

available at www.sciencedirect.comwww.elsevier.com/locate/brainres

**BRAIN
RESEARCH**

Research Report
The role of apparent size in building- and object-specific regions of ventral visual cortex
Anthony D. Cate^{a,*}, Melvyn A. Goodale^b, Stefan Köhler^{b,}**
^aVeterans Affairs Northern California Health Care System, CA, USA

^bUniversity of Western Ontario, Ontario, Canada

ARTICLE INFO
Article history:

Accepted 9 February 2011

Available online 15 February 2011

Keywords:

Perception

Vision

fMRI

Building

Object

Scene

ABSTRACT

Images of buildings and manipulable objects have been found to activate distinct regions in the ventral visual pathway. Yet, many non-categorical properties distinguish buildings from common everyday objects, and perhaps the most salient of these is size. In this fMRI study, we investigated whether or not changes in perceived scale can account for some of the differences in category-specific responses, independent of the influence of semantic or retinotopic image properties. We used independent scans to localize object-specific ROIs in lateral occipital cortex (LO) and scene-specific ROIs in the parahippocampal place area (PPA) and posterior collateral sulcus. We then contrasted the effects of stimulus category and perceived size/distance in these regions in a factorial design. Participants performed an oddball detection task while viewing images of objects, buildings, and planar rectangles both with and without a background that indicated stimulus size/distance via simple pictorial cues. The analyses of fMRI responses showed effects of perceived size/distance in addition to effects of category in LO and the PPA. Interestingly, when simple rectangles were presented in a control condition against the background that indicated size/distance, LO in the right hemisphere responded significantly more to the small/close rectangles than to the large/far ones, in spite of the fact that the rectangles themselves were identical. These findings suggest that ventral stream regions that show category specificity are modulated by the perceived size and distance of visual stimuli.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Buildings and small everyday objects are two categories of images that have been shown to activate distinct regions of the brain in human neuroimaging studies. Among these are the parahippocampal place area (PPA; Epstein and Kanwisher,

1998), which has been shown to respond preferentially to scenes and buildings, and the lateral occipital area (LO; Malach et al., 1995), which has been shown to respond preferentially to everyday objects. It is often assumed that this specificity is based on the unique shape cues that define a particular category.

* Correspondence to: A.D. Cate, Human Cognitive Neurophysiology Laboratory, Veterans Affairs, Northern California Health Care System, 150 Muir Rd., 151-I, P.O. Box 2339 Martinez, CA 94553, USA. Fax: +1 925 372 2561.

** Correspondence to: S. Köhler, Department of Psychology, University of Western Ontario Social Science Centre, Room 7418, London, Ontario, Canada N6A 5C2.

E-mail addresses: acate@ebire.org (A.D. Cate), stefank@uwo.ca (S. Köhler).

Abbreviations: CS, collateral sulcus; LG, lingual gyrus; LO, lateral occipital; PHG, parahippocampal gyrus; PPA, parahippocampal place area; ROI, region of interest

The PPA not only typically responds preferentially to images of many varieties of scenes (Epstein and Kanwisher, 1998; Epstein et al., 1999; Köhler et al., 2002; Goh et al., 2004; Steeves et al., 2004; Epstein et al., 2007) and also responds more to cropped images of single buildings (Maguire et al., 2001) than it does to images of household objects and faces (Epstein and Kanwisher, 1998; Epstein et al., 1999), even though many buildings and common everyday objects share a very similar global rectilinear form. A similar specificity for buildings has been observed in other ventral-stream regions including the right lingual gyrus (LG) (Gorno-Tempini and Price, 2001; Maguire et al., 2001; Rosenbaum et al., 2004) and the posterior collateral sulcus (CS) (Haxby et al., 1999; Ishai et al., 1999; Levy et al., 2001; Hasson et al., 2002; Malach et al., 2002; Avidan et al., 2003; Grill-Spector, 2003; Lerner et al., 2003; Levy et al., 2004). A related review of ventral stream stimulus specificity (Haxby et al., 2000) revealed that a region of building-specific activation was posterior to and distinct from the more scene-general PPA, even though the two may overlap in some studies.

LO is broadly tuned with respect to shape, with its preferred category consisting of any well-formed image of a concrete object (Grill-Spector et al., 1998a,b; Grill-Spector et al., 1999; Kourtzi and Kanwisher, 2000; Lerner et al., 2001; Ferber et al., 2005). Thus, LO will respond to some degree to images of buildings, although not as much as to objects (Grill-Spector, 2003).

The reviewed neuroimaging evidence is consistent with neuropsychological studies showing dissociations between visual object agnosia and topographical or landmark agnosia. Lesions of the anterior LG have been implicated in landmark agnosia, a rare disorder that impairs the recognition of large built edifices but not of small objects. In the relatively few documented cases of this disorder, the patients were able to recognize small objects and shapes but not buildings as a whole (Whiteley and Warrington, 1978); Takahashi and Kawamura, 2002. The effects of LG lesions appear to be distinct from disorders caused by rare focal lesions to the parahippocampal gyrus (PHG). The latter lesions impair the ability to learn the navigation of new routes (Aguirre and d'Esposito, 1999; Barrash et al., 2000), although there is some evidence that PHG lesions also impair the ability to recognize scenes (Lee et al., 2005a,b, and landmarks Takahashi and Kawamura, 2002). In contrast, patient DF who developed visual form agnosia following damage to the lateral posterior occipitotemporal cortex largely centered on area LO (Goodale et al., 1991; Milner et al., 1991; James et al., 2003) shows extremely poor perception of visual objects on the basis of their shape. However, DF does retain some ability to discriminate among scenes, particularly if those scenes have diagnostic color and texture cues (Steeves et al., 2004).

Remarkably, although the role of shape cues in determining category-specific responses in the ventral stream has been well studied, little is known about the role of size, perhaps one of the most salient differences between buildings and common objects. In fact, almost any building will be significantly larger than almost any object. Because they are larger, buildings are more likely to project a larger image on the retina than objects, which means that they occupy larger visual field eccentricities (Levy et al., 2001; Hasson et al., 2002; Levy et al., 2004) and observers may perceive them using lower spatial frequencies (Peyrin et al., 2003; Peyrin et al., 2004). But of

course, in most fMRI studies, these salient differences are not evident because participants are almost always shown pictures of buildings and objects where differences in retinotopic size and spatial frequencies are minimized.

In fMRI research, the image sets used to determine stimulus-specific regions-of-interest (ROIs) in the ventral visual pathway are chosen overwhelmingly on the basis of their membership in a semantic category rather than as groups according to their values on a continuous dimension like size or shape (c.f. Tootell et al., 2008). There are practical reasons for avoiding image size manipulations in neuroimaging experiments, namely that the small bore of an fMRI scanner typically limits the visual angle of stimuli. Also, higher-level ventral stream regions show significantly greater (but not complete, e.g., Op de Beeck and Vogels, 2000) invariance across retinotopic transformations than do early visual areas and so there may have been an expectation that image size would have little effect on their activation. Early studies of region LO reported translation-, scale-, and contrast-invariant BOLD responses, and similar findings have been discussed for the PPA (MacEvoy and Epstein, 2007) and FFA (Kanwisher et al., 1999; Ewbank et al., 2005).

More recently, however, both LO and the PPA have been found to show some sensitivity to retinotopic size and translation, although on a much lower scale than is typically observed in classical occipital retinotopic fields (Grill-Spector and Malach, 2001). In fact, LO appears to respond more strongly to contralateral stimuli (Niemeier et al., 2005; Hemond et al., 2007), and two LO subregions have been shown to be retinotopically mapped with respect to both polar angle and eccentricity (Wandell et al., 2005; Larsson and Heeger, 2006; Sayres and Grill-Spector, 2008). The PPA has been found to be sensitive to the hemifield in which scene images were presented in the fMRI scanner (Schwarzlose et al., 2008) and to encode retinotopic location information more generally (Carlson et al., 2003). Two scene-selective regions of the parahippocampal gyrus have also been found to show a degree of retinotopic organization (Arcaro et al., 2009).

Sensitivity to retinotopic eccentricity has been proposed as a way to account for the apparent category specificity of higher-level visual regions (Levy et al., 2001; Hasson et al., 2002; Levy et al., 2004). Even within the relatively small range of eccentricities feasible in a typical fMRI setup (e.g., within 20°; Levy et al., 2004), it has been found that the activation of building-selective regions in the CS was modulated largely by the eccentricity of visual stimuli, with little regard to what they depicted (Levy et al., 2001; Levy et al., 2004). Specifically, larger eccentricities produced greater activation in the building-selective regions, which the authors attributed to the fact that buildings viewed in the real world extend to larger retinotopic eccentricities than smaller objects. These authors suggested that the CS building-selective region integrates information across broad regions of the visual field, requiring only low acuity and giving rise to the peripheral preference, an interpretation that was bolstered by similar experiments with amblyopic patients (Lerner et al., 2003).

Complementary to findings suggesting that higher-level areas can be sensitive to retinotopic image properties, it has been found that purely perceptual changes in size modulate fMRI activation in retinotopic primary visual cortex. Specifically, simple stimuli positioned in different parts of a pictorial

corridor illusion evoked V1 activation that corresponded to appropriately larger and smaller stimulus visual angles (Murray et al., 2006). Importantly, these effects were found using 2D stimuli both with and without rich cues to pictorial depth (e.g., texture gradients); a simple background consisting of two parallel lines receding to a vanishing point yielded significant behavioral and fMRI results. An important remaining question is whether or not manipulations of perceived size would have an effect on the activity of higher-level visual areas as well.

