

TMS to the Lateral Occipital Cortex Disrupts Object Processing but Facilitates Scene Processing

Caitlin R. Mullin and Jennifer K. E. Steeves

Abstract

■ The study of brain-damaged patients and advancements in neuroimaging have led to the discovery of discrete brain regions that process visual image categories, such as objects and scenes. However, how these visual image categories interact remains unclear. For example, is scene perception simply an extension of object perception, or can global scene “gist” be processed independently of its component objects? Specifically, when recognizing a scene such as an “office,” does one need to first recognize its individual objects, such as the desk, chair, lamp, pens, and paper to build up the representation of an “office” scene? Here, we show that temporary interruption of object processing through repetitive TMS to the left lateral occipital cortex (LO), an area known to selectively process objects, impairs object categorization but surprisingly facilitates

scene categorization. This result was replicated in a second experiment, which assessed the temporal dynamics of this disruption and facilitation. We further showed that repetitive TMS to left LO significantly disrupted object processing but facilitated scene processing when stimulation was administered during the first 180 msec of the task. This demonstrates that the visual system retains the ability to process scenes during disruption to object processing. Moreover, the facilitation of scene processing indicates disinhibition of areas involved in global scene processing, likely caused by disrupting inhibitory contributions from the LO. These findings indicate separate but interactive pathways for object and scene processing and further reveal a network of inhibitory connections between these visual brain regions. ■

INTRODUCTION

The human visual system has the ability to rapidly and accurately categorize complex scenes (Peelen, Fei-Fei, & Kastner, 2009). Typically, a scene is defined as visual information about the immediate environment, often including a combination of both background elements and one or more discrete objects (Henderson & Hollingsworth, 1999). There have been two broad approaches to how we recognize and categorize the multitude of scenes in our visual environment.

Early theories of scene perception emphasized an object-centered approach. In this view, the recognition of a real-world scene emerges from first processing a set of objects contained within it in a bottom-up fashion. Once the component objects are identified, the meaning or gist of the scene is thought to arise on the basis of the arrangement and co-occurrence of those objects (De Graef, Christaens, & d’Ydewalle, 1990; Biederman, 1981, 1987; Freidman, 1979). Additionally, the identification of one or more prominent objects has been thought to be sufficient to activate a scene schema and, thus, facilitate recognition and categorization (Biederman, 1981; Freidman, 1979).

A recent alternative approach to scene perception maintains that the gist of a scene can be processed in a

top-down manner, using global image properties, without the need for first identifying its component objects (Oliva & Torralba, 2001, 2006; Walker Renninger & Malik, 2004; Torralba & Oliva, 2003). Behavioral experiments have suggested that the semantic category of most real-world scenes can be inferred from their spatial layout (Sanocki & Epstein, 1997; Biederman, 1995). Scenes can be identified from low spatial frequency (SF) images that retain the spatial relationship between large-scale structures in the scene but which lack the visual detail required to identify local objects (Schyns & Oliva, 1994). This global approach to scene recognition has been modeled using a variety of spatial properties defined independently of objects within a scene (Oliva & Torralba, 2001). This model, the *Spatial Envelope Model*, proposes a set of perceptual dimensions that represent the dominant spatial structure of a scene.

Further work on this model has yielded support for the theory that rapid categorization of scenes may not be mediated primarily through parts and objects, but rather through global properties (Greene & Oliva, 2006, 2009; Oliva & Torralba, 2001, 2006; Torralba & Oliva, 2003). For instance, natural landscapes tend to have areas with undulating contours and characteristic textures and/or colors, whereas manmade scenes tend to have areas composed of straight cardinal lines (Oliva & Torralba, 2001; Burton & Moorhead, 1987). Global properties reflecting scene

structure, layout, and function could act as primary features for scene processing, and therefore, global aspects of a scene may be processed before (or perhaps in parallel with) the identification of individual objects.

Neuroimaging studies have demonstrated that discrete cortical areas are selectively active when viewing objects and scene images. This includes the lateral occipital cortex (LO; Grill-Spector, Kourtzi, & Kanwisher, 2001) and the parahippocampal gyrus or parahippocampal place area (Epstein & Kanwisher, 1997), respectively. Is object processing required for scene processing? Do these areas comprise a single hierarchical stream for object and scene processing? A study of a patient with bilateral damage to area LO resulting in profound object agnosia, nonetheless, retains a normal ability to categorize scenes, which suggests that object and scene processing are not part of a single hierarchical stream (Steeves et al., 2004). This patient uses global image features, such as color and texture, to categorize scenes despite an inability to visually recognize objects. Furthermore, despite the patient's damage to object-selective cortex fMRI revealed activation in the parahippocampal place area that was modulated by color and texture (Steeves et al., 2004). This single-case study suggests that scene categorization can operate independently of object perception, perhaps in two separate parallel pathways, and that object processing area LO is not required for this ability.

In the following set of experiments, we used repetitive TMS (rTMS) to temporarily disrupt object processing in area LO to examine its contribution to the perception of scenes in the healthy brain. According to the scene-centered model, rTMS to area LO will disrupt object processing but show no effect on the ability to rapidly and accurately categorize scenes; alternatively, according to the object-centered approach where scene categorization is built up from the objects within it, rTMS to area LO will compromise the ability to categorize a scene.

EXPERIMENT 1

Participants

Nine healthy volunteers (six men and three women, age = 26–39 years) participated in all four experiment conditions, including fMRI, to localize area LO. All participants were in good health with normal or corrected-to-normal vision and, according to self-report, had no known contraindications to rTMS or fMRI. Informed consent was obtained, and the experiment was conducted in accordance with the York University Office of Research Ethics, which follows the guidelines outlined by the Declaration of Helsinki.

fMRI Acquisition

Cortical regions selective to objects were localized using fMRI. Scanning was conducted on a 1.5-T GE scanner at the Hospital for Sick Children (Toronto, Canada) using BOLD

imaging and functional volumes were acquired using the GE 8 Channel high-resolution brain array coil. High-resolution anatomical images were acquired (T1-weighted fast spoiled gradient-echo, in-plane resolution = 0.9375×0.9375 mm, 120 axial slices, slice thickness = 1.5 mm, imaging matrix = 256×192 using the square pixel imaging option, field of view = 24×24 cm, echo time = 4.2 msec, repetition time = 9.0 msec, flip angle = 20°). Functional volumes were acquired with EPI (in-plane resolution = 3.75×3.75 mm, slice thickness = 3 mm, imaging matrix = 64×64 , field of view = 24×24 cm, axial slices, repetition time = 2 sec, echo time = 40 msec, flip angle = 90°). Stimuli were presented with a rear-projection system (Avotec, Inc., Stuart, FL) in two separate runs (6 min 52 sec). Functional localizer scans used a one-back paradigm to focus attention on the three categories of visual stimuli: faces, scenes, and objects. Each run began and finished with a fixation cross for 16 sec. Six repetitions of three 16-sec blocks of the three categories of stimuli were presented in a pseudorandom order. Each repetition was interleaved with 16 sec of fixation. Each block contained 16 stimuli presented for 1 sec each. Participants were imaged over two functional runs.

fMRI Analysis

fMRI data were analyzed using BrainVoyager QX (Brain Innovation, Maastricht, The Netherlands) and applying a general linear model. Functional data sets were subjected to a series of preprocessing operations consisting of linear trend removal to exclude scanner-related signal drift, temporal high-pass filtering to remove temporal frequencies lower than three cycles per run, and a correction for small interscan head movements using a rigid body algorithm rotating and translating each functional volume in 3-D space. Each participant's functional images were registered with their anatomical images, and functional data were averaged across the two runs.

