task at fixation rather than the allocation of attention per se.

As in Experiment 1, participants were asked a series of rigorous post-experiment questions after completing the experiment to ascertain whether or not they were aware of the familiar objects suggested on the groundsides of the silhouette borders (see Experiment 1 methods). None of the participants reported having seen these familiar objects.

3.2. Experiment 2 Results

A preliminary analysis of the data revealed that performance on the RSVP task improved substantially from the first to the second half of the experiment. To eliminate these learning effects and avoid ceiling performance, only the first 20 trials of each condition (baseline, LVF *high-competition*, LVF *low-competition*, RVF *high-competition*, and RVF *low-competition*) were analyzed (100 trials total).

Average performance was quite good (65%) and much better than RSVP task accuracy in Experiment 1. Figure 7 graphs the mean accuracy scores on the RSVP task for each condition, as well as accuracy on the test conditions subtracted from accuracy in the baseline condition. Since we are interested in how central RSVP performance is altered by the onset of a task-irrelevant silhouette in the periphery, we conducted our statistical analyses on the difference scores (baseline – test)[4]. A 2 x 2 ANOVA with factors of visual field (LVF/RVF) and test trial silhouette type (*high/low-competition*) revealed a significant interaction between test trial silhouette type and visual field [$F(1,25) = 6.26$, $p = .019$]: For RVF silhouette presentations, accuracy on the RSVP task was significantly lower on trials on which *high-competition* silhouettes appeared than

---

[4] Note that the same effects were observed in an ANOVA conducted on the raw accuracy scores.

on baseline trials [$t(25) = 2.02$, $p = .05$] and RSVP accuracy was significantly higher on trials on which *low-competition* silhouettes appeared than on baseline trials [$t(25) = 2.16$, $p = .045$], whereas for LVF silhouette presentations, RSVP accuracy did not differ significantly from RSVP accuracy on baseline trials, on either *high*- or *low-competition* trials (ps > .55). The interaction subsumed a main effect of silhouette type [$F(1,25) = 14.46$, $p = .001$], indicating that RSVP performance differed from baseline more on *high-competition* than *low-competition* trials.

< Insert Figure 7 about here >

3.3. Experiment 2 Discussion

In Experiment 2, we used an attentional capture task to investigate whether attention was more likely to be drawn to task-irrelevant silhouettes shown in the RVF than the LVF. Our results show that it was. RSVP accuracy was altered relative to baseline accuracy when silhouettes appeared in the RVF, but not when they appeared in the LVF. Specifically, for RVF presentations, accuracy was increased relative to baseline on trials on which *low-competition* silhouettes appeared and was reduced relative to baseline accuracy on trials on which *high-competition* silhouettes appeared. We explain the opposite pattern of results as follows: with RVF presentations, participants learned implicitly that RSVP targets appeared shortly after the (attended) silhouettes and consequently attempted to return their attention to the RSVP task as quickly as possible. Attention could be disengaged from the *low-competition* silhouettes in time to detect the RSVP targets (accounting for improved performance relative to baseline). It may have taken longer to disengage attention from the *high-competition* silhouettes because it takes longer to resolve the greater competition in the latter than the former (Peterson & Enns, 2005; Peterson & Lampignano, 2003), As a consequence, when *high-competition* silhouettes appeared in the RVF, attention was not always reallocated to RSVP stream in time to detect the target, leading to impaired RSVP performance. Neither impaired performance on *high-competition* trials nor improved performance on *low-competition* trials was observed for LVF presentations, consistent with the proposal that LVF presentations did not capture attention, and consequently, participants did not learn the contingency between the silhouette presentation and the RSVP target presentation. Indeed, some previous work indicates that implicit learning requires attention (Jiang & Chun, 2001; Nissen & Bullemer, 1987).

4. General Discussion

In this study, we used fMRI to search for neural evidence that suppressive competition underlies the perception of a single object by assessing suppression of BOLD activation in visual cortex areas that represent the groundside of an object. We presented participants with lateralized task-irrelevant novel enclosed silhouettes whose groundsides either suggested a meaningless shape (*low-competition* silhouettes) or a meaningful familiar object (*high-competition* silhouettes). We found that when stimuli were presented in the RVF, BOLD activation in the cortical representation of the groundside of the silhouettes in LH visual cortex was significantly lower for *high*- vs. *low-competition* silhouettes. No differences in BOLD activation were observed in the cortical representation of the figure side of the *high*- vs. *low-competition* silhouettes. These important findings support the hypothesis that, prior to the perception of a single object, object properties on opposite sides of a shared border competed for object status, and greater competition from the side ultimately determined to be the ground resulted in greater suppression of the ground. The current study provides the first neural evidence in support of behavioral evidence (Peterson & Kim, 2001; Peterson & Skow, 2008; Salvagio et al, 2012) that shape-level competition-mediated ground suppression underlies the perception of a single object.