It was this issue that we addressed in the current study; in other words, we asked whether or not the apparent category-specific activity in areas LO, PPA, and posterior collateral sulcus would be influenced by purely perceived changes in size. Specifically, we hypothesized that category-selective ROIs would respond more strongly to an item that appeared to have the size appropriate for the category in question. We predicted that object-selective LO would respond more when subjects viewed items that appeared to be the size of a small object and that building-selective PPA and posterior CS would respond more strongly when subjects viewed items that appeared to be building-sized. In this sense, our reasoning is analogous to that of previous studies on retinotopic eccentricity bias (Levy et al., 2001; Hasson et al., 2002; Levy et al., 2004). We hypothesized that the introduction of contextual cues to bias interpretation of apparent size of the target items would modulate activation in category-selective ROIs (Cate et al., 2006; Amit et al., 2008) and that such a top-down cognitive mechanism would not rely on truly illusory impressions of size.

In our study, participants viewed images of buildings and objects in sparse pictorial contexts that indicated different real-world sizes. We used 2D displays in which an item's perceived size was altered by varying its apparent relation with other items in the picture, as well as by a very simple perspective cue (two converging floor lines). These size cues were introduced to emphasize conceptual rather than low-level perceptual information about item scale. For similar reasons, and because the activation of regions with scene-specific responses is known to be influenced by the perception of material surface properties (Cant and Goodale, 2007), we avoided rendering the background with any surfaces or textures. Perceived size in the resulting displays was always linked with perceived distance; after all, the apparent size of a stimulus is inversely related to its apparent distance from the observer.

In real-world buildings and objects, category membership and size are inextricably entwined. The current study employed a 2×2 experimental design which included category

and perceived size as crossed factors. We selected experimental stimuli from two subcategories of buildings and objects, namely garages and cameras; these stimulus classes were chosen in order to minimize differences in visual form across categories and again to emphasize conceptual over perceptual category differences. The resulting design allowed us to compare the size of the two main effects directly and to assess whether or not the effects of category and size are independent. We also included additional control conditions to see if the effects of size and category could be established in isolation and to assess whether category cues have different effects when items are presented with or without scene contexts. In all analyses, we specifically focused on functionally defined ROIs that have been linked to category-specific stimuli in past research. Thus, we used localizers to identify areas LO, PPA, and the posterior collateral sulcus.

2. Results

2.1. Behavioral results

Each subject's mean hit rates for oddball detection in the four conditions of the 2×2 factorial design were calculated and submitted to a repeated-measures ANOVA with image category and apparent size/distance as independent factors. The data for the two pairs of control conditions (blank background and rectangle) were compared using paired-sample t-tests.

The ANOVA of RTs showed no significant main effect of image category ($F_{(1,8)}=1.69$, $p=0.23$) nor of apparent size ($F_{(1,8)}=2.12$, $p=0.18$) nor an interaction ($F_{(1,8)}=0.15$, $p=0.71$). The ANOVA of oddball hit rates showed no significant main effect of image category ($F_{(1,8)}=0.24$, $p=0.63$) nor of apparent size ($F_{(1,8)}=0.23$, $p=0.65$), but there was a trend towards an interaction ($F_{(1,8)}=4.41$, $p=0.07$). Mean hit rates for the four conditions were small objects, 88.9% (6.9%); large objects, 84.3% (9.5%); small buildings, 85.2% (6.9%); and large buildings 87.0%, (5.3%). There were no significant differences between the hit rates for object and building stimulus blocks in either the blank background ($t_{(8)}=-0.89$, $p=0.40$) or the rectangle ($t_{(8)}=0.00$, $p=1.00$) control condition.

The non-significant trend towards an interaction in the hit rate results was in a direction suggesting that subjects may have performed slightly better in the conditions representing category/apparent size pairings congruent with typical real-world experiences (e.g., small objects).

Table 1 – Object and scene ROI coordinates and frequencies.

	Talairach coordinates in mm (SD)			N	Subject number								
	x	y	z		1	2	3	4	5	6	7	8	9
Object ROIs													
LO left	-48.2 (1.8)	-63.5 (2.0)	4.1 (2.6)	8	x		x	x	x	x	x	x	x
LO right	40.6 (2.5)	-68.7 (2.5)	2.0 (1.5)	9	x	x	x	x	x	x	x	x	x
Scene ROIs													
PPA left	-24.6 (1.7)	-38.4 (3.2)	-8.4 (1.7)	7		x	x	x	x	x		x	x
PPA right	17.3 (2.4)	-40.8 (2.8)	-9.6 (1.3)	8		x	x	x	x	x	x	x	x
Posterior CS left	-27.0 (1.0)	-68.8 (3.3)	-8.9 (1.8)	9	x	x	x	x	x	x	x	x	x
Posterior CS right	18.6 (2.4)	-63.7 (4.0)	-8.9 (2.4)	7	x			x	x	x	x	x	x

2.2. FSO localizer

Scene-specific ROIs were identified at two sites: in the parahippocampal place area (PPA) and in the posterior collateral sulcus (CS). PPA ROIs that met statistical criteria of significance were identified in the left hemisphere of 7 out of 9 participants and in the right hemisphere of 8 participants (see Table 1). CS ROIs were identified in the left hemisphere of all 9 participants and in the right hemisphere of 7 participants. An object-selective ROI was identified in the lateral occipital cortex (LO) in the left hemisphere of 8 participants and in the right hemisphere of all 9 participants. For illustration, ROIs from one representative subject are shown on an inflated representation of the subject's right cerebral hemisphere. This representation was obtained by segmenting the gray-white matter boundary in the right hemisphere using BrainVoyager QX to create an inflated 3D rendering.

2.3. Experimental scans

Table 2 summarizes the results of the 2×2 ANOVA performed on mean percent signal change averaged across voxels in ROIs from both hemispheres. In addition to *F* and *p* values, the table lists omega-squared (ω^2) values for each effect. Omega-squared is an effect-size measure based on a population-level estimate of the variance accounted for by the independent factor.

LO was characterized by consistently high responses to objects, while building responses depended strongly on small perceived size. Main effects of image category and perceived size were significant in LO (see Table 2), with object images and small perceived sizes receiving higher responses than their corresponding counterparts. There was also a significant two-way interaction, such that perceived size had a significant effect only within the category of building stimuli. The Tukey's HSD test for this ANOVA gave a Tukey's critical *T* value of 0.41, which was the required magnitude for the difference between two conditions to be significant ($\alpha=0.05$). This indicated that the "small building" condition had a significantly larger signal change value than the "large building" condition, but the slight difference between the small and large camera conditions was not significant.

In the PPA, image category and perceived size had independent effects. The PPA showed both main effects without an interaction, with higher responses to building images and large perceived sizes as compared to the other corresponding

conditions; Tukey's HSD (critical *T*=0.12) showed that the "small objects" condition had a significantly smaller response than both the "small building" and "large building" conditions. CS activation was only consistently enhanced by the presence of building stimuli. CS showed only the main effect of image category (building images higher than objects) with a trend towards an interaction; there were no significant pairwise differences (Tukey's critical *T*=0.27). The summarized combined-hemispheres data are plotted for each ROI in Fig. 2.

Table 3 summarizes the results of the 2×2 ANOVA performed using the ROI signal change data separated by hemisphere. The directions of the significant effects were identical to those shown for the data pooled across hemispheres. Significant main effects of image category were found only in left hemisphere LO. Main effects of perceived size were found in both left and right hemisphere LO. As before, there were significant two-way interactions in both hemispheres such that perceived size only modulated responses to the building stimuli. The "small building" condition had a significantly higher response than the "large building" condition, but there were no significant pairwise difference between the small and large object conditions.

Only the left hemisphere PPA showed a significant main effect of image category. Tukey's HSD comparisons revealed that the "large building" condition had a significantly higher response than both the "small object" and "large object" conditions. In contrast, main effects of perceived size were found in the PPA in both hemispheres; the "large building" condition was significantly higher than "small object." Although the pattern of differences in activation in the CS of each hemisphere mirrored the pattern seen when both were combined, there were only trends towards significance for the main effect of image category and for the two-way interaction between image category and perceived size. That being said, *F*-values were somewhat higher in the right hemisphere.

The results of the *t*-tests for the control condition are summarized in Tables 4 and 5. The combined-hemispheres data are shown in Fig. 3. In the corresponding statistical analyses, there were no significant effects in LO for either size alone or category alone, with only a slight trend towards greater responses to the apparently small rectangle. When the data were analyzed separately by hemisphere, LO did show a significantly higher response to apparently small rectangles in the left hemisphere. In contrast, both the PPA and the CS showed no size effect but did show significantly higher

Table 2 – Summary of the experimental conditions ROI analyses, data combined across hemispheres.