A linear balanced contrast of objects versus faces and scenes was used to identify area LO, the rTMS target site within each participant. Area LO was individually identified in each participant by determining the peak object-selective activation in the lateral occipital region of the left hemisphere (lLO; see Figure 1). Whereas a number of studies have shown no hemispheric advantage to object processing (Fize, Fabre-Thorpe, Richard, Doyon, & Thorpe, 2005; Biederman & Cooper, 1991), others have demonstrated a left hemisphere advantage (Zwaan & Yaxley, 2004; Laeng, Shah, & Kosslyn, 1999; Marsolek, 1995). The majority of our participants showed greater activation in the left hemisphere than the right to images of objects. We, therefore, restricted analysis to the left hemisphere.¹

TMS Functional Stereotaxy

Despite the fact that basic organizational patterns of visual areas may be consistent from person to person,

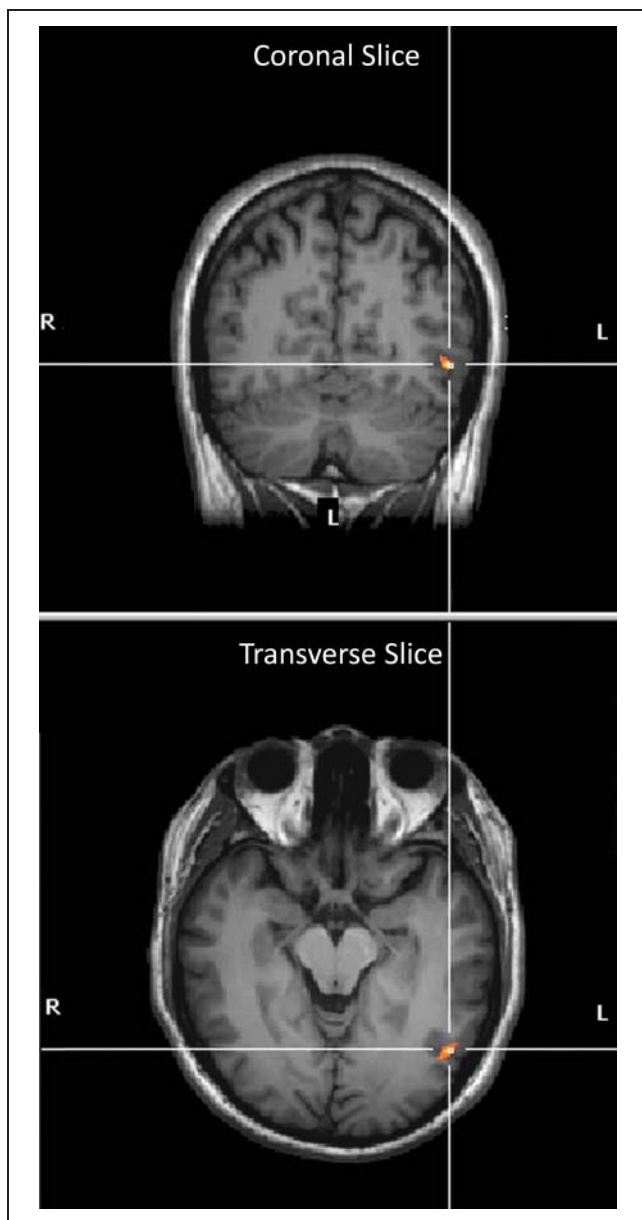


Figure 1. Ventral cluster of object-selective cortex for a typical participant in coronal and transverse views. The crosshairs in each image indicate the peak object-selective functional activation (objects vs. [scenes + faces] contrast) in the lLO region of cortex.

it is well known that the size and anatomical location of visual areas can vary widely across participants. These variations are not ones that can be easily overcome by using anatomical landmarks on an individual's scalp. The essential step is to provide a spatial framework within which areas of interest are specified for each individual. For this reason, we use image-guided TMS with functionally defined target sites.

The functionally defined stimulation sites were localized with Brainsight image-guided coregistration software and hardware (Rogue Research, Montreal, Canada), utilizing individual MRI scans for each participant. Common reference points on both the MRI images and the participant's head

were selected to create a registration matrix. The spatial relationship between these reference points on the MRI images and those on the participant's head were coregistered using a Polaris infrared marker system. The stimulation target site, area LO, was selected for each participant by overlaying his or her activation map from the fMRI object localizer experiment (see above) onto a 3-D reconstruction of the participants brain and scalp within the Brainsight software. Subsequently, image-guided TMS stimulation was achieved by monitoring, in real time, the location and orientation of the TMS coil and targeted stimulation site via infrared markers on the coil and the participant's head.

TMS Stimuli and Experimental Procedure

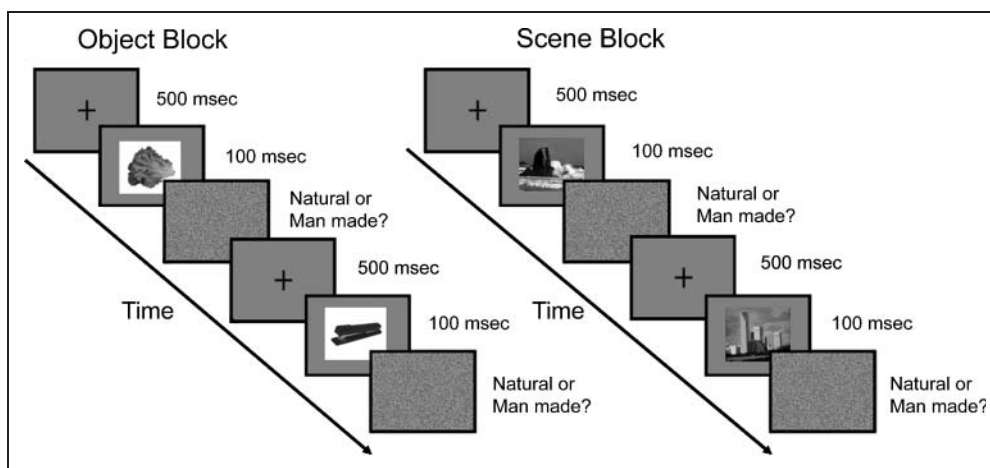
The experimental design consisted of a no-rTMS baseline measure followed by either the experimental condition of rTMS to LO or two control conditions: rTMS to the vertex to control for site specificity of rTMS effects or a sham condition where the coil was placed over the target stimulation site, LO, but was oriented perpendicular to the head so that no current was induced in the brain. Stimulus identity differed on each condition to avoid practice effects.