The results of the current study shed light on existing theories of object perception. One set of theories and models posit that two adjacent regions compete for object status (Craft et al., 2007; Grossberg, 1994; Kienker et al., 1986; Sejnowski & Hinton, 1987) with the winner perceived as the object, and the loser perceived as a shapeless ground. There has been some debate regarding whether neural responses are facilitated for perceived objects (Appelbaum, Wade, Vildavski, Pettet, & Norcia, 2006; Appelbaum, Wade, Pettet, Vildavski, & Norcia, 2008; Lamme, 1995), suppressed for grounds (Likova & Tyler, 2008; Peterson & Skow, 2008; Salvagio et al., 2012; Tsotsos, Culhane, Kei Wai, Lai, Davis, & Nuflo, 1995), or whether both effects occur (Strother et al., 2012). Here, our use of retinotopic mapping and precise localization of the cortical representations of the figure and ground allowed us to provide supporting evidence for ground suppression as a mechanism behind object perception. Moreover, the present experiment extends the previous literature by showing that more suppression is applied to grounds that compete more for object status.

Additionally, that our effect was observed across visual areas V1 through V4 is consistent with previous work showing that suppression of the ground involves feedback from high to low levels of the visual system (Likova & Tyler, 2008; Salvagio et al., 2012; Strother et al., 2012). Our two types of silhouettes differed only at the level of the shape suggested along the vertical extent of the silhouette borders, which were 4° of visual angle in height. At the eccentricity we used (4° from center), receptive fields in V1/V2 are approximately 1-2°—not large enough to encompass the familiar object suggested on the groundside of the borders of the *high-*

competition silhouettes. Thus, the pattern of activation observed even in low levels of visual cortex—that is, reduced activity for *high*- vs. *low-competition* silhouettes—likely originated from higher levels where receptive fields are large enough to encompass the entire vertical border. This high level area could have been V4, where we also observed our effect, or it could be an area higher in the visual hierarchy, such as the inferior or medial temporal lobe. The current study cannot provide information regarding competition occurring higher in the visual hierarchy, as our use of high-resolution scans (1 mm)—which were necessary for our precise ground localization technique—only allowed activation to be assessed up to and including V4. Had we been able to measure activation at higher levels, we would expect effects of competitive suppression to be evident past V4, as previous research has found evidence of competitive suppression as high as the inferior temporal cortex (e.g., Zoccolan et al., 2005). Regardless of whether the competition-mediated ground suppression originated in V4 or in higher-level visual areas, we posit that feedback is accounting for the effects we observed in lower-level visual areas V2 and V1. Our conjecture that the suppression we observed at low levels originated in higher-level brain areas is congruent with previous neuroimaging research that showed neural evidence of top-down mediated ground suppression in object perception (Likova & Tyler, 2008; Strother et al., 2012). Note, however, these previous studies compared their test condition to a control condition in which no segregation between object and ground occurred (e.g., a matched undifferentiated random dot display in Likova and Tyler, 2008). Our study is the first that has incorporated a control for low-level features, and hence, low-level inhibitory interactions (our *low-competition* silhouettes). This important manipulation allows us to extend these previous studies by showing not only that top-down ground suppression can be measured using fMRI, but that fMRI can also reveal evidence of competition and suppression originating at a level beyond edges and features.