	2×2 ANOVA								
	Type			Size			Interaction		
	<i>F</i> (<i>df</i>)	<i>p</i>	ω^2	<i>F</i> (<i>df</i>)	<i>p</i>	ω^2	<i>F</i> (<i>df</i>)	<i>p</i>	ω^2
Object ROI									
LO	8.94 (1,8)	.017*	.50	14.92 (1,8)	.0048**	.64	12.35 (1,8)	.0079**	.59
Scene ROIs									
PPA	9.54 (1,7)	.018*	.55	16.27 (1,7)	.0050**	.68	0.13 (1,7)	.74	<0
Post CS	13.40 (1,8)	.0064**	.61	0.014 (1,8)	.91	<0	4.73 (1,8)	.062	.32

**p* < 0.05.

***p* < 0.01.

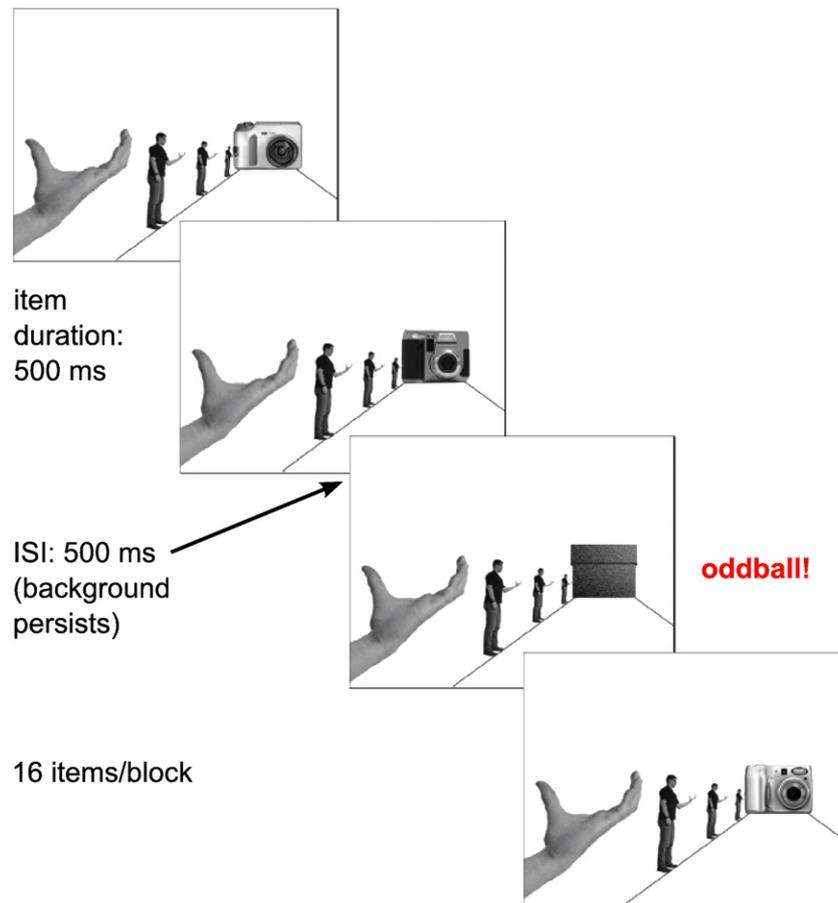


Fig. 1 – Experimental stimuli and task. Subjects fixated exemplars of either objects (cameras) or buildings (garages) that changed every second in a blocked design. In some conditions, stimuli were set against a minimal scene background that suggested different item sizes and distances according to simple pictorial cues. To equalize attention across conditions, the subjects' task was to detect an oddball stimulus that appeared unpredictably either once or twice per block.

responses to building images presented on blank backgrounds than to objects presented on blank backgrounds (see Fig. 3). When examined separately by hemisphere, this pattern was found to be most pronounced in the left PPA.

3. Discussion

This purpose of the present study was to investigate whether perceived size differences modulate category-specific responses in the ventral visual pathway. Towards this end, we manipulated the perceived size of two subcategories of objects and buildings, namely cameras and garages. We found that the BOLD signal in ventral visual pathway regions with typical category-specific responses—as revealed with our functional localizers—is indeed modulated by changes in the apparent size of retinotopically identical images. Although the exact nature of these effects differed across regions, the overall pattern suggests that selectivity for image category may at least in part be explained in terms of real-world (vs. retinotopic) size selectivity and that ventral visual regions are attuned to co-variations between the size and categorical identity of items in the visual environment.

3.1. Area LO

LO was strongly modulated by the apparent size of buildings and abstract shapes but not by the apparent size of images of small objects, i.e., cameras. Specifically, the category sensitivity of LO was observed in the presence of an interacting effect of perceived size. Critically, responses to exemplars of the presumed non-preferred stimulus class of LO (i.e., garages) were significantly increased by the impression of small size. To our knowledge, such an effect has not been reported previously, and it reflects, perhaps, the most important aspect of the current results. A pure effect of perceived size was also revealed in the control conditions, in which no cues to object identity were present. In contrast, LO responses to cameras were consistently high regardless of the induced differences in perceived size, in accord with earlier reports of retinal size invariance in LO (e.g., Malach et al., 1995, Grill-Spector et al., 1998a,b). It may be that, once observers recognized that an item belonged to a category of small things (cameras) by virtue of size-invariant shape features, top-down signals prevented conflicting size/distance cues in the display from modulating LO activation. Our displays' pictorial cues were designed to be high-level cognitive indicators of size/distance and not truly

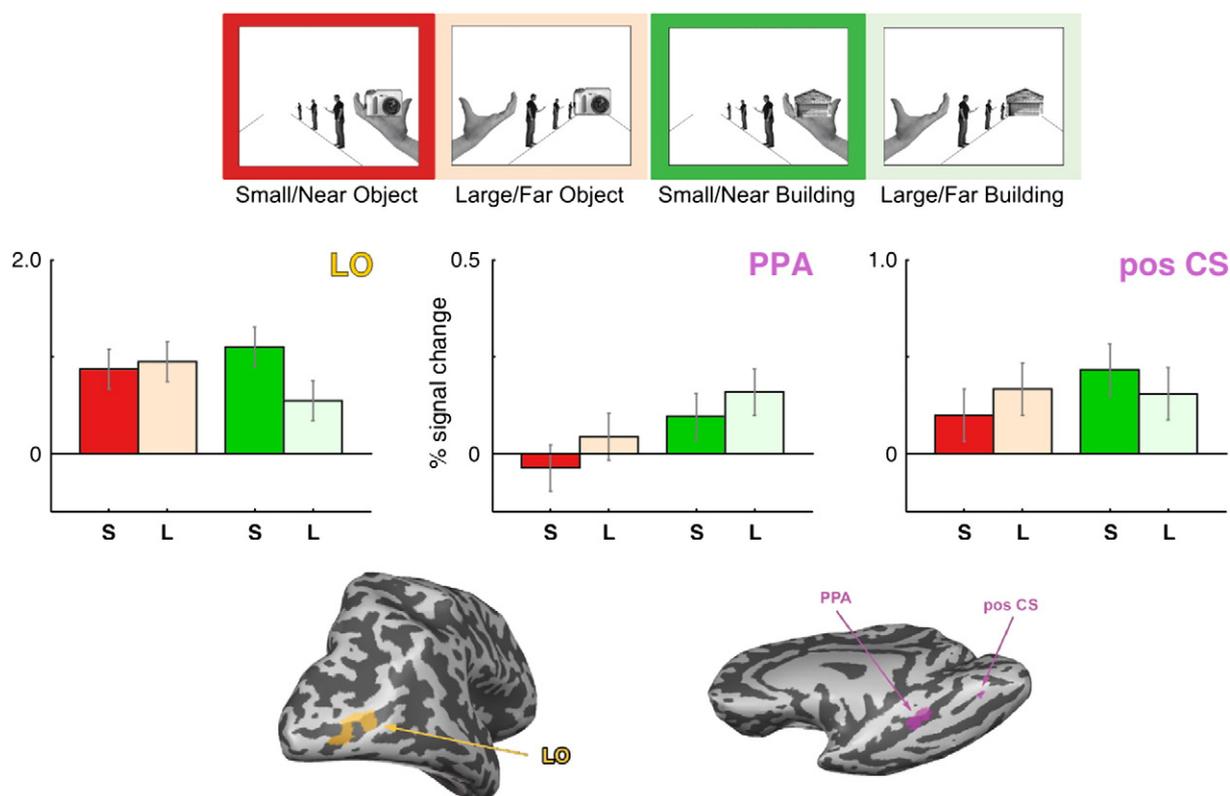


Fig. 2 – Factorial analysis of object and scene ROI activation. Top row: the four conditions used with the 2 × 2 factorial analysis. The constant screen size (~6°) of the focal stimuli corresponded to appropriate sizes for cameras and garages when compared to the familiar size of an individual depicted in the scene. The screen position of the fixated items (left/right) was counterbalanced across blocks. Middle row: bar graphs show ROI mean percent BOLD signal change relative to a fixation baseline for the four conditions at top (data from both hemispheres combined). S indicates focal stimulus in apparently small/far scene context; L, apparently large/near scene context. Error bars represent the critical T value from Tukey's HSD post-hoc comparison of the four conditions; conditions whose error bars do not overlap are significantly different. Note the different scales of the bar graph ordinates. Bottom row: posterolateral and ventromedial views of one subjects' inflated right cerebral hemisphere, with the locations of the ROIs identified in that subject.

compelling illusions. It seems likely that these pictorial size/distance cues would be less salient than the sense of small size associated with cameras as a category. Reasoning along these lines, it would appear that LO activity is not sensitive to the entire range of perceived sizes but rather that LO responds preferentially once small size in particular has been indicated by either shaped-based or other cues. Otherwise, one would expect to see consistently low LO responses to garage images due to their familiar large size. Thus, size-invariant shape cues that indicate a category associated with small size may be one of the primary properties that determines LO's typical response preference for common objects.

