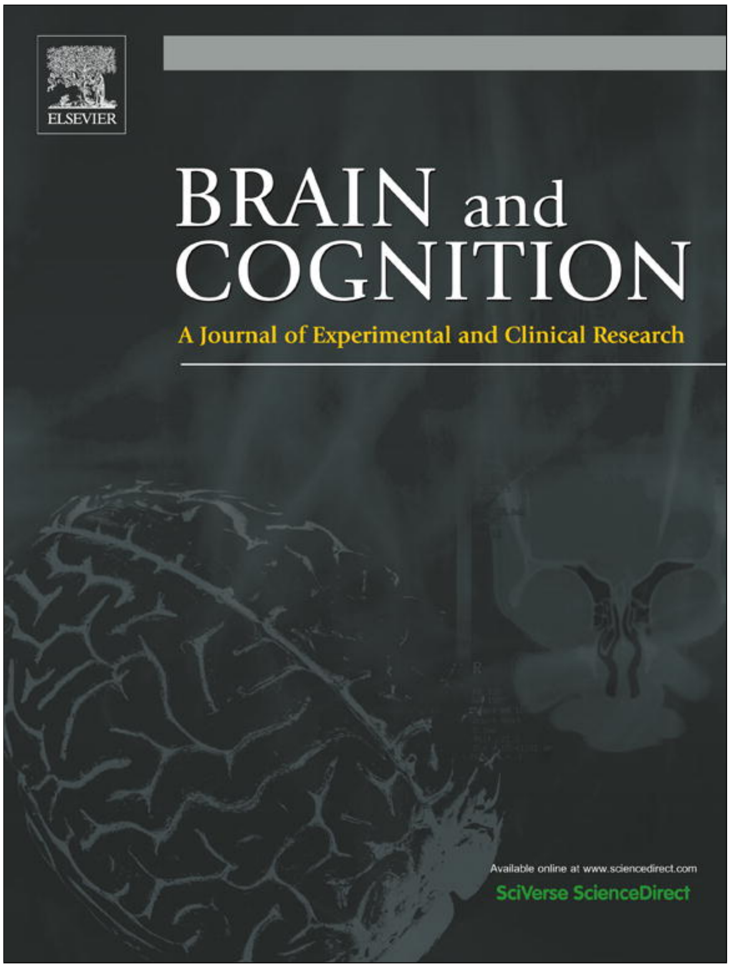
ELSEVIER
BRAIN and COGNITION
A Journal of Experimental and Clinical Research
Available online at www.sciencedirect.com
SciVerse ScienceDirect

Brain and Cognition 80 (2012) 201–213

Contents lists available at SciVerse ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

Do object-category selective regions in the ventral visual stream represent perceived distance information?

Elinor Amit [a,*], Eyal Mehoudar [b], Yaacov Trope [c], Galit Yovel [d]

[a] Department of Psychology, Harvard University, William James Hall 1484, 33 Kirkland Street, Cambridge, MA 02138, United States
[b] Department of Psychology, Tel Aviv University, Tel Aviv, Israel 69987, Israel
[c] New York University, 6 Washington Place, Room 768, New York, NY 10003, United States
[d] School of Psychological Sciences & the Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel 69987, Israel

ARTICLE INFO

Article history:
Accepted 21 June 2012

Keywords:
Ventral visual stream
Distance
Objects
Scenes
Perceived distance

ABSTRACT

It is well established that scenes and objects elicit a highly selective response in specific brain regions in the ventral visual cortex. An inherent difference between these categories that has not been explored yet is their perceived distance from the observer (i.e. scenes are distal whereas objects are proximal). The current study aimed to test the extent to which scene and object selective areas are sensitive to perceived distance information independently from their category-selectivity and retinotopic location. We conducted two studies that used a distance illusion (i.e., the Ponzo lines) and showed that scene regions (the parahippocampal place area, PPA, and transverse occipital sulcus, TOS) are biased toward perceived distal stimuli, whereas the lateral occipital (LO) object region is biased toward perceived proximal stimuli. These results suggest that the ventral visual cortex plays a role in representing distance information, extending recent findings on the sensitivity of these regions to location information. More broadly, our findings imply that distance information is inherent to object recognition.

1. Introduction

It is well established that scenes and objects elicit a highly selective response in specific brain regions in the ventral visual cortex (e.g., Epstein & Kanwisher, 1998; Kanwisher, Chun, McDermott, & Ledden, 1996). Notably, those categories differ is their perceived distance from the observer (i.e. scenes are distal whereas objects are proximal). The goal of the current study is to test to what extent category selective areas in the ventral visual stream play a role in representing distance information, such that scene regions are biased towards distal information whereas object regions are biased towards proximal information, independently from their category-selectivity and retinotopic location.

1.1. The representation of object categories in the ventral visual stream

Neuroimaging studies in the past 15 years have established that certain object categories elicit a highly selective response in specific brain regions in the ventral visual stream (e.g., Aguirre, Zarahn, & D'Esposito, 1998; Epstein & Kanwisher, 1998; Kanwisher et al., 1996; for a recent review, see Op de Beeck, Haushofer, & Kanwisher, 2008). Such selective neural response has been reported for faces (Kanwisher & Yovel, 2006), places (Epstein & Kanwisher, 1998), body parts (Downing, Jiang, Shuman, & Kanwisher, 2001), and words or letter strings (Cohen et al., 2000). The underlying assumption of this line of research is that responses of category-selective regions in the ventral visual stream can be explained by higher-level object representations. For example, a "scene" contains information about the spatial structure of the local environment, and this triggers the response in scene regions (the parahippocampal place area, PPA, and transverse occipital sulcus, TOS, Epstein, 2005; Epstein, Higgins, Jablonski, & Feiler, 2007). With respect to objects, Epstein (2005) suggested that objects are spatially compact entities that one acts upon, in contrast to scenes, which are spatially distributed entities that one acts within. Likewise, Mazer and Gallant (2000) suggested that in order to recognize a single object, constituent patches must be segmented from the background and grouped together into a coherent whole. In other words, in both Epstein and Mazer et al.'s definitions, objects are defined by their inverse relationships with background.