Our data suggest that reentrant connections account for the evidence of competition-mediated ground suppression we observed at low levels, but what is the precise role of this feedback? According to predictive coding models, feedback acts to suppress low-level representations that are *consistent* with predictions posed by higher-level regions (Rao & Ballard, 1999). On these models, the residual error signals between the predicted input and the actual sensory input (i.e., the *inconsistent* representations) remain active and are fed forward. Hierarchical Bayesian inference models, on the other hand, posit the opposite—that reduced activation at low levels is due to top-down suppression of *inconsistent* representations, while consistent information remains active (Lee & Mumford, 2003). Inhibitory feedback is thus implemented in both classes of models—predictive coding and Bayesian inference—but the information being suppressed differs. In our silhouettes, the information that is suppressed is the shape suggested the side of the border that is inconsistent with (or competing against) the best fitting interpretation (i.e., that the inside is the figure). Therefore, although the inhibitory nature of the feedback in the present study is consistent with both classes of models, hierarchical Bayesian inference models better explain our data; feedback suppresses representations that are inconsistent with the percept predicted at higher levels.

Previous research on competition has shown that when multiple items compete for representation, the result is mutual suppression of all competing items (for a review, see Beck & Kastner, 2009). Based on this work, an alternative explanation for our results could be that the competitive suppression we observed was not restricted to the ground and instead was applied to all competitors—that is, both the figure (i.e., the object) and its ground. Under this explanation, we should have observed reduced BOLD activation for *high*- vs. *low-competition* silhouettes on the figure side as well as the groundside. However, our ROI analysis of the figure side of the border revealed no differences in activation between our *high*- and *low-competition* silhouettes, ruling out this alternative explanation. Instead, we found that only the ground is suppressed as a result of losing the shape-level competition for object status. We posit that the figure side of the border is not suppressed because it ultimately "wins" the competition for object status and is consciously perceived as a shaped entity; this occurs for both *high*- and *low-competition* silhouettes. When the competition is resolved and a winner is determined, the competitive suppression is no longer mutual, but rather is concentrated on the "losing" ground region such that the shape suggested there is not consciously perceived. When the shape suggested in the ground is that of a familiar, real-world object as in our *high-competition* silhouettes (rather than a novel object as in our *low-competition* silhouettes), it competes more strongly for object status, and thus requires greater suppression in order for object status to be assigned to the inside of the silhouettes. Our *high*- and *low-competition* silhouettes are differentiated only by the familiarity of the object suggested on the groundside of the silhouettes; thus, the ground is the only place we expected (and observed) a difference in suppression. This finding of competition-mediated suppression of the ground, but not the figure, is predicted from and consistent with previous behavioral work, in which performance on an orientation discrimination task was worse when the target appeared on the groundside of a *high*- vs. *low-competition* silhouette, whereas performance for targets on the figure side did not differ by level of cross-border competition (Salvagio et al., 2012). In other words, the high degree of competition resulted in greater suppression of the losing region, but did not affect the perceived figure region—the same result we observed here.

We note that the region that we defined as the ground does not follow the contour of the silhouette border and therefore is not immediately adjacent to the border. Rather, it is positioned outside the silhouette border closest to the center of the screen (see Figure 3A). This localization method ensured that no portions of the region perceived as the object were included in the region defined as ground. Our investigation of ground activation was thus highly conservative, as suppression must have extended some distance away from the silhouette borders in order for us to observe it. Even so, we were still able to detect differential ground suppression for our *high-* vs. *low-competition* silhouettes in the LH, which speaks to the robustness of our effect.

The current study also speaks to an important debate in the literature on how the visual system processes objects in the environment. On the traditional view, objects are processed via a serial, feedforward sweep through the visual system; feedback is not necessary in this process (e.g., Serre, Oliva, & Poggio, 2007; Thorpe, Fize, & Marlot, 1996; Zhou, Friedman, & von der Heydt, 2000). The alternative view, however, posits that feedforward as well as feedback processing is necessary for conscious object perception (Bullier, 2001; Lamme & Roelfsema, 2000; Nadel & Peterson, 2013; Peterson & Cacciamani, 2013). Specifically, on an initial feedforward pass, objects that might be perceived on opposite sides of a shared border are assessed up to high levels of the visual system (for evidence consistent with this claim, see Cacciamani, Mojica, Sanguinetti, & Peterson, 2014; Peterson, Cacciamani, Mojica, and Sanguinetti, 2012; Sanguinetti et al., 2014). The best interpretation (i.e., the “winner”) is selected through a competitive process that entails suppression of the shape in the “losing” region and the best interpretation is integrated across levels of the visual hierarchy. Our results are consistent with this alternative dynamical view of object perception, given that they suggest that competition-mediated suppression is relayed via feedback to all levels of the visual system, including low levels.