Although the LO ROI was selected based on a category-specific preference for common objects in our functional localizer, LO responses to our experimental stimuli showed relatively weak category specificity. When items were embedded in the scene context that offered cues towards real-world size, only the left hemisphere LO ROI showed a significant category effect. When the two types of stimuli were presented in the blank-background control condition, we found no effect

of category at all. This pattern of results may at first seem puzzling given that LO was selected with our functional localizer based on a response preference for common objects. It should be noted, however, that the stimuli used in our experimental setup differed in meaningful and intended ways from those in the localizer scans and in previous research on category specificity. Critically, experimental stimuli were restricted to images of garages and cameras rather than common objects and buildings more broadly, so as to minimize differences in physical structure independent of size. As it is well established that LO plays a critical role in visual shape analysis (e.g., Kourtzi and Kanwisher, 2000; Sayres and Grill-Spector, 2008), it is not so surprising that the typical category effect is reduced or even eliminated once variations in form are strongly constrained within as well as across categories. The reduction in differential responding may also be linked to an adaptation of the fMRI response across the sequence of highly similar items within each block, a response property that is well documented for LO (e.g., Grill-Spector et al., 1999; Cant and Goddard, 2009). Given the limited

Table 3 – Summary of the experimental conditions ROI analyses, data separated by hemisphere.

	2×2 ANOVA								
	Type			Size			Interaction		
	F (df)	p	ω ²	F (df)	p	ω ²	F (df)	p	ω ²
Object ROIs									
LO left	6.059 (1,7)	.043*	.42	16.19 (1,7)	.0050**	.68	9.68 (1,7)	.017*	.55
LO right	2.75 (1,8)	.14	.18	5.74 (1,8)	.043*	.37	6.01 (1,8)	.040*	.39
Scene ROIs									
PPA left	44.35 (1,6)	.00055**	.88	8.12 (1,6)	.029*	.54	.10 (1,6)	.77	<0
PPA right	1.98 (1,7)	.20	.12	23.79 (1,7)	.0016**	.77	.058 (1,7)	.82	<0
Post CS left	4.11 (1,8)	.077	.28	.00023 (1,8)	.99	<0	2.17 (1,8)	.18	.13
Post CS right	5.44 (1,6)	.058	.43	0.52 (1,6)	.50	<0	5.44 (1,6)	.056	.43

presence of category effects in the present study, one might argue that the significant effect of perceived size on LO responses is even more remarkable.

The observed response to perceived size is in accord with findings of other studies showing that LO is sensitive to specific changes in the visual form of an object (Gerlach et al., 1999; Grill-Spector et al., 2001) but is much less sensitive to changes across exemplars of the same semantic category when those changes are carefully controlled for in terms of changes in visual appearance (Chouinard et al., 2008). In addition, meta-analyses have revealed that activity in this region is more frequently associated with perceptual matching than with naming (Joseph, 2001), a pattern that stands in contrast to the activity observed in more anterior fusiform gyrus regions within the ventral stream (Bar et al., 2001; Gerlach et al., 2002).

LO responded preferentially to items whose size-invariant shape features indicated a small object (a camera), and for other kinds of items (buildings and rectangles), it responded more when stimuli were perceived to be small. This is in accord with studies of extrastriate lesions in humans. Hemimicropsia is a rare disorder resulting from lateral occipital lesions, typically BA 18/19, perhaps overlapping with the fMRI-defined LO (Frassinetti et al., 1999). The impairment is consistently retinotopic: contralateral stimuli are perceived to be smaller than ipsilateral ones (Cohen et al., 1994; Frassinetti et al., 1999; Kassubek et al., 1999), even when the stimuli are parts of the same object (Cohen et al., 1994). The

impairment also appears to be limited to objects in that patients do not report distorted perception of inter-object distance. Although there is at least one case of optic ataxia accompanying micropsia in the affected hemifield (Cohen et al., 1994), the perceptual impairments of hemimicropsia appear to spare the ability to guide actions appropriate for an object's size (Cohen et al., 1994); the lesions typically only affect perception.

Macaque lesion studies have also found size-related perceptual impairments linked to inferior occipitotemporal cortex. Macaques with specific V4 lesions exhibit an asymmetric difficulty in detecting small oddball targets among identically-shaped large standards (Schiller and Lee, 1991; Schiller, 1995) as part of a set of object perception impairments. In addition, it has been reported that prestriate and inferotemporal lesions impair detection of targets that are smaller than standards (Yaginuma et al., 1982). More generally, a previous conclusion that prestriate lesions affect size judgments by impairing the ability to integrate distance information with an object's retinal size is of particular relevance for the findings we report for LO; it also emphasizes a role for perceived (rather than retinal) size and points to more specific mechanisms potentially involved (Ungerleider et al., 1977).

The preference for small item size in LO that appears to accompany selectivity for real-world objects (in the sense of small, discrete items as a category) is also relevant in the context of other fMRI findings on response preferences of this ventral stream region. LO has been found to show a foveal position response bias, in contrast to the peripheral bias for houses (e.g., Levy et al., 2001), but only when the stimuli are images of real objects and not abstract pattern stimuli used for retinotopic mapping (Sayres and Grill-Spector, 2008). LO is also biased to respond to stimuli in the lower visual field (Niemeier et al., 2008; Schwarzlose et al., 2008), which could reflect a specific role in perceiving graspable objects in peripersonal space. Indeed, a recent fMRI study found preferential LO activation for near vs. far objects but only when near objects occupied the visual space surrounding a hand (Makin et al., 2007). Notably, the current study found a strong bias for small/near objects and a category effect that disappeared when the camera and the garage were presented without size/distance cues. It may be that the strength of the perceived size/distance effect was related to the fact that the small/near condition was defined by placing a stimulus on an image of an outstretched palm.

Table 4 – Summary of the control conditions ROI analyses, data combined across hemispheres.

	Paired-samples t-tests of control conditions			
	Type (blank background stimuli)		Size (rectangle stimuli)	
	t (df)	p	t (df)	p
Object ROI				
LO	-1.43 (8)	.19	1.98 (8)	.083
Scene ROIs				
PPA	-2.93 (7)	.022*	0.20 (7)	.85
Post CS	-2.78 (8)	.024*	0.91 (8)	.39

Table 5 – Summary of the control conditions ROI analyses, data separated by hemisphere.

	Paired-samples t-tests of control conditions			
	Type (blank background stimuli)		Size (rectangle stimuli)	
	t (df)	p	t (df)	p
Object ROIs				
LO left	–1.33 (7)	.27	2.80 (7)	.026*
LO right	–1.22 (8)	.26	1.88 (8)	.097
Scene ROIs				
PPA left	–4.62 (6)	.0036**	0.081 (6)	.94
PPA right	–0.89 (7)	.41	0.17 (7)	.87
Post CS left	–1.60 (8)	.15	0.21 (8)	.84
Post CS right	–1.68 (6)	.14	0.19 (6)	.85

3.2. PPA and CS

The PPA was modulated by both category and size and these effects were found to be independent. The category effect was limited to the left hemisphere, suggesting that the PPA may support a more abstract semantic representation of the visual stimulus classes that needed to be differentiated in the current study. Such an interpretation is in line with the fact that many of the typical perceptual differences that exist between common objects and buildings, in particular those pertaining to form, were minimized for the present purpose. A noticeable body of research in visual object priming (Marsolek et al., 1992; Marsolek, 1995) supports the idea that the left hemisphere supports representations of objects that generalize across form and perceptual detail that is irrelevant for object identification, whereas the right hemisphere supports representations that are form-specific. In particular, regions in the left occipitotemporal cortex have been reported to show priming (i.e., reductions) in the BOLD signal across different exemplars of objects. By contrast, in corresponding right-hemisphere regions, priming appears to be limited to repetitions of identical images (Koutstaal et al., 2001; Simons et al., 2003). Support for a more abstract, language-based representation in the left PPA, specifically, comes from research showing that its BOLD response is sensitive to sentence content that refers to places (Aziz-Zadeh et al., 2008).

The category effect we observed could also reflect the contextual processing of objects via learned associations that has been posited to occur in the PPA (Bar and Aminoff, 2003; Bar et al., 2008) After all, a lifetime of experience with large objects in scenes (buildings, trees, mountains) would almost certainly have resulted in strong learned associations between their visual appearance and their size, and this learning could be linked to contextual processing in the PPA. The effects of such learning could have overridden to some degree the pictorial cues to size present in our experimental displays. In other words, seeing a garage would invoke its familiar real-world size independent of its perceived location vis-à-vis the size/distance cues in the display. Indeed, there is evidence that even scenes with toy objects will activate the PPA (Epstein et al., 1999).