Participants were instructed to categorize grayscale photographs of objects and scenes as either "natural" (e.g., a leaf or a beach scene) or "manmade" (e.g., a spoon or an airport scene) as quickly as possible without sacrificing accuracy. Object stimuli were taken from the Bank of Standardized Stimuli (Brodeur, Dionne-Dostie, Montreuil, & Lepage, 2010), and scenes were from a photo image library (also used in Steeves et al., 2004; see Figure 2, e.g., natural and man-made stimuli). All stimuli were grayscale and resized using Adobe Photoshop CS2© version 9.0.2. Participants sat 60 cm from the display, and stimuli were presented centrally on a Dell 19-in. monitor at a visual angle of $9^\circ \times 13.5^\circ$. On each trial, a fixation cross appeared for 500 msec at the center of the display followed by a stimulus image for 100 msec, followed immediately by a mask consisting of static noise pattern that remained on screen until participants responded (see Figure 2). Stimuli were presented in two blocks (objects and scenes) of 20 images each, and block order was counterbalanced across participants. Participants categorized the stimuli by pressing designated keys on a response box (Cedrus, Inc.). Four different test versions were created for the pre-rTMS, post-rTMS, vertex rTMS, and sham rTMS conditions to prevent practice effects, and these different test versions were randomized across stimulation conditions. The baseline condition was always measured first, and participants underwent each of the three remaining conditions in counterbalanced order.

TMS Stimulation

A Magstim Super Rapid 2 stimulator was used to deliver rTMS via a 70-mm fan cooled figure-of-eight coil held in position on the scalp surface by an articulated coil stand

Figure 2. Schematic overview of trial sequence in Experiment 1. Example of two trials (one natural and one man made) for each of the object and scene blocks. Each trial began with a central fixation cross for 500 msec, followed by a stimulus for 100 msec, which was then masked by a static noise pattern that was present until participants responded.



(Magstim; Whitland, UK). The coil was held tangential to the scalp surface with the handle pointed downward. The center of the coil was continually monitored with Brainsight to maintain its position over the target site of interest. rTMS was delivered at low frequency (1 Hz) at 60% of maximal stimulator output similar to a number of previous studies (e.g., Pitcher, Walsh, Yovel, & Duchaine, 2007; Silvanto, Lavie, & Walsh, 2005; Campana, Cowey, & Walsh, 2002) for a total of 420 pulses (7 min). This allowed for several minutes of disruption to the targeted area (Pascual-Leone et al., 1998; Chen et al., 1997), during which time the categorization task was completed, which took approximately 1 min 45 sec. The frequency, intensity, and duration of the rTMS train were well within the safety limits of stimulation (Rossi, Hallett, Rossini, & Pascual-Leone, 2009; Wassermann, 1998). Earplugs were provided to dampen the noise associated with the discharge from the rTMS coil. The vertex was defined as a point midway between the inion and the nasion and equidistant from the left and right intertrigular notches. Fifteen-minute breaks were given after each block.

Results

Accuracy of object discrimination was impaired when rTMS was delivered over ILO but not during stimulation of an unrelated area (the vertex) or during the sham condition (Figure 3). In contrast, scene discrimination was facilitated when rTMS was delivered to ILO but not during the vertex or sham conditions. A 2×4 repeated measures ANOVA of Stimuli (scenes and objects) and Stimulation Condition (baseline, experimental ILO, vertex and sham) revealed a significant interaction [$F_{3, 24} = 12.48, p < .001$]. False discovery rate (FDR) post hoc analysis for object discrimination revealed a significant reduction in accuracy during rTMS to ILO relative to the no-rTMS baseline ($p = .012$), vertex ($p = .021$), and sham conditions ($p = .038$). For scene discrimination, accuracy improved during rTMS to ILO relative to no-rTMS ($p = .009$), vertex ($p < .006$), and sham conditions ($p = .032$). Importantly, the no-rTMS, vertex, and sham conditions were not sig-

nificantly different from each other for either scene or object categories ($ps > .05$).

Mean RTs for objects and scenes were 1105 and 1079 msec, respectively. No significant differences in RT were observed across stimulus conditions ($p > .05$), which is consistent with other similar TMS studies (e.g., Pitcher, Charles, Devlin, Walsh, & Duchaine, 2009).

Discussion

We used rTMS to investigate the contributions of object processing during scene perception. The resistance of scene perception during the temporary interruption of object processing demonstrates that scene perception does not simply rely on the initial identification of objects for categorization to occur. This result confirms the scene-centered approach, that scene processing does not merely operate through a straightforward hierarchical system with object perception as a necessary precursor.

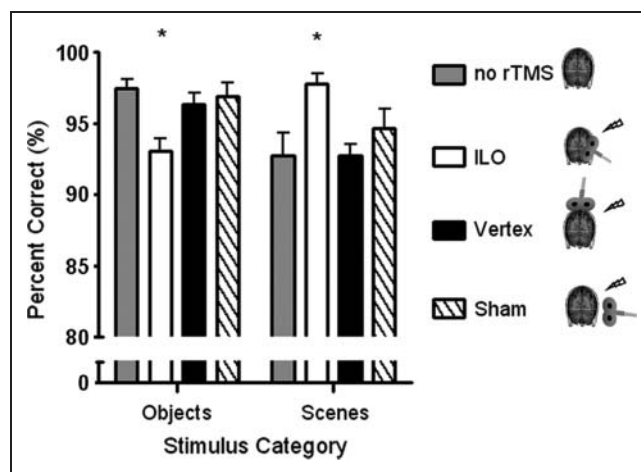


Figure 3. Accuracy scores for categorization tasks performed under each of the four stimulation conditions: no-rTMS, ILO, vertex, and sham. * indicates a significant difference from all other conditions within the same stimulus category (FDR, $p < .05$). The Y axis is truncated to better illustrate the significant effect of the post-rTMS condition to cortical area ILO compared with the other conditions. Chance performance was 50%.

An unanticipated finding was the significant facilitation of scene processing during the interruption of object processing. Our initial hypothesis was that if the scene-centered theory held, scene processing would remain unaffected when rTMS was delivered to area LO, because these operations likely operate in parallel. This facilitation in scene processing may be the result of a release of inhibitory interactions on scene pathways from object pathways.

It has been suggested that different areas of visual cortex compete with each other for limited resources such as blood flow and pathway access to other brain regions (Walsh, Ellison, Battelli, & Cowey, 1998). As a result, the facilitatory effect of rTMS seen here may be because of a release of inhibition from area ILO on scene processing pathways following the disruption to object-selective cortex. Our psychophysiological data in the present study are consistent with previous perceptual theories of object and scene processing. Biederman (1981) outlined a number of pathways by which scene information could be perceptually computed—one involving the identification of prominent objects, whereas another involved processing global aspects of what he termed “scene emergent features.” This work was later extended when Oliva and colleagues examined the dissociation of the global representation of scenes from its local parts or objects (Greene & Oliva, 2009; Torralba & Oliva, 2003; Oliva & Torralba, 2001), concluding that a global scene-centered pathway is a plausible coding mechanism of visual scenes in the brain. This theory of scene processing was not meant to be an alternative to the object-centered approach but rather the global scene-centered pathway may act in parallel with an object-centered pathway creating a complementary system allowing for a fast and accurate representation of a scene (Oliva & Torralba, 2006). Our data support this view with the addition of inhibitory connections between these pathways.

EXPERIMENT 2

The results of Experiment 1 indicated that rTMS to ILO differentially affected both object and scene categorization. Specifically, stimulation to the peak cortical area of selectivity for objects produced a significant reduction in object categorization and a significant increase in scene categorization performance compared with baseline. In Experiment 2, we sought to replicate the findings of Experiment 1 using an on-line rTMS technique, which, in addition, allows us to assess the temporal aspects of our original findings. The application of on-line rTMS to area LO will allow us to measure the time course of its effects on object and scene processing. In addition, we compared this effect in each of the right and the left hemispheres. Previous research has suggested that the two hemispheres may be tuned to different SFs (Robertson & Ivry, 2000).