In this study, our main result is a reduction of the BOLD signal elicited by the ground region, which we take as evidence for suppression. We acknowledge that the BOLD signal is an indirect measure of neural activity; previous work has shown that the BOLD signal correlates most highly with local field potentials, which can arise from either excitatory or inhibitory synaptic events (Logothetis & Wandell, 2004; Logothetis, 2002). A reduction in the observed signal can therefore arise in multiple ways, including (but not limited to) a decrease in inhibitory interneuron activity, a decrease in inter-regional recurrent activity, or a decrease in excitatory activity. Although we believe that our observed reduction in BOLD signal might be due to a decrease in local excitation, our use of fMRI does not allow us to definitively distinguish between these neural mechanisms. Future studies could address this question using a more direct measure of neural firing, such as single-cell recordings.

4.1. Laterality and Attention

Interestingly, the ground suppression arising from shape-level competition in Experiment 1 was only observed in the LH with RVF presentation, not in the RH with LVF presentation. Does this laterality effect reduce the support for our interpretation? We think not. Other researchers have shown that attention is more likely to be captured by task-irrelevant stimuli in the RVF than the LVF (Newman et al., 2013; Reuter-Lorenz et al., 1990; Takio et al., 2013). It is possible that the differences in BOLD responses for the grounds of *high-competition* vs. *low-*

competition silhouettes can only be observed when attention is directed to the silhouettes because attention amplifies neural response differences (Poort et al., 2012), allowing them to be observed. Experiment 2 used an attentional capture paradigm to investigate this hypothesis. The results showed that attention is captured by silhouettes that appear in the RVF, but not the LVF. We note that the connection between Experiments 1 and 2 is indirect; thus, we cannot ascertain that our attention explanation directly accounts for our lateralized fMRI results or is the only explanation for them. We can say, though, that attention is one potential explanation for the lateralized effects observed in Experiment 1. Future research will explore the generality of the laterality effects in research examining competition.

Our results are consistent with other research showing that attention can modulate the outcome of visual competition. For instance, binocular rivalry studies have shown that attention affects the likelihood of observing suppression arising from competition (Ling & Blake, 2012) and that attention modulates rivalry-dependent responses in V1 (Watanabe, Cheng, Murayama, Ueno, Asamizuya, Tanaka, & Logothetis, 2011; Zhang, Jamison, Engel, He, & He, 2011). Some of these studies posit that attention is necessary for competition to occur (Zhang et al., 2011), while others posit that attention simply enhances suppression resulting from competition (Ling & Blake, 2012). Consistent with these findings regarding competition in binocular rivalry, research investigating figure-ground segregation has shown that attention directed towards two regions sharing a border can enhance the modulation of neural responses by figure-ground status (Craft et al., 2007; Poort et al., 2012; Qiu, Sugihara, & von der Heydt, 2007).

The influence of attention on suppressive interactions in visual cortex is a well-investigated area of research; many studies, in fact, conclude that directing attention to a stimulus enhances its signal at the expense of its neighbors (Desimone & Duncan, 1995; Kastner et al., 1998). As mentioned earlier, in the present study, we found no difference in BOLD activation on the figure side of *high*- vs. *low-competition* silhouettes. Thus, we have no evidence that more attention was directed to the insides of the silhouettes under conditions of higher competition. Consistent results have been found in Salvagio et al.'s (2012) behavioral work, where performance on an orientation discrimination task was significantly worse for targets appearing on the groundside of *high-competition* than *low-competition* silhouettes, suggestive of greater suppression of the groundside of the former than the latter, whereas performance did not differ for targets appearing on the inside of *high-competition* vs. *low-competition* silhouettes. Had more attention been directed to the inside of the *high-competition* silhouettes to resolve the greater competition there, one might expect better orientation discrimination performance for targets shown on the figure side (the inside) of the borders of *high-competition* silhouettes.

4.2. Conclusion

Our results show that perception of a single visual object relies on suppressive extrastriate mechanisms that are sensitive to high-level competition for object status at a border shared by two regions in the visual input. This competition results in suppression of the losing region—the region ultimately perceived as the groundside of the figure. We have shown that competition at a relatively high level (the level of object shape) drives suppression at relatively low levels of visual processing (e.g., V2/V1) as well as higher levels (e.g., V4); we hypothesize that reentrant connections from higher levels mediate the suppression evident at low levels. Computational models of competition for object status focus on competition occurring between low-level features, whereas our data supports proposals that high-level properties such as shape familiarity enter into the competition as well.