Consistent with the research on the ventral stream, the ventral/dorsal stream theory (Mishkin, Ungeleider, & Macko, 1983) suggests that as visual information exits the occipital lobe, it follows two main paths, or "streams." The dorsal stream terminates in the parietal lobe and represents the location of an object (the "where pathway") or an "action" towards the object (Goodale, Milner, Jakobson, & Carey, 1991). The ventral stream in the temporal

---

[*] Corresponding author.
E-mail address: amit@wjh.harvard.edu (E. Amit).

0278-2626/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.bandc.2012.06.006

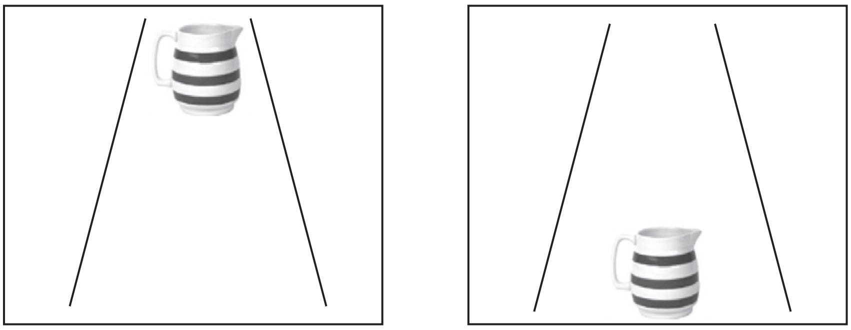

Fig. 1. An example of the Ponzo illusion.

lobe is involved with object identification (the "what pathway"). Lesions to this region of the brain often produce difficulties in recognizing, identifying and naming different categories of objects (Habib & Sirigu, 1987; McNeil & Warrington, 1993; Moscovitch, Winocur, & Behrmann, 1997).

1.2. Beyond shape: location biases in the ventral visual stream

Recent studies have shown that at least some of the object category areas in the ventral "what" stream are sensitive not only to shape, but also to retinotopic location – that is, the location of a target relative to fixation (center versus periphery). Specifically, it was found that scene-selective regions are biased toward stimuli presented in the upper visual field, whereas object-selective regions are biased toward stimuli presented in the lower visual field (Schwarzlose, Swisher, Dang, & Kanwisher, 2008). Furthermore, object areas were found to be biased towards foveal information, whereas scene areas are biased towards peripheral information (Hasson, Levy, Behrmann, Hendler, & Malach, 2002).

In addition to these retinotopic biases, recent studies also suggest that these category-selective regions are sensitive to the size of the stimuli. Specifically, it was shown that scene regions (PPA and TOS) are biased toward stimuli that are generally classified as large (henceforth, "large stimuli"), whereas object region (lateral occipital, LO) is biased toward small stimuli (Cate, Goodale, & Kohler, 2011; Konkle & Oliva, 2010). In sum, the findings above suggest that the ventral stream includes information other than just the category of the target (e.g., face, object, or place) but also its size and location. However, a unified account of these location and size biases and the selectivity for objects and scenes is still lacking.

1.3. Perceived distance as an integrating variable

One factor that may be common to all four types of biases (i.e., object category, location, retinotopy, and size) is the perceived distance between the observer and the visual stimuli – that is, *egocentric distance*. In particular, scenes are typically farther away whereas objects are more proximal. Furthermore, distal stimuli usually occupy the upper part of the visual field (e.g., Bruno & Cutting, 1988; Cutting & Vishton, 1995; Epstein, 1966; Gibson, 1950, p.180; Yonas, Elleiff, & Arterberry, 2002) and do not require foveal vision for recognition, whereas proximal stimuli usually occupy the lower part of the visual field and do require foveal vision for recognition (Levy, Hasson, Avidan, Hendler, & Malach, 2001). Thus, the location and eccentricity biases in scene and object regions may reflect an egocentric distance bias. Finally, size and distance are inherently related to each other (in the realm of visual perception). Specifically, when presented with two items identical in size, the distal item is perceived to be larger than the proximal item (e.g., Murray et al., 2006; Smith, 1958). Thus, it is possible that effects that were attributed to size could actually be explained by distance. While size and distance are typically confounded, given that object and scene selective areas also showed biases to location (down/up) and eccentricity – with both being associated with distance but not size – we believe that it is distance and not size that may underlie all reported biases in object and scene areas.

In sum, perceived distance can account for the location, eccentricity, and size biases reported above. Consistent with the view of the ventral/dorsal stream theory (Mishkin et al., 1983), it is possible that this distance bias in the "what" stream reflects the ideal distance of a given stimulus from the observer to allow for intact identification.

The current study tested whether category selective areas in the ventral visual stream represent egocentric distance information. In order to test the hypothesis, in both Experiments 1 and 2, we took advantage of the well-known distance illusion: the Ponzo illusion (e.g., Yi Li & Guo, 1995). This illusion is comprised of two lines that converge at the upper end (see Fig. 1 for an illustration). As a result, the upper end is perceived as more distal than the lower end. The evidence for the existence of such an illusion comes from the estimation of a "distal" (upper) item as larger than the "proximal" (lower) item, when in reality they are equal in size. We first pretested the effectiveness of the distance illusion using a psychophysical size estimation task and then examined the response of object and scene category regions to the distance manipulation.