The CS showed a main effect of category, although only when responses were calculated by averaging across hemispheres, with similar results in the control condition tests. The most notable aspect of the preference for garage over camera stimuli in the CS was that it occurred in the absence of a main effect of perceived size. The fact that it was present even in the control condition in which the scene context that offered extrinsic size/distance cues was entirely absent converges nicely on other studies linking this area specifically to the identification of buildings as opposed to scenes more generally

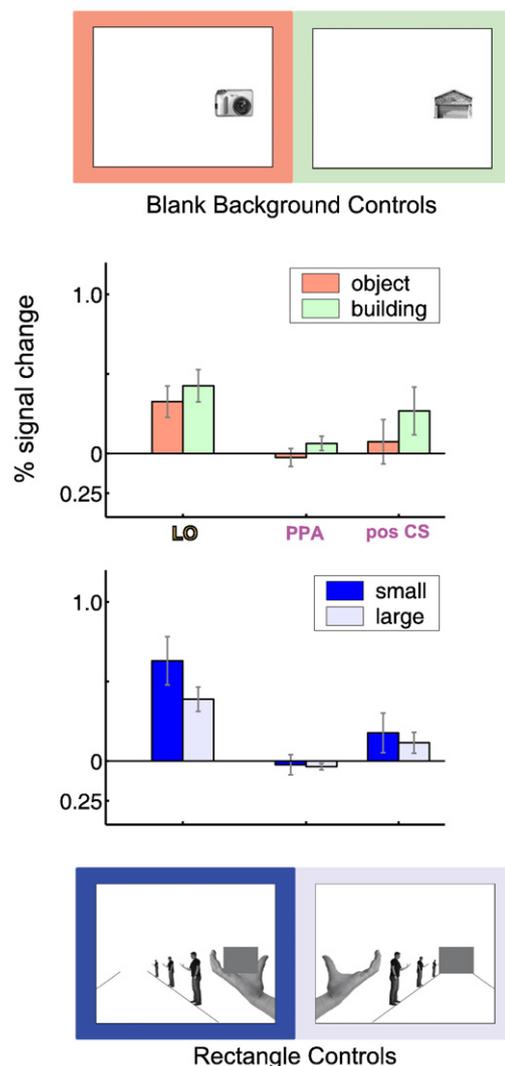


Fig. 3 – Control conditions ROI analysis. Top: example screens from the two blank background control conditions used to isolate effects of stimulus identity. Stimuli were identical to those used in the scene background conditions. Bar graphs plot ROI mean percent BOLD signal change from each of the ROIs (data from both hemispheres combined). Error bars represent standard error of the mean. Bottom: examples of the rectangle stimulus control conditions used to isolate effects of pictorial size/distance. Focal stimuli were rectangles with the same average size and luminance as the object and building stimuli.

(Aguirre et al., 1998). Moreover, the observation that CS activity, unlike PPA activity, was not consistently modulated by our high-level size cues suggests that this region, which is situated posterior to the PPA, is more attuned to perceptual than to conceptual cues. Accordingly, the category specificity in CS could arise from low-level image features that distinguish buildings from objects irrespective of any contextually mediated processing of associations. There was a non-significant trend ($p=0.056$) towards a two-way interaction in the ANOVA that mirrored a similar pattern of results in the behavioral data. Notably, hit rates in oddball detection for blocks with familiar identity/size pairings (e.g., large garages) were slightly higher than those for unfamiliar pairings (e.g., small garages), which was reflected in a non-significant two-way interaction in the ANOVA on hit rates ($p=0.07$). This pattern could suggest that activity in the posterior CS is particularly sensitive to increases in task difficulty due to lack of stimulus familiarity.

3.3. Perceived size or perceived distance?

In our simple 2D displays, perceived size and perceived distance were intrinsically related, i.e., the apparent size of a stimulus was inversely related to its apparent distance from the observer. However, the simple pictorial cues we used are more likely to have influenced perceived stimulus size than perceived distance for a number of reasons. The use of a familiar object (a person) in the scene provided a cue to size that was not dependent on other cues in the display. While a cue to size is ideally also a cue to distance, human observers are poor at judging absolute distance based on familiar size (Gogel and Mertens, 1967; Gogel, 1969; Fitzpatrick et al., 1982; Predebon, 1992; Predebon, 1994). The other size/distance cues in the displays were also unlikely to have provided precise distance information. The corridor illusion we employed has been shown to be useful for making judgments of relative size, even in non-human primates (Barbet and Fagot, 2002; Barbet and Fagot, 2007) but not for absolute size (Greene and Gentner, 2001). Finally, the coarse distance information present in the displays was unlikely to have interacted with the familiar size cues since it has been shown that perceived distance often has little effect on size estimation for familiar objects (Haber and Levin, 2001). In sum, it is therefore more likely that our neuroimaging results reflect manipulations of the perceived size of objects rather than their perceived distance conveyed by the pictorial cues.

3.4. Relation of perceived size effects to retinotopic eccentricity effects

The apparent size effects we observed can be related to the proposed retinotopic size/eccentricity account of ventral stream functional organization (Levy et al., 2001; Hasson et al., 2002; Levy et al., 2004). The eccentricity bias account hypothesizes that the opposing needs to discriminate fine details (faces, foveal stimuli) and to integrate across particular portions of the field of view (buildings, peripheral stimuli) is the basis of the topographical layout of the apparently category-specific ventral stream areas. We found that changes in apparent size modulated object- and building-specific responses in two ventral stream regions, even when retino-

topic size and location were held constant. Indeed, even though eccentricity and retinotopic size may have played a critical role in shaping the evolution and/or ontogenetic development of the functional organization of the ventral stream, the fact that perceived size modulates activity in both LO and the PPA suggests that these apparently category-specific regions may have ultimately escaped complete dependence on eccentricity and retinotopic size as a cue to the nature of the stimuli. Put differently, size selectivity may be considered to be free from spatial specificity in the same way that shape selectivity is typically size- and position-invariant in a number of ventral stream regions (e.g., Schwartz et al., 1983). From this perspective, our results provide a bridge between the traditional category-based account of ventral stream organization and the retinotopic/eccentricity proposal.

A recent study illustrates how apparent size can have the same effects as true retinotopic size differences on fMRI responses. Murray et al. (2006) found that images of spheres with perceived size altered by a pictorial illusion evoked V1 activation that corresponded to the perceived, not the retinal size. Such effects suggest that activity in early visual areas is modulated by top-down information about size. Accordingly, in the current study, top-down information that influences perceived size may have produced the same effects as changing the eccentricity of building and face images did in earlier studies.

We would like to speculate that, once an item is perceived to have the size of a small object, then processing of its shape information engages more heavily an object-recognition pathway that includes LO. Similarly, recognition of items perceived to be large will rely less on activity in LO and more on a pathway including the posterior CS and the PPA. There is anatomical evidence for two parallel processing pathways in anterior ventromedial temporal cortex in the macaque which originate in posterior temporal and parietal regions, progress through different medial temporal regions (the perirhinal and posterior parahippocampal regions in the parahippocampal gyrus, respectively), remain segregated through the entorhinal cortex (lateral and medial), and even project to different subregions in the hippocampus (Lavenex and Amaral, 2000; Witter et al., 2000). Evidence for a similar segregation is also emerging from research based on the examination of resting-state connectivity of different MTL structures in humans (Kahn et al., 2008). Accumulating evidence for perceptual processing in cortical MTL regions, including suggestions that perirhinal cortex is preferentially involved in object as compared to scene processing (Bussey et al., 2003; Bussey and Saksida, 2005; Lee et al., 2005a,b; Bussey et al., 2006; O'Neil et al., 2009), provides functional support for such an organization as well.

In any case, it remains to be investigated exactly how visual processing in such dual pathways would differ. The present experiments, certainly, do not point to anything other than a possible role for perceived size as a “routing” criterion. However, it does seem unlikely that processing of objects and buildings would be strictly segregated in different regions of the temporal and occipital cortex posteriorly; the observation that LO and the posterior CS responded well above fixation to various types of stimuli in the present study provides support for this cautionary note. It seems more likely

that shape processing for both small objects and large buildings would be distributed to some extent across ventral visual cortex (Ishai et al., 1999).

3.5. Conclusion

We conclude that the progressively stronger invariance to retinal size and position found in successively more anterior ventral stream regions does not necessarily imply that size and position information itself is disregarded in these regions. On the contrary, spatial dimensions like size and allocentric position are among the most important visual properties of items in the world; they determine many properties of the retinal images projected by an item, as well as being strong cues to other important properties including distance from the viewer, weight, and clues to probable identity. Thus, it may not be surprising that top-down information about size can be just as important as retinal position and eccentricity in shaping the response profile of regions that are thought to be category-specific.

4. Experimental procedures

4.1. Participants

Nine healthy individuals (two females) participated in the experiment. All participants (mean age, 23.9 years; SD, 5.0; range, 18–34) were right handed, reported normal or corrected-to-normal visual acuity, gave their informed consent to participate in the study, and had no known history of neurological disorder. The participants were members of the University of Western Ontario community. The procedures and protocols for this study were approved by the Research Ethics Board for Health Sciences Research at the University of Western Ontario (London, Ontario, Canada), and this study was therefore performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki. All participants gave their informed consent prior to inclusion in the study.

4.2. Stimuli

4.2.1. Face-scene-object (FSO) localizer stimuli

We used a protocol that has been routinely used in our laboratory to identify category specific ROIs (e.g., Ganel et al., 2006; Ganel et al., 2006; Cant and Goodale, 2007). The stimuli consisted of grayscale photographs of faces, scene images (equal numbers of landscapes, exterior views of houses, and house interiors), and objects (both living and nonliving). A condition consisting of scrambled versions of all three image types was also presented, although this condition did not enter the contrasts employed to define our ROIs. All categories of objects, including scrambled images, in this face-scene-object (FSO) localizer were 250×250 pixels in size. Scrambled images were created by dividing face, scene, and object images into 100 25×25 pixel sections, which were then randomly rotated and redistributed across the image grid. This scrambling destroyed the familiar appearance of the objects while preserving local image features. Although faces were included

in the standard set of stimuli we used for localization, we did not analyze activity in any face-specific regions.

4.2.2. Experimental stimuli

The experimental scans employed 6 types of stimuli. Four of these constituted the crossing of two image parameters and formed the basis of a 2×2 analysis: the category of item presented at fixation (object or building) and the apparent size/distance of the fixated item (near/small or far/large). These four stimulus categories were thus small/near object, large/far object, small/near building, and large/far building (Fig. 1). Object items were photographs of cameras with similar rectangular outline shapes, taken from a frontal viewpoint that minimized perspective foreshortening (a cue to viewing distance). Building items were photographs of garages from a residential neighborhood in London, Ontario. These garages shared similar layouts and outline shapes and were photographed to avoid a view of the roof sloping in the depth direction (another cue to viewing distance). Stimuli included eight different cameras and garages. All images were grayscale, adjusted to have the same mean pixel intensity, and were histogram equalized to equate contrast.