Acknowledgements

The authors would like to thank Sarah Shomstein for her valuable input in the development of Experiment 2. MAP acknowledges support from the Office of Naval Research (N00014-14-1-0671) while working on this article.

References

Appelbaum, L. G., Wade, A. R., Vildavski, V. Y., Pettet, M. W., & Norica, A. M. (2006). Cue-invariant networks for figure and background processing in human visual cortex. *Journal of Neuroscience, 26(45),* 11695-11708.

Appelbaum, L. G., Wade, A. R., Pettet, M. W., Vildavski, V. Y., & Norcia, A. M. (2008). Figure–ground interaction in the human visual cortex. *Journal of Vision,8(*9), 8.

Beck, D. M., & Kastner, S. (2005). Stimulus context modulates competition in human extrastriate cortex. *Nature neuroscience, 8(*8), 1110-1116.

Beck, D. M., & Kastner, S. (2007). Stimulus similarity modulates competitive interactions in human visual cortex. *Journal of Vision, 7*(2), 19.

Beck, D. M., & Kastner, S. (2009). Top-down and bottom-up mechanisms in biasing competition in the human brain. *Vision research*, *49*(10), 1154-1165.

Bles, M., Schwarzbach, J., De Weerd, P., Goebel, R., & Jansma, B. (2006). Receptive field size-dependent attention effects in simultaneously presented stimulus displays. *Neuroimage, 30,* 506-511.

Bullier, J. (2001). Integrated model of visual processing. *Brain Research Reviews, 36,* 96-107.

Cacciamani, L., Mojica, A. J., Sanguinetti, J. L., & Peterson, M. A. (2014). Semantic access occurs outside of awareness for the ground side of a figure.*Attention, Perception, & Psychophysics*, *76*(8), 2531-2547.

Craft, E., Schutze, H., Niebur, E., & von der Heydt, R. (2007). A neural model of figure-ground organization. *Journal of neurophysiology*, *97*(6), 4310-4326.

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage, 9, 179-194.*

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual review of neuroscience*, *18*(1), 193-222.

Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis. II. Inflation, flattening, and a surface-based coordinate system. *Neuroimage, 9, 195-207.*

Fischl, B., Liu, A., & Dale, A.M. (2001). Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex. *IEEE Trans Med Imaging, 20, 70-80.*

Folk, C. L., Leber, A. B., & Egeth, H. E. (2002). Made you blink! Contingent attentional capture produces a spatial blink. *Perception & psychophysics*,*64*(5), 741-753.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human perception and performance*, *18*(4), 1030.

Gibson, B. S., & Peterson, M. A. (1994). Does orientation-independent object recognition precede orientation-dependent recognition? Evidence from a cueing paradigm. *Journal of Experimental Psychology: Human Perception and Performance, 20, 299-316.*

Grossberg, S. (1994). 3-D vision and figure-ground separation by visual cortex. *Perception & psychophysics, 55, 48-121.*

Jenkinson, M., Bannister, P. R., Brady, J. M., Smith, S. M. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage, 17, 825-841.*

Jiang, Y., & Chun, M. M. (2001). Selective attention modulates implicit learning.*The Quarterly Journal of Experimental Psychology: Section A*, *54*(4), 1105-1124.

Kastner, S., de Weerd, P., Desimone, R., & Ungerleider, L. G. (1998). Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. *Science*, *282*(5386), 108-111.

Kienker, P. K., Sejnowski, T. J., Hinton, G. E., & Schumacher, L. E. (1986). Separating figure from ground with a parallel network. *Perception, 15, 197-216.*

Kogo, N., Strecha, C., Van Gool, L., & Wagemans, J. (2010). Surface construction by a 2-D differentiation–integration process: A neurocomputational model for perceived border ownership, depth, and lightness in Kanizsa figures.*Psychological review*, *117*(2), 406.

Lamme, V. A. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *The Journal of Neuroscience*, *15*(2), 1605-1615.

Lamme, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neuroscience, 23,* 571-579.

Lee, T. S., & Mumford, D. (2003). Hierarchical Bayesian inference in the visual cortex. *JOSA A*, *20*(7), 1434-1448.