Our main hypothesis concerns the representation of perceived distance in the ventral visual stream. Although there are several candidate regions for testing the distance hypothesis, we chose, based on the literature reviewed above, four representative regions: scene regions (the parahippocampal place area, PPA, Epstein & Kanwisher, 1998; and transverse occipital sulcus, TOS, Epstein et al., 2007) and object regions (LO; and posterior fusiform gyrus: pFs; both comprises together the LOC: the Lateral Occipital Cortex, e.g., Malach et al., 1995). We hypothesized that object regions (LO and pFs) would be biased toward perceived proximal stimuli, whereas scene regions (PPA and TOS) would be biased toward perceived distal stimuli.

In Experiment 1 we tested this hypothesis by presenting buildings and objects within the Ponzo lines, in a "distal" (upper) or "proximal" (lower) locations. Objects and buildings were used as they were found to elicit a strong response in the LOC object area and the scene areas respectively (e.g., Malach et al., 1995 Epstein and Kanwisher, 1998). From a theoretical point of view, the PPA response to buildings may be due to the fact that buildings are stable objects that often play a role in defining the spatial structure of the environment (Epstein, 2005). In Experiment 2 we tested the effect of distance independently of location of the target item, by presenting the stimuli in a fixed location at the center of the screen, while the Ponzo lines are presented in either the upper half or the lower half of the screen. Importantly, in both experiments, our perceived distance manipulation entailed egocentric and not retinotopic distance. In other words, stimuli were always presented foveally and perceived distance was defined as a function of the location of the stimulus in the Ponzo lines rather than retinotopic location. Finally, we collected eye position data outside of the scanner with the exact same design and stimuli as the experiment in the scanner, to confirm that subjects fixate on the stimuli and to rule out the possibility that retinotopic and not spatiotopic effects accounts for our results.

2. Materials and methods

Our study included two fMRI experiments. Each fMRI experiment was preceded by a psychophysical experiment that took place outside the scanner, in which we assessed whether the particular Ponzo display that we used in the fMRI experiment indeed generates a distance illusion. Finally, we collected eye position data outside of the scanner.

2.1. Methods for psychophysical experiments

To assess whether the Ponzo lines generated a distance illusion we took advantage of the well-established effect of distance on size perception (e.g., Murray, Boyaci, & Kersten, 2006; Smith, 1958) (see Fig. 1 for an illustration). Participants were asked to compare the size of two same-identity stimuli, which appeared simultaneously in the upper and lower parts of two lines that converged toward the upper end (i.e., the Ponzo lines). The participants were asked to judge which stimulus was larger. A distance illusion is evident if the stimulus at the upper end of the Ponzo lines is judged to be larger than the stimulus at the lower end of the Ponzo lines, despite the fact that they are identical in size.

2.1.1. Participants

Ten New York University undergraduates participated in the two psychophysical manipulation-check experiments, for partial fulfillment of course credit. The study was approved by the University Committee on Activities Involving Human Subjects (UCAIHS).

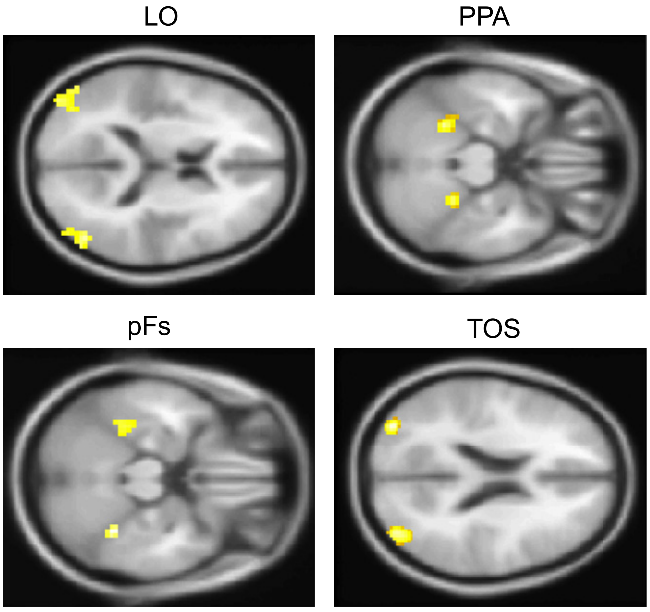

Fig. 2. Regions of interests of representative participants: Lateral occipital cortex (LO – object selective); posterior fusiform gyrus (pFs – object selective); parahippocampal place area (PPA – scene selective); and transverse occipital sulcus (TOS – scene selective).

2.1.2. Stimuli

The stimuli consisted of black and white photographs of 20 objects and 20 buildings. Images of buildings were taken from the site http://www.zoomap.co.il/default.asp, and were selected randomly (See Fig. 2 for an illustration). Objects were adapted from Epstein and Kanwisher (1998). A pair of vertically oriented converging straight lines, the Ponzo illusion, was used to create the illusion of depth. In Experiment 1, the Ponzo lines transversed the entire computer screen. In Experiment 2, the Ponzo lines appeared either at the upper half or the lower half of the screen. In both experiments, two same-identity stimuli appeared simultaneously at the upper end and at the lower end of the Ponzo lines. The stimuli appeared in five different sizes. In Experiment 1 (full screen Ponzo), the standard image that was used in the fMRI experiment and labeled here “100%,” was 4.5 cm in width, and 4.5 cm in length. In Experiment 2 (half screen Ponzo) the width of the standard image was 2.6 cm and the length was 2.6 cm. We created four additional stimuli that were 98%, 99%, 101% and 102% of the size of the standard stimuli.