Object and building images were superimposed on a background scene that was designed to convey a sense of depth and perspective using minimal pictorial cues. The scene included two lines converging towards a vanishing point to one side of the image, along which were placed four copies of the same photograph of a person, scaled progressively smaller to indicate receding distance. The foremost person image was depicted to be so close to the viewer that only the outstretched palm was in the field of view of the scene. This hand subtended 9.4° of arc at its widest point. Since the hand had an extent of 14 cm in reality, the nearest image of the hand represented a distance of 85 cm from the observer. The farthest person image subtended 2.8° of arc in height; since the photographed person was 1.9 m tall, he would appear to stand at a distance of about 40 m from the viewer. Note that, at the close position, a camera image would appear to have its appropriate real-world size (approximately 6.9 cm tall), and at the far position, a garage would assume its appropriate size (i.e., 4.2 m tall). The apparent size/distance of the object and building images was manipulated by placing the images so that they either appeared to be resting on the palm (small/near) or on the ground with the frontal edge at the distance of the smallest person (large/far). These two positions had the same vertical position in the frame of the background scene and were equidistant from the scene's center. Two mirror-image versions of the background scene were used to ensure that identification of apparent size/distance was not confounded with position on the presentation screen. Participants were instructed to look directly at the cameras or buildings regardless of its position on the screen. Fixating off-center screen positions could alter ROI activation relative to fixating the screen's center because one side of the screen will extend farther into the retinotopic periphery. However, because use of positions was counterbalanced, with objects and buildings as well as the small/near and large/far cues appearing at both screen positions equally often, any such bias would not be expected to affect any comparisons between the experimental conditions.

Four control conditions were also included in the experimental runs to examine (i) the effects of apparent size/distance in the absence of cues related to stimulus familiarity, and (ii) the baseline responses to objects and buildings in the absence of depth cues (Fig. 3). The size/distance control images were created by drawing rectangles, rather than objects or buildings, at the near or far positions of the background scene. Eight different gray rectangles were employed, with slight variations in brightness and aspect ratio. For the stimulus-baseline control images, the camera and garage images appeared against blank backgrounds, at the same screen positions as were used with the scene background.

To allow for the administration of an oddball-detection task in association with the stimuli described, a single oddball stimulus was designed to be equally difficult to detect when presented amid object, building, and rectangle stimuli. The oddball was rendered in 3D using ray tracing software (POV-Ray v. 3.1, copyright 1991–2010 POV-Team) as two adjoined rectangular solids with a granite texture. The stimulus was a grayscale frontal view of this oddball, with the intensity histogram equalized to be the same as the other stimuli.

4.3. Apparatus

Stimulus presentation was controlled by SuperLab Pro version 2.0.4 (Cedrus Corporation, San Pedro, CA). Each image was projected via an LCD projector (NEC VT540 [NEC Corporation, Tokyo, Japan], screen resolution of 800×600) onto a screen mounted to the outside of the scanner. The distance from the participant's eyes, via a mirror, to the screen was 60 cm. Participants pressed a key on a response pad located on the thigh, and the hit rates and reaction times (RTs) were recorded during performance of the oddball task.

4.4. Procedures

4.4.1. FSO localizer scans

A single run of the FSO localizer consisted of randomly presented blocks of intact face, scene, or object stimuli (4 blocks of each) interleaved with blocks of scrambled images from each category. No images were repeated within or across blocks. Two separate 7.5 min runs (with different block orders) were carried out with each participant prior to the experimental runs. Participants were instructed to maintain central fixation while passively viewing the images. All blocks contained 32 images, each presented for 360 ms and separated by 30 ms interstimulus intervals.

4.4.2. Experimental scans

Prior to entering the scanner, participants were given task instructions and shown examples of the stimuli. Participants were instructed to scrutinize the changing camera and garage stimuli shown in the background scene and to make a button press when they detected the oddball stimulus (Fig. 1). Participants were shown an example of the background scene with no focal stimuli present and also background scenes that had rectangles in the near and far positions; they were instructed to consider anything appearing on the close palm as being near and palm-sized and anything appearing next to the smallest

person image as being appropriately distant and large. Next, participants were shown examples of a camera, a garage, and the oddball stimulus, all on blank backgrounds. The experimenter explained that, during some trial blocks rectangles, cameras and garages would appear embedded in the background scene and that, regardless of an item's familiar size (if any), they should consider it to have the size and distance indicated by its position in the scene. In other words, they were told to expect to see giant cameras and tiny garages in addition to those shown in their familiar size. Participants were also instructed that, in some trial blocks, these items would appear against blank backgrounds.

Participants performed four experimental runs. The screen position of focal target stimuli was constant during a given run (i.e., either towards the left or the right of the screen) and was counterbalanced across experimental conditions, with objects, buildings, and oddball targets occupying the same screen positions equally often in the large and small conditions across blocks and runs. Each run consisted of 16 trial blocks with 5 fixation blocks interspersed; all blocks lasted 18 s. Fixation blocks beginning and ending each run were not analyzed. During fixation blocks, participants fixated a cross that appeared at the center of a blank white background. The trial blocks included all eight unique conditions, repeated twice in a counterbalanced order (two different block orders were used). The main experimental conditions were small/near object, large/far object, small/near building, and large/far building. Additional control conditions were also run. They involved presentation of small/near rectangles, large/far rectangles, objects against a blank background, and buildings against a blank background (see Fig. 3).

A trial block presented 16 focal stimuli, including oddballs. All eight unique exemplars of the relevant category were repeated twice in a pseudorandom order in which all consecutive stimuli differed. Either one or two of these stimuli were replaced with the oddball stimulus within each block; one- and two-oddball blocks occurred with equal but pseudorandom frequency. Oddballs were equally likely to appear at any position in the trial order, except that, when there was a second oddball, it occurred with equal probability at one of the final four trial order positions and it never directly followed the first oddball. This trial-order procedure was designed to encourage participants to attend equally well across the entire block.

All trials within a block belonged to the same stimulus condition. During a trial, the background image and its embedded focal target item appeared for 500 ms. During a 500 ms intertrial interval, the focal item was absent but the background image persisted. Thus, the participants perceived a stable background with focal target items changing quickly. The salience of the house and garage stimuli that changed across trials against a constant background, as well as the fact that the oddball stimuli were presented in the same (off-center) location that was occupied by the changing houses and objects in a given block, made it most likely that participants fixated on these focal target locations rather than on the surrounding background cues throughout each block in all experimental conditions. Also, it should be noted that oddball detection performance was not at ceiling, suggesting that performance was attention-demanding and that participants optimized fixation behavior accordingly.

4.5. Magnetic resonance imaging acquisition

The experiments were carried out with a 4.0 Tesla Siemens-Varian (Erlangen, Germany; Palo Alto, CA) whole-body imaging magnetic resonance imaging system at the Robarts Research Institute, using a head coil. A series of scout images were collected for each participant to select 25 contiguous, 5 mm-thick functional slices in an oblique axial orientation that covered the entire cortex. Functional volumes were collected using a T2*-weighted, navigator echo-corrected, slice interleaved 4 shot spiral imaging pulse sequence (repetition time [TR]=3 s, matrix size=64×64, flip angle=45°, echo time [TE]=15 ms, field of view=220 mm, voxel size=3.4×3.4×5.0 mm). After all the functional scans were completed, T1-weighted anatomical images were obtained using axial slice orientation (3D spiral acquisition with inversion time=1300 ms, TE=3 ms, TR=50 ms, 256×256, slices=128, voxel size=0.86×0.86×1.25 mm).

4.6. fMRI data analysis

Functional data were analyzed using BrainVoyager QX software (Brain Innovation, Maastricht, The Netherlands). Data were preprocessed to remove linear trends across the time course of each voxel and to correct head motion. The FSO localizer data, but not the experimental scan data, were spatially smoothed using a 6 mm FWHM Gaussian kernel. The BOLD time course of each voxel was normalized to percent signal change relative to baseline blocks (scrambled images for the FSO localizer scans and fixation blocks for the experimental scans) and corrected for serial correlation. Finally, all anatomical and functional data were transformed to Talairach stereotaxic space (Talairach and Tournoux, 1988), and each subject's functional images were coregistered to her anatomical volume.

Scene- and object-specific ROIs were determined by contrasting activation evoked by the stimulus of interest with that from the two other FSO stimulus categories. For example, the PPA was identified by contrasting activity in blocks involving presentation of scenes against activity in blocks involving faces as well as those involving objects. Regressors for each FSO condition were convolved with a standard hemodynamic function and fit to the time series of the BOLD percent signal change for each voxel in a general linear model (GLM). A random effects analysis was conducted on the group data to identify regions that were selective for scenes and objects across all subjects, and then each individual's FSO localizer scans were examined to identify the specific voxels meeting the ROI criteria in those general regions. LO ROIs were confined to the lateral occipital lobe in the vicinity of the inferior middle occipital gyrus; posterior fusiform regions that appear to be functionally distinct (pFus; e.g., Grill-Spector, 2003) were not incorporated. PPA ROIs occupied the medial bank of the collateral sulcus and the adjacent parahippocampal gyrus. Posterior CS ROIs occupied the posterior part of the collateral sulcus adjacent to the lingual gyrus. ROIs were identified in both hemispheres when possible. ROIs consisted of clusters of at least 9 contiguous voxels that exceeded a $p < 0.005$ threshold (uncorrected; $t_{(47)} > 3.20$). In order to obtain data from as many participants and hemispheres as possible, a more liberal t -threshold was used when a cluster did not exceed the original threshold. In this

case, ROIs were defined using the most conservative possible t -threshold. In general, activity in reported ROIs met a threshold of at least $t > 2.0$ (uncorrected $p < 0.05$).