This may contribute to the findings of Experiment 1, in that stimulation to the left hemisphere may disrupt processing of high SF, which could impede the perception of objects over scenes as scenes can be processed using low SF information alone (Schyns & Oliva, 1994).

Participants

Nine healthy volunteers (six men and three women, age = 27–40 years) participated in all conditions of Experiment 2, including fMRI to localize LO. Seven of them (four men and three women) participated in both Experiments 1 and 2.

fMRI Acquisition and Analysis/TMS Functional Stereotaxy

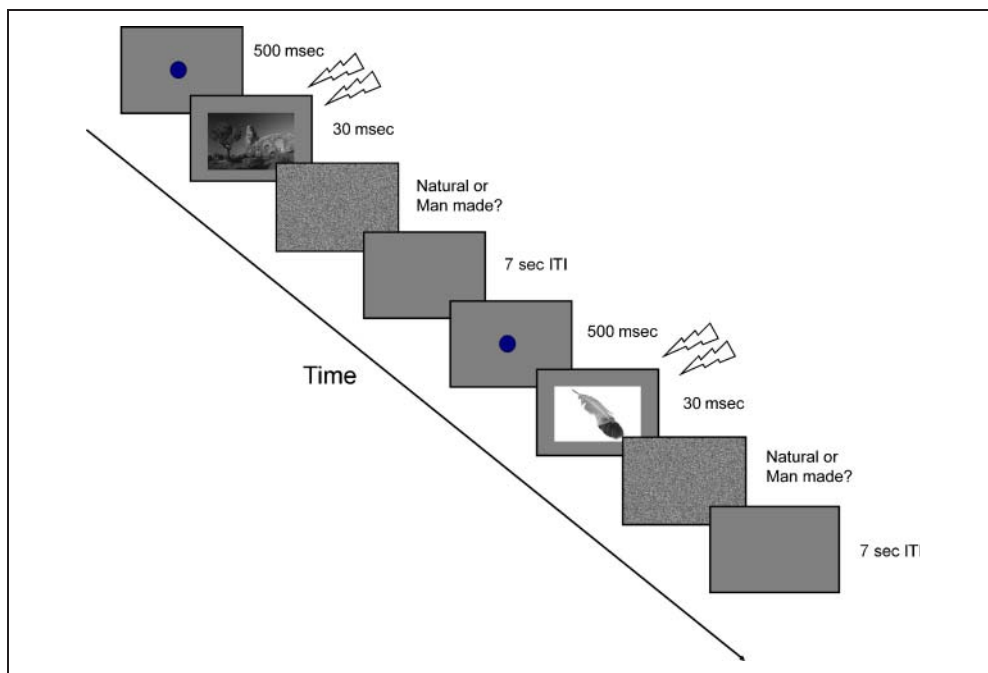
The acquisition of the fMRI data was identical in both experiments. The analysis of the neuroimaging data for Experiment 2 was virtually identical to that of Experiment 1, with the addition of localizing the peak object-selective activation in both hemispheres. Functional stereotaxy was identical to Experiment 1.

TMS Stimuli and Experimental Procedure

The experimental design consisted of a double pulse of rTMS, separated by 100 msec, to area LO in the left and right hemispheres at six different stimulation onset asynchrony (SOA) (0, 40, 80, 120, 180, and 220 msec) relative to stimulus presentation, allowing for the creation of short deficit time windows. Additionally, the baseline measure included trials of no stimulation termed “no-rTMS.” The seven stimulation conditions, including no-rTMS, were presented in random order. In addition, the experiment was also conducted in a separate run with stimulation to the vertex of each participant as a control condition. The order of stimulation was as follows: All participants were first tested in the left hemisphere, counterbalanced with stimulation to the vertex. Participants then returned for a session of stimulation to the right hemisphere counterbalanced with a second vertex condition. Data from the vertex conditions were collapsed across both testing sessions. Analysis of the no-rTMS condition revealed no significant differences from the participants’ first visit to their second (Objects: $F_{2, 16} = 0.254, p = .779$; Scenes: $F_{2, 16} = 0.267, p = .769$). Each experimental run (LO or vertex) consisted of 40 trials of each of the seven SOAs for 280 trials per stimulation site. Breaks were given every 70 trials.

The task was identical to that of Experiment 1. Participants were required to categorize grayscale photographs of objects and scenes as either “natural” or “manmade.” Object stimuli were taken from the Bank of Standardized Stimuli (Brodeur et al., 2010), and new scene stimuli were acquired from the SUN database (Xiao, Hays, Ehinger,

Figure 4. Schematic overview of trial sequence in Experiment 2. Example of two trials (one natural scene and one natural object). Each trial began with a central fixation dot for 500 msec, followed by a stimulus for 30 msec, which was then masked by a static noise pattern that was present until participants responded. This was followed by a 7-sec intertrial interval. A double-pulse of rTMS was delivered at one of six different time points between 0 and 220 msec poststimulus onset.



Oliva, & Torralba, 2010). Stimuli were rendered in gray-scale and resized in an identical manner to Experiment 1 (see Figure 4 for examples of stimuli). Stimulus identity differed between Experiments 1 and 2 to avoid practice effects as several participants took part in both experiments. Before the current experiment, all stimuli were tested on a group of 20 naive participants to determine whether all stimuli could be reliably categorized. No stimuli were miscategorized significantly more often than any others, indicating that they were reliable. Participants sat 60 cm from the display, and stimuli were presented centrally on a 19-in. Macintosh monitor at a visual angle of $9^\circ \times 13.5^\circ$. On each trial, a fixation dot appeared at the center of the display for 500 msec, followed by a stimulus image for 30 msec, followed immediately by a mask consisting of static noise pattern that remained on screen until participants responded (see Figure 4). Object and scene stimuli were presented randomly throughout the experiment.

TMS Stimulation

The equipment used to deliver rTMS and navigate the coil was identical to that of Experiment 1. VPixx custom software and Datapixx hardware (VPixx, Inc., Saint-Bruno, QC, Canada) externally triggered the Magstim Super Rapid 2 at the specified SOAs relative to the stimulus onset. A double pulse of TMS was delivered to the ILO and right LO (rLO) and the vertex at six different asynchronies relative to stimulus onset: 0 and 100 msec, 40 and 140 msec, 80 and 180 msec, 120 and 220 msec, 180 and 280 msec, and 220 and 320 msec. The pulses were delivered at a frequency of 10 Hz at 60% of maximal stimulator output.

The sequence of both the stimuli and the delays was randomized. The maximum duration of each trial was 1.5 sec, followed by a 7-sec intertrial interval. The frequency, intensity, and duration of the rTMS train were within the safety limits of stimulation (Rossi et al., 2009; Wassermann, 1998).

Results

We fit a linear mixed model, using the MIXED procedure of SPSS v19.0, with stimulation site (LO and vertex), SOA (0, 40, 80, 120, 180, 220, and no rTMS), and their interaction as fixed factors to our data. The model also contained random corrections to each fixed factor because of interparticipant variability. These random corrections were assumed to be multivariate normal with a standard “variance components” covariance structure. Fitting was performed separately for each hemisphere. Planned linear contrasts were then performed on the fixed factors of the model to test two hypotheses: (1) Performance at each SOA differed significantly from the no-rTMS baseline within stimulation site, and (2) performance at each SOA differed across stimulation sites.