2.1.3. Procedure

Stimuli were presented using Superlab software. Trials of objects and buildings were presented in separate blocks, with a rest period in the middle of each block. Each trial consisted of two stimuli that appeared simultaneously at the upper end and at the lower end of the Ponzo lines. The two stimuli included the standard stimulus (size 100%) and same-identity stimulus in one of the five different sizes (98%, 99%, 100%, 101% or 102%). The location of the standard stimuli was at the upper part of the screen on half of the trials, and the lower part on the other half.

In Experiment 1, the design was as follows: 2 Categories (Buildings, Objects) × 5 Sizes (98%, 99%, 100%, 101%, 102%) × 2 Locations of the standard stimulus (Upper, Lower). Each of the combinations was presented twice for each of the 20 Building and the 20 Object stimuli, making a total of 800 presentations. On each trial, the Ponzo lines with the two stimuli embedded in them appeared for 800 ms.

In Experiment 2, the design was as follows: 2 Categories (Buildings, Objects) × 5 Sizes (98%, 99%, 100%, 101%, 102%) × 2 Locations of the standard stimulus (Upper, Lower) × 2 Locations of the Ponzo lines (Upper screen, Lower screen). Because this experiment included an additional condition (the location of the Ponzo), to keep the length of this experiment similar to Experiment 1, each of the unique combinations was presented once for each of the 20 Building and 20 Object stimuli, making a total of 800 presentations. On each trial, the Ponzo lines with the two stimuli embedded in them appeared for 800 ms.

In both experiments, after viewing the stimulus, the following question was presented: “Please indicate which of the two stimuli seemed larger: Press (A) if it was the upper, (H) if the lower or (L) if they seemed equal.”

2.2. Methods for the fMRI experiments

The fMRI experiments tested the prediction that object regions (the two parts of the LOC: LO, and pFs) would be biased toward perceived proximal information whereas scene regions (PPA and TOS) would be biased toward perceived distal information. We therefore ran a localizer task to identify object and scene regions in each subject (See Fig. 3 for representative participants). Subsequently, we examined the response of these areas to the different conditions of the main task (i.e., perceived distance).

2.2.1. Participants

Twenty-six healthy participants participated in Experiment 1 and 14 participants participated in Experiment 2. All participants

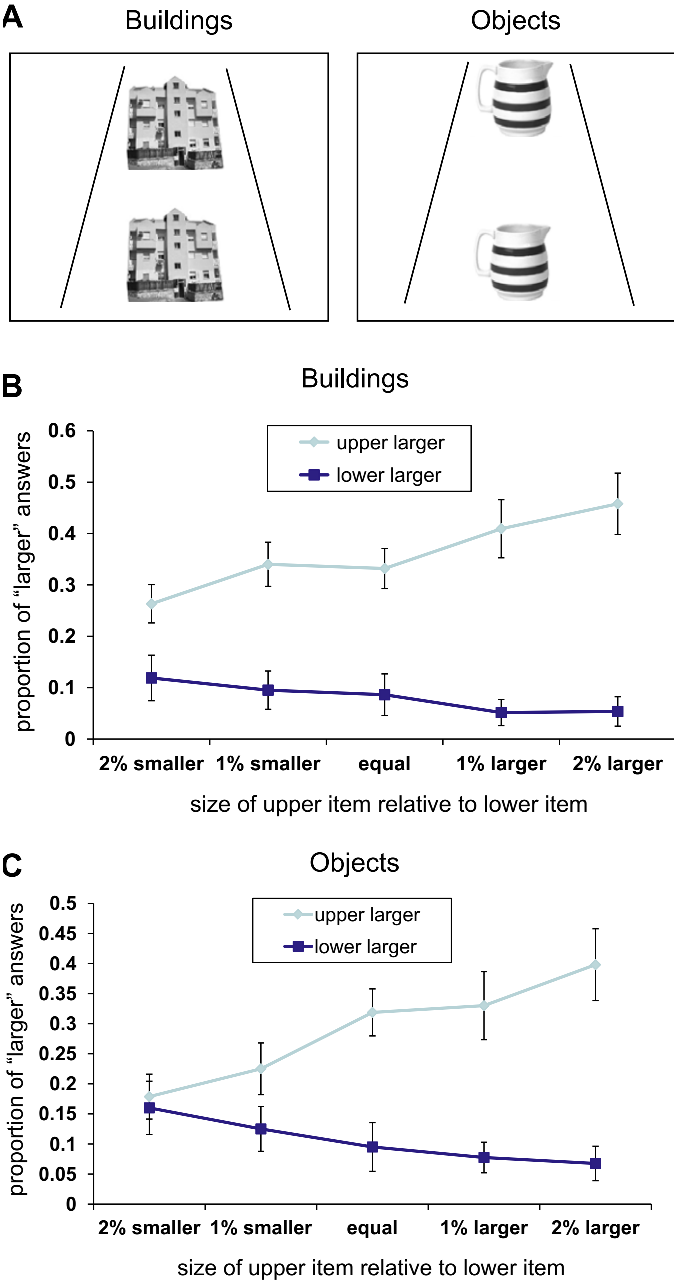

Fig. 3. A. The psychophysical experiment included objects or buildings that were presented simultaneously on the distal and proximal ends of two converging lines (the Ponzo illusion). The size of the two stimuli was either identical or different (±1 and ±2% difference). Participants were asked to indicate which of the two stimuli seemed larger: the upper, the lower, or whether they seem equal. B. The target stimuli were buildings. *Y* axis represents the proportion of "larger" answers as a function of size of the upper stimulus relative to the lower stimulus. Upper stimulus was judged as larger than the lower stimulus in all size conditions. C. The target stimuli were objects. *Y* axis represents the proportion of "larger" answers as a function of size of the upper stimulus relative to the lower stimulus. Upper stimulus was judged as larger for all size conditions except when the upper stimulus was smaller in 2% than the lower stimulus.

had normal or corrected to normal vision and signed a consent form that was approved by the Helsinki committee of the Sourasky Medical Center, Israel.