Likova, L. T., & Tyler, C. W. (2008). Occipital network for figure/ground organization. *Experimental Brain Research, 189, 257-267.*

Ling, S., & Blake, R. (2012). Normalization regulates competition for visual awareness. *Neuron*, *75*(3), 531-540.

Logothetis, N. K. (2002). The neural basis of the blood–oxygen–level–dependent functional magnetic resonance imaging signal. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*,*357*(1424), 1003-1037.

Logothetis, N. K., & Wandell, B. A. (2004). Interpreting the BOLD signal. *Annu. Rev. Physiol.*, *66*, 735-769.

Luck, S. J., Chelazzi, L., Hillyard, S. A., & Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *Journal of neurophysiology*, *77*(1), 24-42.

MacEvoy, S. P., & Epstein, R. A. (2009). Decoding the representation of multiple simultaneous objects in human occipitotemporal cortex. *Current Biology*, *19*(11), 943-947.

Miller, E. K., Gochin, P. M., & Gross, C. G. (1993). Suppression of visual responses of neurons in inferior temporal cortex of the awake macaque by addition of a second stimulus. *Brain research*, *616*(1), 25-29.

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, *229*(4715), 782-784.

Nadel, L., & Peterson, M. A. (2013). The hippocampus: Part of an interactive posterior representational system spanning perceptual and memorial systems. *Journal of Experimental Psychology: General, 142*(4), 1242-1254.

Newman, D. P., O'Connell, R. G., & Bellgrove, M. A. (2013). Linking time-on-task, spatial bias and hemispheric activation asymmetry: a neural correlate of rightward attention drift. *Neuropsychologia*, *51*(7), 1215-1223.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive psychology*, *19*(1), 1-32.

Peterson, M. A., & Cacciamani, L. (2013). Toward a dynamical view of object perception. In *Shape Perception in Human and Computer Vision* (pp. 443-457). Springer, London.

Peterson, M. A., Cacciamani, L., Mojica, A. J., & Sanguinetti, J. L. (2012). Meaning can be accessed for the ground side of a figure. *Gestalt Theory*,*34*(3/4).

Peterson, M. A., Cacciamani, L., Barense, M. D., & Scalf, P. E. (2012). The perirhinal cortex modulates V2 activity in response to the agreement between part familiarity and configuration familiarity. *Hippocampus*, *22*(10), 1965-1977.

Peterson, M. A., & Enns, J. T. (2005). The edge complex: Implicit memory for figure assignment in shape perception. *Perception & psychophysics*, *67*(4), 727-740.

Peterson, M. A., Harvey, E. M., & Weidenbacher, H. J. (1991). Shape recognition contributions to figure-ground reversal: which route counts?. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(4), 1075.

Peterson, M. A., & Kim, J. H. (2001). On what is bound in figures and grounds.*Visual Cognition*, *8*(3-5), 329-348.

Peterson, M. A., & Lampignano, D. W. (2003). Implicit memory for novel figure-ground displays includes a history of cross-border competition. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(4), 808.

Peterson, M. A., & Skow, E. (2008). Suppression of shape properties on the ground side of an edge: Evidence for a competitive model of figure assignment. *Journal of Experimental Psychology: Human Perception and Performance, 34, 251-267.*

Peterson, M. A., & Skow-Grant, E. (2003). Memory and learning in figure–ground perception. *Psychology of Learning and Motivation*, *42*, 1-35.

Poort, J., Raudies, F., Wannig, A., Lamme, V. A., Neumann, H., & Roelfsema, P. R. (2012). The role of attention in figure-ground segregation in areas V1 and V4 of the visual cortex. *Neuron, 75*(1), 143-156.

Qiu, F. T., Sugihara, T., & von der Heydt, R. (2007). Figure-ground mechanisms provide structure for selective attention. *Nature neuroscience,10(11*), 1492-1499.

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature neuroscience*, *2*(1), 79-87.

Reddy, L., & Kanwisher, N. (2007). Category selectivity in the ventral visual pathway confers robustness to clutter and diverted attention*.* *Current Biology,17*(23), 2067-2072.

Remington, R. W., Johnston, J. C., & Yantis, S. (1992). Involuntary attentional capture by abrupt onsets. *Perception & Psychophysics, 51*(3), 279-290.

Reuter-Lorenz, P. A., Kinsbourne, M., & Moscovitch, M. (1990). Hemispheric control of spatial attention*.* *Brain and cognition, 12(2*), 240-266.