Object Categorization Accuracy

Separate linear contrasts, with FDR corrections for multiple comparisons, were performed comparing stimulation to LO at each SOA to the no-rTMS baseline conditions within the same stimulation site. Stimulation to area LO in the left hemisphere revealed a significant effect of SOA, with the three earliest SOAs resulting in significantly worse performance compared with the no-rTMS baseline

[0 msec: $t_{96.157} = -4.592, p < .001$; 20 msec: $t_{96.157} = -4.161, p < .001$, 80 msec: $t_{96.157} = -3.157, p = .004$]. No significant differences were observed at the later three SOAs compared with the no rTMS baseline [120 msec: $t_{96.157} = -1.435, p = .186$; 180 msec: $t_{96.157} = -0.861, p = .391$; 220 msec: $t_{96.157} = -1.435, p = .186$] (see Figure 5A).

rTMS to the rLO shows a similar pattern of results with the three earliest SOAs showing a subtle drop in accuracy compared with the no-rTMS baseline condition; however, corrections for multiple comparisons reveal that only the 20-msec SOA is significantly different from no rTMS [0 msec: $t_{96.140} = -1.666, p = .198$; 20 msec: $t_{96.140} = -3.253, p = .012$; 80 msec: $t_{96.140} = -1.772, p = .198$]. No significant differences were observed in the later three SOAs [120 msec: $t_{96.140} = -0.741, p = .461$; 180 msec: $t_{96.140} = -1.402, p = .246$; 220 msec: $t_{96.140} = -1.296, p = .248$] (see Figure 6A).

Additional comparisons between the experimental site (LO) and the control site (vertex) revealed a similar pattern of results across the different SOAs. Performance on the three earliest SOAs was significantly interrupted when rTMS was delivered to the lLO in comparison with the vertex at the same SOA [0 msec: $t_{75.598} = -3.086, p = .018$; 20 msec: $t_{75.598} = -2.624, p = .030$; 80 msec: $t_{75.598} = -2.309, p = .048$]. No significant differences were observed at the later three SOAs compared with the vertex condition at the same SOAs [120 msec: $t_{75.598} = 0.380, p = .859$; 180 msec: $t_{75.598} = -0.365, p = .714$; 220 msec: $t_{75.598} = -0.031, p = .976$] (see Figure 5A).

rTMS to the rLO shows a similar pattern of results with the three earliest SOAs showing a subtle drop in accuracy compared with their vertex counterparts; however, no significant differences were observed [0 msec: $t_{79.915} = -1.074, p = .429$; 20 msec: $t_{79.915} = -2.355, p = .126$;

Figure 5. A and B. Accuracy scores as a function of each SOA with rTMS to lLO. (A) The performance for object categorization. (B) The performance for scene categorization. ** indicates a significant difference from the no-rTMS baseline condition and vertex counterpart (FDR, $p < .05$). * indicates marginally significant differences from the no-rTMS baseline condition and vertex counterpart (FDR, $p < .10$). Chance performance was 50%.

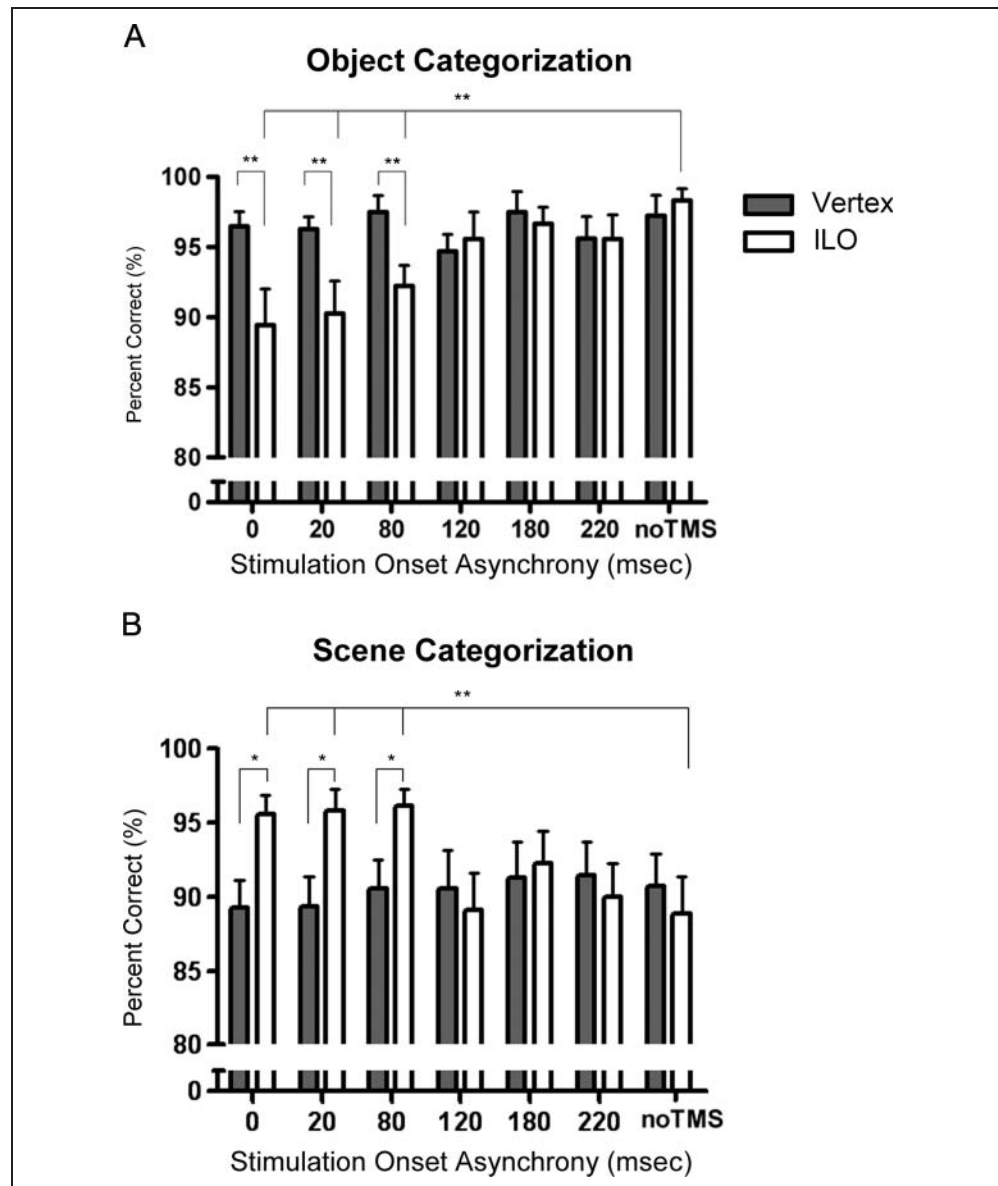
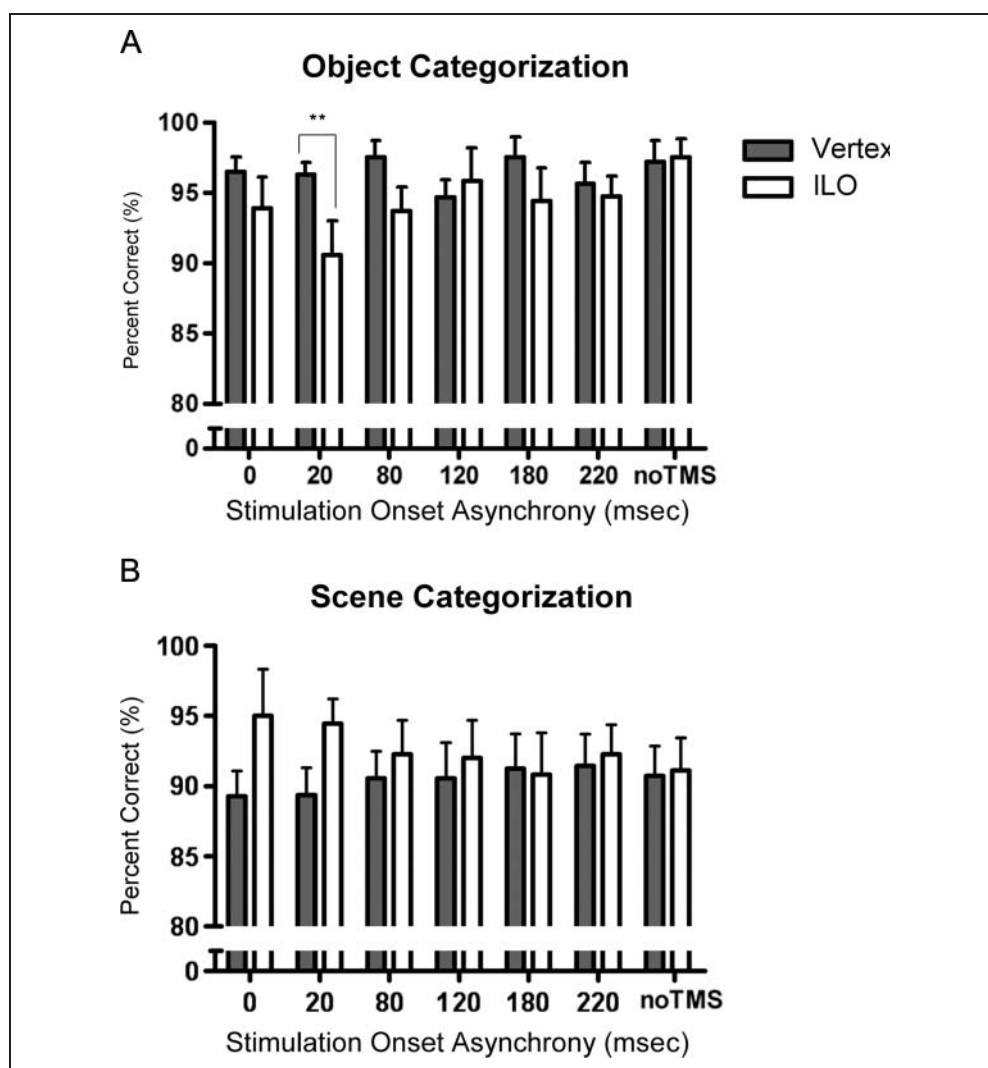