2.2.2. Stimuli

2.2.2.1. Localizer. In order to define functional regions of interest (fROI) we ran a localizer experiment. The stimuli consisted of three categories of items: Scenes, Objects, and Scrambled objects. Scrambled objects were designed as Fourier power-spectrum controls for the object pictures. They were constructed by performing a two dimensional (2D) Fourier transformation on high-pass filtered pictures of the same objects (Malach et al., 1995). Each category included 80 different images. The images were digitized grayscale photographs (300 × 300 pixels) that were adapted from previous studies (Epstein, Harris, Stanley, & Kanwisher, 1999; Epstein & Kanwisher, 1998).

2.2.2.2. Main task. The stimuli – buildings and objects - were the "standard stimuli" used in the psychophysical experiments. Thus, the stimuli in the fMRI experiments came in one size. In Experiment 1, the stimuli were 4.5 cm in width, and 4.5 cm in length. In Experiment 2, the stimuli were 2.6 cm and the length was 2.6 cm. We used the same Ponzo illusion that was used in the psychophysical experiments. In Experiment 1 (full screen Ponzo), the stimuli (buildings or objects) were presented either at the upper ("distal") or the lower ("proximal") part of the Ponzo lines. In Experiment 2 (half screen Ponzo), the stimuli, buildings and objects, were placed in a fixed location at the center of the screen, thus appearing either at the upper end of the lower Ponzo lines ("distal") or at the lower end ("proximal") of the upper Ponzo lines.

2.2.3. fMRI data acquisition

Scanning was performed with a GE 3.0-T scanner at the Tel Aviv Sourasky Medical Center, Israel. For the functional scans, we collected 30 slices aligned parallel to the temporal lobe, TR = 2000 ms TE = 35 ms, flip angle = 90. We also collected high-resolution anatomical images (SPGR) including 170 sagittal slices (TR = 6.5 ms; TE = 2 ms, FOV = 256).

2.2.4. Procedure

Stimuli were presented using Psychtoolbox implemented in MATLAB (Brainard, 1997). Each fMRI experiment consisted of two runs of the localizer and two runs of the experimental task, which were presented in an interleaved manner, starting with the localizer.

2.2.4.1. Localizer scan. The localizer aimed to localize scene and object regions, and included 2 scans, each consisting of 17 consecutive 16 s epochs, four for each category (Scenes, Objects and Scrambled objects) and five fixation blocks (total time 272 s). Each block included 20 stimuli. The serial position of the categories was counterbalanced across and within scans. Each image was presented for 300 ms with 500 ms blank intervals between stimuli. Participants were instructed to press a key whenever they noticed two images repeated in a row (1-back). This happened twice per epoch and ensured that participants were awake and attentive.

2.2.4.2. Main task. A 2 × 2 design in which we manipulated Perceived Distance (Distal, Proximal) × Category (Buildings, Objects) was used. Each condition was presented in a separate 16 s block. Each scan consisted of 21 blocks, four of each of the four experimental conditions interleaved with five fixation blocks. Each block included 20 stimuli. Each stimulus was presented for 750 ms with 50 ms inter stimulus interval. On each trial, participants saw a background image of the Ponzo lines, with an object or a building embedded in it, either in a "proximal" location, or a "distal" location. Participants were instructed to fixate on the stimuli and press a key whenever they noticed two identical images repeated in a row (1-back).

2.2.5. Data analysis

2.2.5.1. Region of Interest (ROI) analysis. The object-selective regions were defined as voxels in the lateral occipital (LO) cortex and in the posterior fusiform gyrus (pFs) that show a higher response to objects than scrambled objects ($p < .0001$, uncorrected for multiple comparisons). The scene-selective regions were defined as voxels in the parahippocampal gyrus (PPA) and the transverse occipital sulcus (TOS) bilaterally that show a higher response to scenes than objects ($p < 0.0001$, uncorrected). Table 1 indicates the number of participants who showed each of the regions of interest (ROIs) within each hemisphere in each of our fMRI experiments.[1]

2.2.5.2. Data analysis of main task. To test the perceived distance hypothesis, we compared the responses of each of the scene- and object-selective regions to perceived distal and perceived proximal buildings and objects. The peak of the hemodynamic response (averaged across the 6th and 10th TRs – temporal resolution units) was used as the dependent measure. For both Experiment 1 and 2, we found no main effect of hemispheric laterality or interaction of laterality with any other variable; therefore laterality was excluded as a factor from all subsequent analyses.

For Experiment 1 (full screen Ponzo), we report a repeated measure ANOVA of percent signal change, in which Region (Scene-ROI, Object-ROI), Category (Buildings, Objects) and Perceived Distance (Distal, Proximal) were treated as within-subject factors. Because not all our participants showed activation in all four areas, in order to better represent the data we collected from all of our participants we performed separate analyses for each pair of scene and object ROIs. For paired-comparisons of the perceived distance effect within each ROI, we applied a one-tailed t-test in the direction predicted by the perceived distance hypothesis.