4.7. ROI analysis of experimental scans

The ROIs identified with the FSO localizer scans were used to evaluate activation during the experimental scans. GLM regressors for each of the eight experimental task conditions were fit to BOLD time course averaged across all voxels in a given ROI. These regressor beta weights were used to calculate the mean percent signal change of the BOLD response for each experimental condition. Separate analyses examined hemisphere-specific and combined-hemispheres ROIs.

The mean percent signal change measures for the four conditions comprising the 2×2 design (small/near object, large/far object, small/near building, and large/far building) were analyzed in repeated-measures ANOVAs that included image category (object or building) and perceived size (near/small, far/large) as factors. Pairwise comparisons of the four conditions were examined using Tukey's HSD post-hoc test (even though these were planned, not post-hoc comparisons) with $\alpha = 0.05$. The two sets of control conditions (blank background object and building images; small/near and large/far rectangle stimuli images) were analyzed using paired-sample t -tests.

Acknowledgments

This research was funded by the Canadian Institutes of Health Research (CIHR) and the Natural Sciences and Engineering Research Council of Canada (NSERC).

REFERENCES

- Aguirre, G.K., d'Esposito, M., 1999. Topographical disorientation: a synthesis and taxonomy. *Brain* 122 (Pt 9), 1613–1628.
- Aguirre, G.K., Zarahn, E., et al., 1998. An area within human ventral cortex sensitive to "building" stimuli: evidence and implications. *Neuron* 21 (2), 373–383.
- Amit, E., Trope, Y., et al., 2008. A distance principle of organization of the ventral visual stream. *J. Vis.* 8 (6). doi:10.1167/8.6.329 article 329.
- Arcaro, M.J., McMains, S.A., et al., 2009. Retinotopic organization of human ventral visual cortex. *J. Neurosci.* 29 (34), 10638–10652.
- Avidan, G., Levy, I., et al., 2003. Spatial vs. object specific attention in high-order visual areas. *Neuroimage* 19 (2 Pt 1), 308–318.
- Aziz-Zadeh, L., Fiebach, C.J., et al., 2008. Modulation of the FFA and PPA by language related to faces and places. *Soc. Neurosci.* 3 (3–4), 229–238.
- Bar, M., Aminoff, E., 2003. Cortical analysis of visual context. *Neuron* 38 (2), 347–358.
- Bar, M., Aminoff, E., et al., 2008. Famous faces activate contextual associations in the parahippocampal cortex. *Cereb. Cortex* (New York, N.Y.: 1991) 18 (6), 1233–1238 <http://www.ncbi.nlm.nih.gov/pubmed/17934188>.
- Bar, M., Tootell, R.B., et al., 2001. Cortical mechanisms specific to explicit visual object recognition. *Neuron* 29 (2), 529–535.
- Barbet, I., Fagot, J., 2002. Perception of the corridor illusion by baboons (*Papio papio*). *Behav. Brain Res.* 132 (1), 111–115.
- Barbet, I., Fagot, J., 2007. Control of the corridor illusion in baboons (*Papio papio*) by gradient and linear-perspective depth cues. *Perception* 36 (3), 391–402.

- Barrash, J., Damasio, H., Adolphs, R., Tranel, D., 2000. The neuroanatomical correlates of route learning impairment. *Neuropsychologia* 38 (6), 820–836.
- Bussey, T.J., Saksida, L.M., 2005. Object memory and perception in the medial temporal lobe: an alternative approach. *Curr. Opin. Neurobiol.* 15 (6), 730–737.
- Bussey, T.J., Saksida, L.M., et al., 2003. Impairments in visual discrimination after perirhinal cortex lesions: testing “declarative” vs. “perceptual–mnemonic” views of perirhinal cortex function. *Eur. J. Neurosci.* 17 (3), 649–660.
- Bussey, T.J., Saksida, L.M., et al., 2006. Perirhinal cortex and feature-ambiguous discriminations. *Learn. Mem.* 13 (2), 103–105.
- Cant, J.S., Goodale, M.A., 2007. Attention to form or surface properties modulates different regions of human occipitotemporal cortex. *Cereb. Cortex* 17 (3), 713–731.
- Carlson, T.A., Schrater, P., et al., 2003. Patterns of activity in the categorical representations of objects. *J. Cogn. Neurosci.* 15 (5), 704–717.
- Cate, A.D., Goodale, M.A., et al., 2006. The influence of perceived size/distance on object and place ROIs. *J. Vis.* 6 (6). doi:10.1167/6.6.623 article 623.
- Chouinard, P.A., Morrissey, B.F., et al., 2008. Repetition suppression in occipital–temporal visual areas is modulated by physical rather than semantic features of objects. *Neuroimage* 41 (1), 130–144.
- Cohen, L., Gray, F., et al., 1994. Selective deficit of visual size perception: two cases of hemimicropsia. *J. Neurol. Neurosurg. Psychiatry* 57 (1), 73–78.
- Epstein, R., Harris, A., et al., 1999. The parahippocampal place area: recognition, navigation, or encoding? *Neuron* 23 (1), 115–125.
- Epstein, R., Kanwisher, N., 1998. A cortical representation of the local visual environment. *Nature* 392 (6676), 598–601.
- Epstein, R.A., Parker, W.E., et al., 2007. Where am I now? Distinct roles for parahippocampal and retrosplenial cortices in place recognition. *J. Neurosci. Off. J. Soc. Neurosci.* 27 (23), 6141–6149 <http://www.ncbi.nlm.nih.gov/pubmed/17553986>.
- Ewbank, M.P., Schluppeck, D., et al., 2005. fMR-adaptation reveals a distributed representation of inanimate objects and places in human visual cortex. *Neuroimage* 28 (1), 268–279.
- Ferber, S., Humphrey, G.K., et al., 2005. Segregation and persistence of form in the lateral occipital complex. *Neuropsychologia* 43 (1), 41–51.
- Fitzpatrick, V., Pasnak, R., et al., 1982. The effect of familiar size at familiar distances. *Perception* 11 (1), 85–91 <http://www.ncbi.nlm.nih.gov/pubmed/7133938>.
- Frassinetti, F., Nichelli, P., et al., 1999. Selective horizontal dysmetropsia following prestriate lesion. *Brain* 122 (Pt 2), 339–350.
- Ganel, T., Gonzalez, C.L., et al., 2006. The relationship between fMRI adaptation and repetition priming. *Neuroimage* 32 (3), 1432–1440.
- Gerlach, C., Aaside, C.T., et al., 2002. Brain activity related to integrative processes in visual object recognition: bottom-up integration and the modulatory influence of stored knowledge. *Neuropsychologia* 40 (8), 1254–1267.
- Gerlach, C., Law, I., et al., 1999. Perceptual differentiation and category effects in normal object recognition: a PET study. *Brain journal neurology* 122 (Pt 11), 2159–2170.
- Gogel, W.C., 1969. The effect of object familiarity on the perception of size and distance. *Q. J. Exp. Psychol.* 21 (3), 239–247.
- Gogel, W.C., Mertens, H.W., 1967. Perceived size and distance of familiar objects. *Percept. Mot. Skills* 25 (1), 213–225.
- Goh, J.O.S., Siong, S.C., et al., 2004. Cortical areas involved in object, background, and object-background processing revealed with functional magnetic resonance adaptation. *J. Neurosci.* 24 (45), 10223–10228.
- Goodale, M.A., Milner, A.D., et al., 1991. A neurological dissociation between perceiving objects and grasping them. *Nature* 349 (6305), 154–156.
- Gorno-Tempini, M.L., Price, C.J., 2001. Identification of famous faces and buildings: a functional neuroimaging study of semantically unique items. *Brain* 124 (Pt 10), 2087–2097.
- Greene, E., Gentner, S., 2001. Further consideration of size illusions in random dot stereograms. *Percept. Mot. Skills* 93 (1), 205–212.
- Grill-Spector, K., 2003. The neural basis of object perception. *Curr. Opin. Neurobiol.* 13 (2), 159–166.
- Grill-Spector, K., Kourtzi, Z., et al., 2001. The lateral occipital complex and its role in object recognition. *Vis. Res.* 41, 1409–1422.
- Grill-Spector, K., Kushnir, T., et al., 1999. Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* 24 (1), 187–203.
- Grill-Spector, K., Kushnir, T., et al., 1998a. Cue-invariant activation in object-related areas of the human occipital lobe. *Neuron* 21 (1), 191–202.
- Grill-Spector, K., Kushnir, T., et al., 1998b. A sequence of object-processing stages revealed by fMRI in the human occipital lobe. *Hum. Brain Mapp.* 6 (4), 316–328.
- Grill-Spector, K., Malach, R., 2001. fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychol. (Amst)* 107 (1–3), 293–321.
- Haber, R.N., Levin, C.A., 2001. The independence of size perception and distance perception. *Percept. Psychophys.* 63 (7), 1140–1152.
- Hasson, U., Levy, I., et al., 2002. Eccentricity bias as an organizing principle for human high-order object areas. *Neuron* 34 (3), 479–490.
- Haxby, J.V., Ungerleider, L.G., et al., 1999. The effect of face inversion on activity in human neural systems for face and object perception. *Neuron* 22, 189–199.
- Haxby, J.V., Ishai, A., et al., 2000. Object–form topology in the ventral temporal lobe. *TICS* 4 (1), 3–4.
- Hemond, C.C., Kanwisher, N.G., et al., 2007. A preference for contralateral stimuli in human object- and face-selective cortex. *PLoS ONE* 2 (6), e574.
- Ishai, A., Ungerleider, L.G., et al., 1999. Distributed representation of objects in the human ventral visual pathway. *Proc. Natl. Acad. Sci. U. S. A.* 96, 9379–9384.
- James, T.W., Culham, J., et al., 2003. Ventral occipital lesions impair object recognition but not object-directed grasping: an fMRI study. *Brain* 126 (Pt 11), 2463–2475.
- Joseph, J.E., 2001. Functional neuroimaging studies of category specificity in object recognition: a critical review and meta-analysis. *Cogn. Affect. Behav. Neurosci.* 1 (2), 119–136.
- Kanwisher, N., Stanley, D., et al., 1999. The fusiform face area is selective for faces not animals. *NeuroReport* 10 (1), 183–187.
- Kassubek, J., Otte, M., et al., 1999. Brain imaging in a patient with hemimicropsia. *Neuropsychologia* 37 (12), 1327–1334.
- Kohler, S., Crane, J., et al., 2002. Differential contributions of the parahippocampal place area and the anterior hippocampus to human memory for scenes. *Hippocampus* 12 (6), 718–723.
- Kourtzi, Z., Kanwisher, N., 2000. Cortical regions involved in perceiving object shape. *J. Neurosci.* 20 (9), 3310–3318.
- Koutstaal, W., Wagner, A.D., et al., 2001. Perceptual specificity in visual object priming: functional magnetic resonance imaging evidence for a laterality difference in fusiform cortex. *Neuropsychologia* 39 (2), 184–199.
- Larsson, J., Heeger, D.J., 2006. Two retinotopic visual areas in human lateral occipital cortex. *J. Neurosci.* 26 (51), 13128–13142.
- Lavenex, P., Amaral, D.G., 2000. Hippocampal–neocortical interaction: a hierarchy of associativity. *Hippocampus* 10 (4), 420–430.
- Lee, A.C., Buckley, M.J., et al., 2005a. Specialization in the medial temporal lobe for processing of objects and scenes. *Hippocampus* 15 (6), 782–797.
- Lee, A.C., Bussey, T.J., et al., 2005b. Perceptual deficits in amnesia: challenging the medial temporal lobe “mnemonic” view. *Neuropsychologia* 43 (1), 1–11.