Figure 6. Accuracy scores as a function of each SOA with rTMS to rLO. (A) The performance for object categorization. (B) The performance for scene categorization. ** indicates a significant difference from the no-rTMS baseline condition and vertex SOA counterpart (FDR, $p < .05$). * indicates marginally significant differences from the no-rTMS baseline condition and vertex SOA counterpart (FDR, $p < .10$). Chance performance was 50%.



80 msec: $t_{79,915} = -1.577, p = .357$; 120 msec: $t_{79,915} = 0.471, p = .711$; 180 msec: $t_{79,915} = -1.275, p = .424$; 220 msec: $t_{79,915} = -0.372, p = .711$] (see Figure 6A).

Scene Categorization Accuracy

Separate linear contrasts, with FDR corrections for multiple comparisons, were performed, comparing stimulation to LO at each SOA to the no-rTMS baseline conditions within the same stimulation site. Stimulation to area LO in the left hemisphere revealed a significant effect of SOA, with the three earliest SOAs resulting in significantly improved performance compared with the no-rTMS baseline [0 msec: $t_{96,169} = 2.751, p = .014$; 20 msec: $t_{96,169} = 2.865, p = .014$; 80 msec: $t_{96,169} = 3.003, p = .014$]. No significant differences were observed at the later three SOAs compared with the no-rTMS baseline [120 msec: $t_{96,169} = 0.927, p = .138$; 180 msec: $t_{96,169} = 1.375, p = .206$; 220 msec: $t_{96,169} = 0.458, p = .648$] (see Figure 5B).

rTMS to the rLO shows a similar pattern of results with the three earliest SOAs showing an improvement in accuracy compared with the no-rTMS baseline condition;

however, no significant differences were observed across any of the SOAs [0 msec: $t_{96,160} = 1.487, p = .618$; 20 msec: $t_{96,160} = 1.274, p = .618$; 80 msec: $t_{96,160} = 0.446, p = .882$; 120 msec: $t_{96,160} = 0.340, p = .882$; 180 msec: $t_{96,160} = -0.106, p = .916$; 220 msec: $t_{96,140} = 0.446, p = .882$] (see Figure 6B).

Additional comparisons between the experimental site (LO) and the control site (vertex) revealed a similar pattern of results across the different SOAs. Performance on the three earliest SOAs reflected a marginally significant interruption when rTMS was delivered to the ILO in comparison with the vertex at the same SOA [0 msec: $t_{41,988} = 2.233, p = .090$; 20 msec: $t_{41,988} = 2.302, p = .090$; 80 msec: $t_{41,988} = 1.990, p = .100$]. No significant or marginally significant differences were observed at the later three SOAs compared with the vertex condition at the same SOAs [120 msec: $t_{41,988} = -0.514, p = .731$; 180 msec: $t_{41,988} = 0.346, p = .731$; 220 msec: $t_{41,988} = -0.514, p = .731$] (see Figure 5B).

rTMS to the rLO shows a similar pattern of results with the three earliest SOAs showing an increase in accuracy compared with their vertex counterparts; however, no

significant differences were observed [0 msec: $t_{54.53} = 1.692, p = .417$; 20 msec: $t_{54.53} = 1.503, p = .417$; 80 msec: $t_{54.53} = 0.509, p = .902$; 120 msec: $t_{54.53} = 0.427, p = .902$; 180 msec: $t_{54.53} = 0.509, p = .902$; 220 msec: $t_{54.53} = -0.246, p = .806$] (see Figure 6B).

Reaction Times

No significant differences in RT were observed across stimulation sites and SOAs ($ps > .05$), which is consistent with previous studies of this nature (Camprodon, Zohary, Brodbeck, & Pascual-Leone, 2010; Cohen Kadosh, Walsh, & Cohen Kadosh, 2010; Pitcher et al., 2007, 2009). Mean RTs with rTMS delivered to lLO were 619 msec for objects and 631 msec for scenes. With stimulation delivered to rLO mean RTs were 640 msec for objects and 629 msec for scenes. Stimulation to the vertex resulted in mean RTs of 622 msec for objects and 643 msec for scenes.

Discussion

We used rTMS to investigate the time course of the contributions of the LO in each hemisphere during object and scene categorization. We observed site-specific rTMS-induced object categorization impairments in the three earliest SOAs in the left hemisphere (see Figure 5), suggesting that the lLO's contribution to object processing takes place within the first 180 msec of stimulus onset. Although a similar pattern of results was observed in the right hemisphere, the effect of rTMS to rLO was more subtle and reached significance for a single comparison. This may be because of the fact that most of our participants exhibited greater object-selective activation in the left hemisphere. Although rTMS was being applied to rLO, the dominant left hemisphere was still intact, allowing object and scene processing to continue with relatively little disruption. Further potential explanations are considered in the General Discussion.