In Experiment 2 (half-screen Ponzo), since the psychophysical experiment showed that the illusion was effective for building stimuli but less so for object stimuli (see below), the analysis for buildings and for objects was conducted separately. We first report a repeated measure ANOVA of percent signal change, with Region (Scene-ROI, Object-ROI), and Perceived Distance (Proximal, Distal) as within-subject factors for buildings, and then for objects. For paired-comparisons of the distance effect within each ROI, a one-tailed t-test was applied, in the direction predicted by the distance hypothesis.

2.3. Methods for fixation checks

2.3.1. Participants and procedure

Eight participants participated in each of the two experiments. Participants were seated about 45 cm from the monitor and were presented with exactly the same display and procedure that we used in the fMRI experiments. Informed consent was obtained and approved by the ethics committee of the Psychology Department at Tel Aviv University.

2.3.2. Apparatus

We used an SMI remote eye-tracker (RED system, SensoryMotoric Instruments, Germany), and sampled pupil centroid at 60 Hz. The default nine point calibration and validation sequences were administered at the beginning of the experiment. Calibration was repeated when maximum error at validation was more than 0.9°.

2.3.3. Software

Fixation and area of interest (AOI) data were analyzed through the use of custom built Python-based software libraries. Stimuli were presented using Psychtoolbox implemented in MATLAB (Brainard, 1997).

Table 1
Show the number of participants for whom ROI were identified, in each of the two fMRI experiments. In Experiment 1, 17 participants showed activation in all regions, and in Experiment 2, 11 participants showed activation in all regions. Number of participants that showed pattern of activation that is consistent with the prediction in each region was the following. Experiment 1: *buildings*: LO-17; pFs-11; PPA-18; TOS-15. *Objects*: LO-19; pFs-12; PPA-13; TOS-10. Experiment 2: *buildings:* LO-12; pFs-8; PPA-8; TOS-9. *Objects*: LO-4; pFs-4; PPA-9; TOS-11.

Region	Experiment 1 ($n = 26$)	Experiment 2 ($n = 14$)
Right PPA	23	11
Left PPA	21	11
Right TOS	17	11
Left TOS	18	11
Right LO	24	14
Left LO	23	14
Right pFs	23	14
Left pFs	23	11

2.3.4. AOI analyses

We defined the AOI as a rectangular area around the object/building and counted the relative frequencies of fixations within and outside the AOI. We then assessed the frequencies of fixations (separately for each of the two experiments) within the rectangular area as a function of the experimental conditions. Invalid fixations (i.e., fixations outside the screen or blink related errors) were excluded in this analysis (3.8% in the full-screen Ponzo experiment and 2.7% in the half-screen Ponzo experiment).

2.4. Behavioral results of the 1-back task

Participants were instructed to fixate on the stimuli and press a key whenever they noticed two identical images repeated in a row (1-back). We recorded subject key presses during the experiment and calculated the proportion of hits for the 1-back task. We then assessed hit rate as a function of the experimental condition separately for each of the two experiments.

3. Results

3.1. Experiment 1: full screen ponzo display

3.1.1. Psychophysical experiment

The stimulus display and the results are presented in Fig. 3A–C. We compared the proportion of responses “upper stimulus is larger” to the proportion of responses “lower stimulus is larger” (note that the proportion of the third response, “stimuli are equal”, is complementary to the other two responses and thus redundant.) A higher proportion of responses that indicates that the upper stimulus looks larger than the lower would reflect the existence of a distance illusion.

We conducted a 5 (size of the non-standard stimulus: 2% smaller than the standard, 1% smaller than the standard, equal, 1% larger than the standard, 2% larger than the standard) × 2 (response: upper stimulus is larger, lower stimulus is larger) × 2 (stimulus: buildings, objects) × 2 (standard stimulus (100%) located at the upper versus lower end of the Ponzo lines) repeated measures analysis of variance. There was no main effect or interaction with the location of the standard stimulus. There was a main effect for size $F(4, 36) = 10.53$, $p < .0001$, such that the larger the non-standard stimulus was, the less likely subjects were to respond “equal” (compared to “upper larger” and “lower larger”). There was also a main effect for response, $F(1, 9) = 21.2$, $p < .0001$, such that subjects

---

[1] In defining regions of interests, the contrast between the ROI conditions (e.g., places versus objects) did not elicit any clear activation for some subjects. In these cases (which were a minority), we say that the “regions were not found” – meaning that the functional contrast between the two ROI experimental conditions failed to create reliable regions. For further elaboration on this topic, check Saxe, Brette, and Kanwisher (2006).