Reynolds, J. H., Chelazzi, L., & Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. *The Journal of Neuroscience*, *19*(5), 1736-1753.

Roelfsema, P. R., Lamme, V. A., Spekreijse, H., & Bosch, H. (2002). Figure—ground segregation in a recurrent network architecture. *Journal of Cognitive Neuroscience*, *14*(4), 525-537.

Rolls, E. T., & Tovee, M. J. (1995). Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. *Journal of Neurophysiology*,*73*(2), 713-726.

Salvagio, E., Cacciamani, L., & Peterson, M. A. (2012). Competition-strength-dependent ground suppression in figure–ground perception. *Attention, Perception, & Psychophysics*, *74*, 964-978.

Sanguinetti, J. L., Allen, J. J. B., & Peterson, M. A. (2014). The ground side of an object: perceived as shapeless yet processed for semantics. *Psychological Science.*

Sanguinetti, J. L., Trujillo, L. T., Schnyer, D. M., Allen, J. J., & Peterson, M. A. (2014). Increased alpha band activity indexes inhibitory competition across a border during figure assignment. *Journal of Vision*, *14*(10), 49-49.

Scalf, P. E., Basak, C., & Beck, D. M. (2011). Attention does more than modulate suppressive interactions: attending to multiple items. *Experimental brain research*, *212*(2), 293-304.

Scalf, P. E., & Beck D. M. (2010). Competition in visual cortex impedes attention to multiple items. *Journal of Neuroscience, 30,* 161-169.

Sejnowski, T. J., & Hinton, G. E. (1987). Separating figure from ground with a Boltzmann machine. In *In MA Arbib & A. Hanson (Eds.), Vision, brain, and cooperative computation*. MIT Press, Cambridge, pp 704-724.

Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., Rosen, B. R., & Tootell, R. B. H. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science, 268, 889-893.*

Serre, T., Oliva, A., & Poggio, T. (2007). A feedforward architecture accounts for rapid categorization. *Proceedings of the National Academy of Sciences,104*(15), 6424-6429.

Shipp, S., Watson, J. D. G., Frackowiak, R. S. J., & Zeri, S. (1995). Retinotopic maps in human prestriate visual cortex: the demarcation of areas V2 and V3. *Neuroimage*, *2*(2), 125-132.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M., & Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage, 23, 208-219.*

Snowden, R. J., Treue, S., Erickson, R. G., & Andersen, R. A. (1991). The response of area MT and V1 neurons to transparent motion. *The Journal of Neuroscience*, *11*(9), 2768-2785.

Straw, A. D. (2008). Vision Egg: An Open-Source Library for Realtime Visual Stimulus Generation. *Frontiers in Neuroinformatics. doi: 10.3389/neuro.11.004.2008*

Strother, L., Lavell, C., & Vilis, T. (2012). Figure-ground representation and its decay in primary visual cortex. *Journal of Cognitive Neuroscience, 24:4, 905-914.*

Takio, F., Koivisto, M., Tuominen, T., Laukka, S. J., & Hämäläinen, H. (2013). Visual rightward spatial bias varies as a function of age. *Laterality: Asymmetries of Body, Brain and Cognition*, *18*(1), 44-67.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), 520-522.

Trujillo, L. T., Allen, J. J., Schnyer, D. M., & Peterson, M. A. (2010). Neurophysiological evidence for the influence of past experience on figure–ground perception. *Journal of Vision, 10, 1-21.*

Tsotsos, J. K., Culhane, S. M., Kei Wai, W. Y., Lai, Y., Davis, N., & Nuflo, F. (1995). Modeling visual attention via selective tuning. *Artificial Intelligence, 78(1-2), 507-545.*

Vecera, S. P., & O'reilly, R. C. (1998). Figure-ground organization and object recognition processes: an interactive account. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(2), 441.

Watanabe, M., Cheng, K., Murayama, Y., Ueno, K., Asamizuya, T., Tanaka, K., & Logothetis, N. (2011). Attention but not awareness modulates the BOLD signal in the human V1 during binocular suppression. *Science*, *334*(6057), 829-831.

Woolrich, M. W., Ripley, B. D., Brady, J. M., & Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of FMRI data. *NeuroImage,* 14, 1370-1386.

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: voluntary versus automatic allocation. *Journal of Experimental Psychology: Human perception and performance*, *16*(1), 121.