The effects of rTMS to lLO are consistent with previous research on the temporal dynamics of visual recognition. ERP studies have shown that the visual processing required for identification or categorization can be achieved in under 150 msec (Schendan, Ganis, & Kutas, 1998; Thorpe, Fize, & Marlot, 1996). Similar findings have been demonstrated with magneto-encephalography (Halgren, Raji, Marinkovic, Jousmäki, & Hari, 2000). However, more recent studies have suggested that contributions from area LO to categorical processing can take place much more rapidly. Murray et al. (2002) found a modulation of visual-evoked potentials at 88–100 msec over lateral-occipital cortex during object recognition.

Studies using rTMS to investigate the temporal dynamics of visual processing have reported visual suppression even earlier (20–60 msec) when applied to primary visual cortex (Camprodon et al., 2010; Corthout, Hallett, & Cowey, 2002; Corthout, Uttl, Walsh, Hallett, & Cowey, 1999). Such early suppression in primary visual cortex seems natural,

given the putative hierarchy of the cortical visual system. However, a recent review of multiple recording techniques in primates has shown that the onset latencies across visual areas are inconsistent with their placement within the putative anatomical hierarchy and that higher-order visual areas can become active at latencies similar to or even earlier than the primary visual cortex (Michel, Seeck, & Murray, 2004). This notion supports the current findings of impaired object categorization with early stimulation (0, 20, or 80 msec) to the lLO relative to stimulus onset.

In addition, with stimulation to the lLO, we observed a site-specific rTMS-induced improvement in scene categorization at the three earliest SOAs compared with the vertex control conditions and no-rTMS baseline condition (see Figure 5B). Comparable to the results for object categorization, stimulation to the rLO produced a similar pattern of behavioral results to that of the left hemisphere, but this did not reach significance. As with Experiment 1, this finding supports the scene-centered model, where intact object processing is not required to form global scene representations. In fact, as object processing is disrupted, scene categorization accuracy improves. The time course of facilitation of scene processing exactly mirrors the time course of the object disruption in a reciprocal manner. These results indicate an inhibitory relationship between the object and scene processing streams that can be unmasked using rTMS.

GENERAL DISCUSSION

The goal of the current experiments was to assess the contribution of object processing to scene processing through rTMS-induced disruption to the LO region. The initial results, using off-line TMS, which produces a longer window of disruption to lLO, found a disruption of object processing but a facilitation of scene processing. This is likely because of the unmasking of inhibitory connections between scene- and object-selective pathways. We replicated this effect using on-line double-pulse rTMS to assess the temporal aspects of our findings. We show a strong reciprocal connection between object and scene processing at the three earliest SOAs (within 180 msec), whereby object processing is impaired and scene processing is improved.

The results of both Experiments 1 and 2 demonstrate support for separate parallel pathways carrying object and scene information. Behavioral studies have suggested that these parallel pathways interact such that scene context can facilitate the identification of objects (Boyce & Pollatsek, 1992; Rayner & Pollatsek, 1992) and contextual categorization of scenes can be impaired by the presence of a salient object in the scene, particularly when the object is incongruent with the scene context (Joubert, Rousselet, Fize, & Fabre-Thorpe, 2007). Although interactions between these parallel pathways are what likely gives rise to the perceived richness of scene identity, we have demonstrated that these interactions may be in the form of

inhibitory effects. The administration of rTMS to lLO likely unmasked these inhibitory interactions, thereby facilitating scene perception.

The competition between scene and object processing could conceivably be driven by SF differences between scene and object images. Previous research has suggested that the two hemispheres may be tuned to different SFs (Robertson & Ivry, 2000). Because scenes can be processed using low SF information alone (Schyns & Oliva, 1994), it is conceivable that rTMS to lLO disrupts high SF processing while leaving low SF processing intact. Although the SF content of our stimuli were highly variable, we observed a similar pattern of behavioral results with stimulation of the two hemispheres, although the effect was not significant in the right hemisphere for reasons previously discussed. The similar pattern of behavioral results observed between the left and right hemisphere stimulation may suggest that this effect is not dependent on SF and is driven by the competition between objects and scenes alone, with objects coded dominantly in the left hemisphere. However, the lack of significance observed during rTMS to rLO could imply that SF does indeed contribute to this effect and that stimulation to the right hemisphere does not interrupt high SFs to the same degree, resulting in less of a pronounced effect. It is feasible that both theories contribute to the current findings. Future research should attempt to systematically vary the SF content of object and scene stimuli during rTMS to lLO and rLO.

The present findings are consistent with imaging studies showing distinct processing areas for object (Grill-Spector et al., 2001) and scene stimuli (Epstein & Kanwisher, 1997), as well as patient data showing that scene categorization can operate independent of object processing (Steeves et al., 2004). Moreover, the present data extend our understanding of how objects and scenes are coded at a cortical level. Until recently, we have often relied on patient data to form causal links between brain areas and behavior; however, in many instances, the patient brain is an imperfect model, given that lesions are rarely restricted to only one cortical area. Our results using rTMS in the intact brain to transiently interrupt cortical processing provide causal support for the scene-centered approach to scene processing.

In summary, perceiving scenes in the real world likely involves the concurrent extraction, in parallel pathways, of the global image properties of a scene and the objects contained within the scene. Furthermore, interactions between these parallel pathways likely result in the perceived richness of scene identity and the speed at which we can process. We have demonstrated that these interactions can also be in the form of inhibitory effects of object processing on scene processing that are released using rTMS.

Acknowledgments

We thank Laurence Harris and Ruth Weiss for helpful comments on an earlier draft and Patrick Byrne for helpful statistical coun-

seling. This work was supported by grants from the Canada Foundation for Innovation, the Ontario Research Fund, and the Natural Sciences and Engineering Research Council of Canada.

Reprint requests should be sent to Jennifer K. E. Steeves, Centre for Vision Research, Department of Psychology, York University, 4700 Keele St., Toronto, Ontario, M3J 1P3, Canada, or via e-mail: steeves@yorku.ca.

Note

1. Two of the participants showed greater peak activation in the right hemisphere and were also tested using the right hemisphere LO as the rTMS target site. No significant differences were observed when comparing stimulation of the right hemisphere to that of the left hemisphere in these participants.