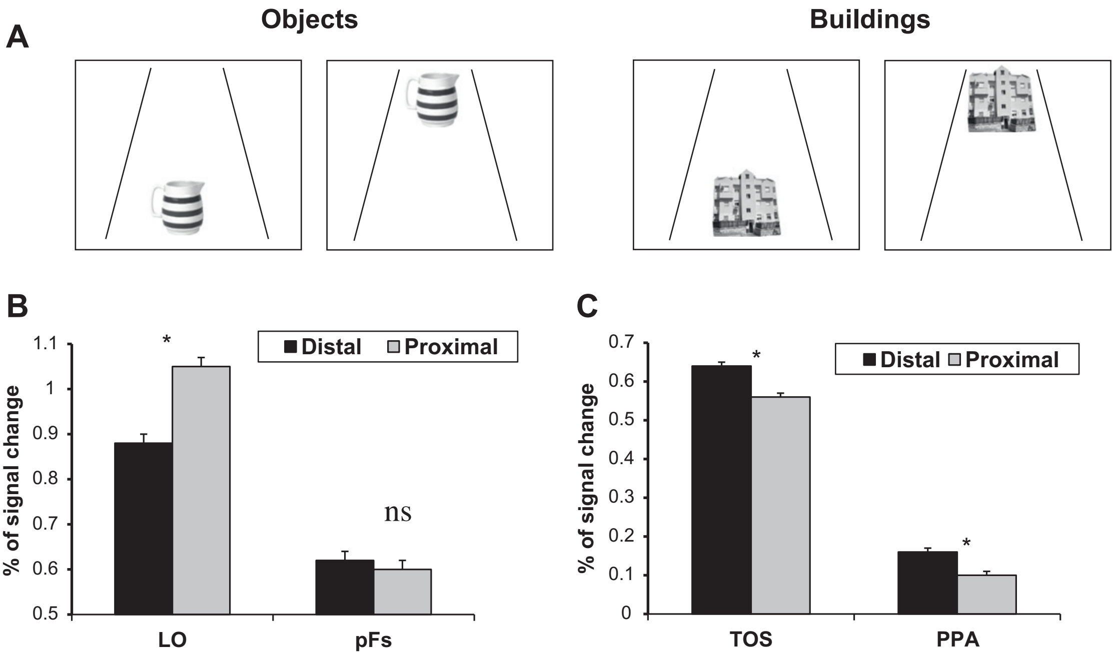

Fig. 4. A. The fMRI experiment included blocks of objects or buildings presented at the proximal or distal ends of two converging lines (the Ponzo illusion). Participants fixated on the stimuli and performed a 1-back task. B. The lateral occipital (LO) object area showed higher response to stimuli (objects and buildings) presented in the proximal than distal end of the screen. The posterior fusiform (pFs) object area showed no distance bias. $^{*} < .05$. C. Scene-selective regions (PPA and TOS) showed higher response to stimuli (objects and buildings) presented in the distal than proximal end of the Ponzo lines.

were more likely to respond "upper is larger" than "lower is larger". In addition, there was an interaction between size and response, $F(4, 36) = 20.3$, $p < .0001$, such that the larger was the upper stimulus, the larger was the difference between the two responses (upper larger, lower larger). Finally, there was an interaction between stimulus (objects, buildings) and response, $F(1, 9) = 7.1$, $p < .05$, such that the difference between the responses "upper larger" and "lower larger" was larger for buildings than for objects. Due to this interaction we looked separately at the simple effects of response within each size condition, for buildings and for objects. Since there was no effect for the location of the standard stimulus, we collapsed the data across this factor. The results show that for buildings, the upper stimulus was judged as larger in all size conditions, $p < .05$. For objects, the upper stimulus was judged as larger when it was actually 2%, 1% larger, or equal to the lower stimulus, $p < .05$. The same effect was marginal when the lower stimulus was 1% larger, $p = .07$ and not significant when it was 2% smaller, $p > .05$.

In sum, as expected, the psychophysical experiment suggests that the Ponzo illusion is effective in producing a distance illusion for buildings. For objects, the effect was somewhat weaker but still appeared when the stimuli were equal and even (to a lesser extent) when the upper object was 1% smaller.

3.1.2. fMRI experiment

The stimulus display and the results are presented in Fig. 4A–C, which shows the effect of perceived distance on the response of the scene- and object-selective regions across stimulus category. In all analyses reported below, the interaction of Region and Category was significant ($p < .01$), indicating the expected higher response to objects than buildings in the object-selective regions (LO and pFs) and to buildings than objects in the scene-selective regions (TOS and PPA). To examine our main hypothesis that scene regions are biased to perceived distal stimuli and object regions are biased to perceived proximal stimuli, we report below the direction of the perceived distance effect in each of the regions and then the interaction of Region and Perceived Distance. Because activation in all of the four areas (LO, pFs, TOS, PPA) was not found in all of our participants, the number of participants included in each of these analyses is different.

The results are consistent with our perceived distance hypothesis. In the LO (object region, see Fig. 4B) responses were higher to perceived proximal than to perceived distal stimuli (*Ms* = 1.04 and 0.88, respectively, $t(23) = 5.3$, $p < .01$).[2] In contrast, in the PPA (scene region, see Fig. 4C) responses were higher to perceived distal than perceived proximal stimuli (*Ms* = 0.17 and 0.1, respectively, $t(23) = 2.27$, $p < .05$). Similarly to the PPA, in the TOS (scene region) responses were higher to perceived distal than perceived proximal stimuli (*Ms* = 0.64 and 0.56, respectively $t(18) = 2.6$, $p < .05$). The effect of distance on response in the pFs was not significant.

3.1.2.1. LO versus PPA. Consistent with the perceived distance hypothesis, analysis of the 22 participants who had activations in both the LO and the PPA showed a significant interaction between Region and Distance, $F(1, 21) = 36.6$, $p < .01$. The 3-way interaction between Region, Distance, and Category was not significant, $F < 1$.

3.1.2.2. LO versus TOS. In line with the perceived distance hypothesis, analysis of the 18 participants who had activations in both the LO and the TOS showed a significant interaction between Region and Distance, $F(1, 17) = 28.47$, $p < .01$.

In addition, a three-way interaction of Region, Distance, and Category, $F(1, 17) = 5.24$, $p < .05$ indicates that the perceived proximal bias in LO did not interact with Category, $F(1, 17) < 1$, whereas for the TOS, the perceived distal bias was larger for the buildings than the objects (*Ms* for the difference = 0.13 and 0.03, respec-

---

[2] In this and subsequent analyses, *t*-tests for each area included all participants who showed activation in that region. When the effect is compared across two areas using repeated measures ANOVA, participants that showed activation of both regions (e.g., PPA and LO) are included.

tively) (a category by distance interaction $F(1, 17) = 6.08$, $p < .05$ in TOS).

The interaction of Perceived Distance and Region was not significant for pFs with each of the scene-selective areas PPA and TOS.