- Lerner, Y., Hendler, T., et al., 2001. A hierarchical axis of object processing stages in the human visual cortex. *Cereb. Cortex* 11 (4), 287–297.
- Lerner, Y., Pianka, P., et al., 2003. Area-specific amblyopic effects in human occipitotemporal object representations. *Neuron* 40 (5), 1023–1029.
- Levy, I., Hasson, U., et al., 2001. Center–periphery organization of human object areas. *Nat. Neurosci.* 4 (5), 533–539.
- Levy, I., Hasson, U., et al., 2004. Functional analysis of the periphery effect in human building related areas. *Hum. Brain Mapp.* 22 (1), 15–26.
- MacEvoy, S.P., Epstein, R.A., 2007. Position selectivity in scene- and object-responsive occipitotemporal regions. *J. Neurophysiol.* 98 (4), 2089–2098.
- Maguire, E.A., Frith, C.D., et al., 2001. Distinct neural systems for the encoding and recognition of topography and faces. *Neuroimage* 13 (4), 743–750.
- Makin, T.R., Holmes, N.P., et al., 2007. Is that near my hand? Multisensory representation of peripersonal space in human intraparietal sulcus. *J. Neurosci.* 27 (4), 731–740.
- Malach, R., Levy, I., et al., 2002. The topography of high-order human object areas. *TICS* 6 (4), 176–184.
- Malach, R., Reppas, J.B., et al., 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc. Natl Acad. Sci. USA* 92 (18), 8135–8139.
- Marsolek, C.J., 1995. Abstract visual-form representations in the left cerebral hemisphere. *J. Exp. Psychol. Hum. Percept. Perform.* 21, 375–386.
- Marsolek, C.J., Kosslyn, S.M., et al., 1992. Form-specific visual priming in the right cerebral hemisphere. *J. Exp. Psychol. Learn. Mem. Cogn.* 18 (3), 492–508.
- Milner, A.D., Perrett, D.I., et al., 1991. Perception and action in “visual form agnosia”. *Brain* 114, 405–428.
- Murray, S.O., Boyaci, H., et al., 2006. The representation of perceived angular size in human primary visual cortex. *Nat. Neurosci.* 9 (3), 429–434.
- Niemeier, M., Goltz, H.C., et al., 2005. A contralateral preference in the lateral occipital area: sensory and attentional mechanisms. *Cereb. Cortex* 15 (3), 325–331.
- Niemeier, M., Singh, V.V., et al., 2008. The perceptual consequences of the attentional bias: evidence for distractor removal. *Exp. Brain Res.* 189 (4), 411–420.
- O’Neil, E.B., Cate, A.D., et al., 2009. Perirhinal cortex contributes to accuracy in recognition memory and perceptual discriminations. *J. Neurosci.* 29 (26), 8329–8334.
- Op de Beeck, H., Vogels, R., 2000. Spatial sensitivity of macaque inferior temporal neurons. *J. Neurophysiol.* 426 (4), 505–518.
- Peyrin, C., Baciou, M., et al., 2004. Cerebral regions and hemispheric specialization for processing spatial frequencies during natural scene recognition. An event-related fMRI study. *Neuroimage* 23 (2), 698–707.
- Peyrin, C., Chauvin, A., et al., 2003. Hemispheric specialization for spatial frequency processing in the analysis of natural scenes. *Brain Cogn.* 53 (2), 278–282.
- Predebon, J., 1992. The influence of object familiarity on magnitude estimates of apparent size. *Perception* 21 (1), 77–90.
- Predebon, J., 1994. Perceived size of familiar objects and the theory of off-sized perceptions. *Percept. Psychophys.* 56 (2), 238–247 <http://www.ncbi.nlm.nih.gov/pubmed/7971124>.
- Rosenbaum, R.S., Ziegler, M., et al., 2004. “I have often walked down this street before”: fMRI studies on the hippocampus and other structures during mental navigation of an old environment. *Hippocampus* 14 (7), 826–835.
- Sayres, R., Grill-Spector, K., 2008. Relating retinotopic and object-selective responses in human lateral occipital cortex. *J. Neurophysiol.*
- Schiller, P.H., 1995. Effect of lesions in visual cortical area V4 on the recognition of transformed objects. *Nature* 376, 342–344.
- Schiller, P.H., Lee, K., 1991. The role of the primate extrastriate area V4 in vision. *Science* 251, 1251–1253.
- Schwartz, E.L., Desimone, R., et al., 1983. Shape recognition and inferior temporal neurons. *Proc. Natl Acad. Sci. USA* 80 (18), 5776–5778.
- Schwarzlose, R.F., Swisher, J.D., et al., 2008. The distribution of category and location information across object-selective regions in human visual cortex. *Proc. Natl. Acad. Sci. U. S. A.* 105 (11), 4447–4452.
- Simons, J.S., Koutstaal, W., et al., 2003. Neural mechanisms of visual object priming: evidence for perceptual and semantic distinctions in fusiform cortex. *Neuroimage* 19 (3), 613–626.
- Steeves, J.K., Humphrey, G.K., et al., 2004. Behavioral and neuroimaging evidence for a contribution of color and texture information to scene classification in a patient with visual form agnosia. *J. Cogn. Neurosci.* 16 (6), 955–965.
- Takahashi, N., Kawamura, M., 2002. Pure topographical disorientation—the anatomical basis of landmark agnosia. *Cortex* 38 (5), 717–725.
- Talairach, J., Tournoux, P., 1988. *Co-Planar Stereotaxic Atlas of the Human Brain: 3-D Proportional System: An Approach to Cerebral Imaging*. Thieme0865772932.
- Tootell, R.B., Devaney, K.J., et al., 2008. fMRI mapping of a morphed continuum of 3D shapes within inferior temporal cortex. *Proc. Natl Acad. Sci. USA* 105 (9), 3605–3609.
- Ungerleider, L., Ganz, L., et al., 1977. Size constancy in rhesus monkeys: effects of pulvinar, prestriate, and inferotemporal lesions. *Exp. Brain Res.* 27 (3–4), 251–269.
- Wandell, B.A., Brewer, A.A., et al., 2005. Visual field map clusters in human cortex. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360 (1456), 693–707.
- Whiteley, A.M., Warrington, E.K., 1978. Selective impairment of topographical memory: a single case study. *J. Neurol. Neurosurg. Psychiatry* 41 (6), 575–578.
- Witter, M.P., Naber, P.A., et al., 2000. Cortico-hippocampal communication by way of parallel parahippocampal–subicular pathways. *Hippocampus* 10 (4), 398–410.
- Yaginuma, S., Niihara, T., et al., 1982. Further evidence on elevated discrimination limens for reduced patterns in monkeys with inferotemporal lesions. *Neuropsychologia* 20 (1), 21–32.