REFERENCES

- Biederman, I. (1981). On the semantics of a glance at a scene. In M. Kubovy & J. Pomerantz (Eds.), *Perceptual organization* (pp. 213–263). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Biederman, I. (1987). Recognition by components: A theory of human image understanding. *Psychological Review*, *94*, 115–147.
- Biederman, I. (1995). Visual object recognition. In S. F. Kosslyn & D. Osherson (Eds.), *An invitation to cognitive science. Vol. 2: Visual Cognition* (2nd ed., pp. 121–165). Cambridge, MA: MIT Press.
- Biederman, I., & Cooper, E. E. (1991). Object recognition and laterality: Null effects. *Neuropsychologia*, *29*, 685–694.
- Boyce, S. J., & Pollatsek, A. (1992). Identification of objects in scenes: The role of scene background in object naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 531–543.
- Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The Bank of Standardized Stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PLoS One*, *5*, e10773.
- Burton, G. J., & Moorhead, I. R. (1987). Color and spatial structure in natural scenes. *Applied Optics*, *26*, 157–170.
- Campana, G., Cowey, A., & Walsh, V. (2002). Priming of motion direction and area V5/MT: A test of perceptual memory. *Cerebral Cortex*, *12*, 663–669.
- Camprodon, J. A., Zohary, E., Brodbeck, V., & Pascual-Leone, A. (2010). Two phases of V1 activity for visual recognition of natural images. *Journal of Cognitive Neuroscience*, *22*, 1262–1269.
- Chen, R., Classen, J., Gerloff, C., Celnik, P., Wassermann, E. M., Hallett, M., et al. (1997). Depression of motor cortex excitability by low-frequency transcranial magnetic stimulation. *Neurology*, *48*, 1398–1403.
- Cohen Kadosh, K., Walsh, V., & Cohen Kadosh, R. (2010). Investigating face property specific processing in the right OFA. *Social Cognitive and Affective Neuroscience*, *6*, 58–65.
- Corthout, E., Hallett, M., & Cowey, A. (2002). Early visual cortical processing suggested by transcranial magnetic stimulation. *NeuroReport*, *13*, 1163–1166.
- Corthout, E., Uttl, B., Walsh, V., Hallett, M., & Cowey, A. (1999). Timing of activity in early visual cortex as revealed by transcranial magnetic stimulation. *NeuroReport*, *10*, 2631–2634.
- De Graef, P., Christaens, D., & d'Ydewalle, G. (1990). Perceptual effects of scene context on object identification. *Psychological Research*, *52*, 317–329.

- Epstein, R., & Kanwisher, N. (1997). A cortical representation of the local environment. *Nature*, *392*, 598–601.
- Fize, D., Fabre-Thorpe, M., Richard, G., Doyon, B., & Thorpe, S. J. (2005). Rapid categorization of foveal and extrafoveal natural images: Associated ERPs and effects of lateralization. *Brain and Cognition*, *59*, 145–158.
- Freidman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology*, *108*, 316–355.
- Greene, M. R., & Oliva, A. (2006). Natural scene categorization from the conjunction of ecological global properties. *Proceedings of the Twenty-Eighth Annual Conference of the Cognitive Science Society*, *1*, 291–296.
- Greene, M. R., & Oliva, A. (2009). The briefest of glances: The time course of natural scene understanding. *Psychological Science*, *20*, 464–472.
- Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, *41*, 1409–1422.
- Halgren, E., Raij, T., Marinkovic, K., Jousmäki, V., & Hari, R. (2000). Cognitive response profile of the human fusiform face area as determined by MEG. *Cerebral Cortex*, *10*, 69–81.
- Henderson, J., & Hollingsworth, A. (1999). High-level scene perception. *Annual Review of Psychology*, *50*, 243–271.
- Joubert, O. R., Rousselet, G. A., Fize, D., & Fabre-Thorpe, M. (2007). Processing scene context: Fast categorization and object interference. *Vision Research*, *47*, 3286–3297.
- Laeng, B., Shah, J., & Kosslyn, S. (1999). Identifying objects in conventional and contorted poses: Contributions of hemisphere-specific mechanisms. *Cognition*, *70*, 53–85.
- Marsolek, C. J. (1995). Abstract visual-form representations in the left cerebral hemisphere. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 375–386.
- Michel, C. M., Seeck, M., & Murray, M. M. (2004). The speed of visual cognition. *Supplements to Clinical Neurophysiology*, *57*, 617–627.
- Murray, M. M., Wylie, G. R., Higgins, B. A., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). The spatiotemporal dynamics of illusory contour processing: Combined high-density electrical mapping, source analysis, and functional magnetic resonance imaging. *Journal of Neuroscience*, *22*, 5055–5073.
- Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: As holistic representation of the spatial envelope. *International Journal of Computer Vision*, *42*, 145–175.
- Oliva, A., & Torralba, A. (2006). Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research: Visual Perception*, *155*, 23–36.
- Pascual-Leone, A., Tormos, J. M., Keenan, J., Tarazona, F., Cañete, C., & Catalá, M. D. (1998). Study and modulation of human cortical excitability with transcranial magnetic stimulation. *Journal of Clinical Neurophysiology*, *15*, 333–343.
- Peelen, M. V., Fei-Fei, L., & Kastner, S. (2009). Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature*, *460*, 94–97.
- Pitcher, D., Charles, L., Devlin, J. T., Walsh, V., & Duchaine, B. (2009). Triple dissociation of faces, bodies and objects in the extrastriate cortex. *Current Biology*, *19*, 319–324.
- Pitcher, D., Walsh, V., Yovel, G., & Duchaine, B. (2007). TMS evidence for the involvement of the right occipital face area in early face processing. *Current Biology*, *17*, 1568–1573.
- Rayner, K., & Pollatsek, A. (1992). Eye movements and scene perception. *Canadian Journal of Psychology*, *46*, 342–376.
- Robertson, L. C., & Ivry, R. (2000). Hemispheric asymmetries: Attention to visual and auditory primitives. *Current Directions in Psychological Science*, *9*, 59–63.
- Rossi, S., Hallett, M., Rossini, P. M., & Pascual-Leone, A. (2009). Safety of TMS Consensus Group. Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clinical Neurophysiology*, *120*, 2008–2039.
- Sanocki, T., & Epstein, W. (1997). Priming spatial layout of scenes. *Psychological Science*, *8*, 374–378.
- Schendan, H. E., Ganis, G., & Kutas, M. (1998). Neurophysiological evidence for visual perceptual categorization of words and faces within 150 ms. *Psychophysiology*, *35*, 240–251.
- Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science*, *5*, 195–200.
- Silvanto, J., Lavie, N., & Walsh, V. (2005). Double dissociation of V1 and V5/MT activity in visual awareness. *Cerebral Cortex*, *15*, 1736–1741.
- Steeves, J. K. E., Humphrey, G. K., Culham, J. C., Menon, R. S., Milner, A. D., & Goodale, M. A. (2004). Behavioral and neuroimaging evidence for a contribution of color and texture information to scene classification in a patient with visual form agnosia. *Journal of Cognitive Neuroscience*, *16*, 955–965.
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*, 520–522.
- Torralba, A., & Oliva, A. (2003). Statistics of natural images categories. *Networks: Computation in Neural Systems*, *14*, 391–412.
- Walker Renninger, L., & Malik, J. (2004). When is scene identification just texture recognition? *Vision Research*, *44*, 2301–2311.
- Walsh, V., Ellison, A., Battelli, L., & Cowey, A. (1998). Task-specific impairments and enhancements induced by magnetic stimulation of human visual area V5. *Proceedings of the Royal Society of London, Series B, Biological Sciences*, *265*, 537–543.
- Wassermann, E. M. (1998). Risk and safety of repetitive transcranial magnetic stimulation: Report and suggested guidelines from the International Workshop on the Safety of Repetitive Transcranial Magnetic Stimulation, June 5–7, 1996. *Electroencephalography and Clinical Neurophysiology*, *108*, 1–16.
- Xiao, J., Hays, J., Ehinger, K., Oliva, A., & Torralba, A. (2010). *SUN database: Large-scale scene recognition from abbey to zoo*. Paper presented at the IEEE Conference on Computer Vision and Pattern Recognition (pp. 3485–3492). San Francisco, CA.
- Zwaan, R. A., & Yaxley, R. H. (2004). Lateralization of object-shape information in semantic processing. *Cognition*, *94*, B35–B43.