Zhang, P., Jamison, K., Engel, S., He, B., & He, S. (2011). Binocular rivalry requires visual attention. *Neuron*, *71*(2), 362-369.

Zhou, H., Friedman, H. S., & Von Der Heydt, R. (2000). Coding of border ownership in monkey visual cortex. *The Journal of Neuroscience*, *20*(17), 6594-6611.

Zoccolan, D., Cox, D. D., & DiCarlo, J. J. (2005). Multiple object response normalization in monkey inferotemporal cortex. *The Journal of Neuroscience*,*25*(36), 8150-8164.

Figure Captions

Figure 1: Sample novel silhouettes used in the current experiment (see also Peterson & Kim, 2001; Peterson & Skow, 2008; Salvagio et al., 2012; Trujillo et al., 2010). (A) *Low-competition* silhouettes which suggest novel, meaningless shapes on the groundsides. (B) *High-competition* silhouettes which suggest portions of familiar, real-world objects on their groundsides. The familiar objects suggested are (from left to right) seahorses, maple leaves, and table lamps.

Figure 2: Trial structure for one block. Shown here is a block of *high-competition* silhouettes in the right visual field. Stimuli have been enlarged for illustrative purposes. ISI = interstimulus interval.

Figure 3: Functional localization of the ground region. (A) The vertical dashed line demarcates the silhouettes' closest edge to fixation in each visual field (this was the same for every silhouette). The green area in each visual field is the region localized as the ground; voxels activated by a Gabor presented in this location were included in the ground ROI. Voxels activated by a Gabor presented in the red area were removed from the ground ROIs. The lines, the color coding, and the silhouette are shown here for illustrative purposes; in the localizer scans, only the fixation cross and a Gabor patch located in the green or the red area were presented. (B) Ground ROIs (*g*) in each visual area resulting from the localization procedure in (A) in one participant. LVF = left visual field, RVF = right visual field, LH = left hemisphere.

Figure 4: Experiment 1 ground ROI analysis results. Mean parameter estimates extracted from the ground region of *low-competition* (red) and *high-competition* (blue) silhouettes for contralateral and ipsilateral stimulus presentation. (A-B) Data collapsed across visual area; (C-D) Separated by visual area. Error bars represent standard error of the mean of the difference scores (*low-competition – high-competition*). Contra = contralateral visual field presentation; Ipsi = ipsilateral visual field presentation. * $p < .05$

Figure 5: Individual participants' parameter estimates from the ground ROI analysis in Experiment 1. Shown are difference scores (low-competition – high-competition) for both contralateral (green) and ipsilateral (light purple) presentation. Error bars represent standard error of the mean of the difference scores. * p < .05

Figure 6: Experiment 1 figure ROI analysis results. Mean parameter estimates extracted from the figure region of *low-competition* (red) and *high-competition* (blue) silhouettes for contralateral and ipsilateral stimulus presentation. (A-B) Data collapsed across visual area; (C-D) Data separated by visual area. Note that the y-axis scale here differs from those shown in Figure 4 (ground ROIs). Error bars represent standard error of the mean of the difference scores (*low-competition – high-competition*). Contra = contralateral visual field presentation; Ipsi = ipsilateral visual field presentation.

Figure 7: Results of Experiment 2. (top) Mean proportion correct on the RSVP task as a function of condition (baseline and test trials), and within the latter, silhouette type and visual field presentation. (bottom) Proportion correct on test trials subtracted from proportion correct on baseline trials. Values above 0 indicate impaired performance on test trials; values below baseline indicate enhanced performance on test trials. Error bars represent standard error of the mean of the difference scores. LVF = left visual field; RVF = right visual field. LH = left hemisphere; RH = right hemisphere. * $p < .05$

Appendix A

List of familiar objects suggested along the groundside of the *high-competition* silhouettes, listed alphabetically:

Anchor
Axe
Bear
Bell
Butterfly
Bone
Boot
Bunny
Coffee pot
Dog
Duck
Eagle
Elephant
Face
Faucet
Flower
Foot
Grapes
Guitar
Hand
Horn
House
Hydrant
Jet
Lamp
Leaf
Mickey mouse
Owl
Palm tree
Pig
Pineapple
Rhino
Seahorse
Snowman
Spray bottle
Umbrella
Train
Watering can
Woman
Wrench