3.1.3. Discussion of Experiment 1 (full screen Ponzo)

The results of Experiment 1 are consistent with the perceived distance hypothesis for the LO, TOS and PPA but not for the pFs. We found that both scene-selective regions (PPA and TOS) were biased toward perceived distal information, whereas the LO object-selective region was biased toward perceived proximal information for both objects and buildings. These findings support our hypothesis that category-selective regions are sensitive to perceived distance information for both preferred and non-preferred stimuli.

In the Ponzo illusion, perceived distal stimuli occupy the upper part of the screen, whereas perceived proximal stimuli occupy the lower part of the screen. Recent findings suggest a similar location bias in the same category-selective regions (e.g., Schwarzlose et al., 2008). It is noteworthy that Schwarzlose and colleagues showed this location bias for stimuli presented in the periphery while participants were fixating at the center of the screen, whereas our data suggest that a similar effect may be found also when participants are fixating on the stimuli that are presented at the upper and lower parts of the screen. The goal of Experiment 2 was to determine whether the effect that we found in Experiment 1 simply reflects a location bias (i.e., bias toward upper versus lower stimuli) or can be attributed to a perceived distance bias. In order to test this question, we presented the object/building stimuli in a fixed location at the center of the screen, with the Ponzo lines being presented in either the upper half or the lower half of the screen. This way, a centrally presented object would appear in a perceived proximal location when the Ponzo lines are located in the upper half of the screen, and in a perceived distal location when the Ponzo lines are located in the lower half of the screen. If the perceived distance bias is independent of location, we should see a stronger response in scene-selective regions when the stimulus is presented with lower Ponzo lines and consequently perceived to be distal, than when presented with upper Ponzo lines and therefore perceived more proximal. In contrast, in the object region LO, we expect a stronger response when the stimulus is presented with upper Ponzo lines (proximal) than lower Ponzo lines (distal).

Notably, the design of Experiment 2 may generate a competition between the location of the stimulus and the Ponzo display. Specifically, although the location of the stimulus is fixed, perceived distal stimuli (preferred by scene regions) are associated with a visual display (the Ponzo lines) that occupies the lower part of the screen (preferred by object regions), and perceived proximal stimuli (preferred by object regions) are associated with a visual display that mostly occupies the upper part of the screen (preferred by scene regions). Thus, the effect of location of the display is in the opposite direction of the effect of perceived distance of the stimulus. If the location of the Ponzo lines has a stronger effect on activation than does the object-perceived distance, then we would expect the PPA and TOS to respond more strongly to the upper Ponzo than to the lower Ponzo, and the LO to respond more strongly to the lower than to the upper Ponzo. However, if the stimulus-perceived distance has a stronger effect than the location of the Ponzo lines, then we would expect the PPA and TOS to respond more strongly to the lower Ponzo (stimulus is perceived as distal), and the LO to respond more strongly to the upper Ponzo (stimulus is perceived as proximal). We first examined whether the half screen Ponzo line display indeed generates a distance illusion in a psychophysical experiment. We then assessed the response of scene-selective and object-selective regions across the two conditions.

3.2. Experiment 2: half screen Ponzo display

3.2.1. Psychophysical experiment

The stimulus display and results are presented in Fig. 5A–E. As in the previous psychophysical experiment, we compared only the proportion of responses in which the upper stimulus was larger with the proportion of responses that the lower stimulus was larger. We conducted the analysis separately for the upper screen and lower screen Ponzo display conditions.

3.2.1.1. Ponzo lines located in the lower half of the screen. We conducted a 5 (size of the non-standard stimulus: 2% smaller than the standard, 1% smaller than the standard, equal, 1% larger than the standard, 2% larger than the standard) × 2 (stimulus: buildings, objects) × 2 (response: upper stimulus is larger, lower stimulus is larger) × 2 (standard stimulus located at the upper versus lower end of the lines) repeated measures analysis of variance. There was no main or interaction effect for the location of the standard stimulus. There was a main effect for size $F(4, 36) = 14.34$, $p < .0001$, such that the larger the non-standard stimulus was, the less likely subjects were to respond ‘‘equal’’ (compared to ‘‘upper is larger’’ and ‘‘lower is larger’’). There was a main effect for stimulus $F(1, 9) = 14.95$, $p < .01$, such that there were less ‘‘equal’’ responses for buildings than for objects. There was also a main effect for response, $F(1, 9) = 34.7$, $p < .0001$, such that subjects were more likely to respond ‘‘upper is larger’’ than ‘‘lower is larger’’. In addition, there was an interaction between stimulus and size $F(4, 36) = 3.1$, $p < .05$, such that there was a steeper reduction in the ‘‘equal’’ responses for buildings than for objects the larger the upper stimulus was. Further, there was an interaction between size and response, $F(4, 36) = 23.25$, $p < .001$, such that the larger the upper stimulus was, the bigger was the difference between the responses ‘‘upper is larger’’ and ‘‘lower is larger’’. Finally, there was an interaction between stimulus and response $F(1, 9) = 13.8$, $p < .05$, such that the difference between the responses ‘‘upper larger’’ and ‘‘lower larger’’ was larger for buildings than for objects. Due to this interaction we looked separately at each response within each size condition, for buildings and objects. The results show that for both buildings and objects, subjects responded more ‘‘upper is larger’’ than ‘‘lower is larger’’ for all size conditions $p < .05$.

3.2.1.2. Ponzo lines located in the upper half of the screen. We conducted an identical ANOVA as we did for the ‘‘lower half’’ Ponzo. There was a main effect for size $F(4, 32) = 8.2$, $p < .001$, such that the larger the upper stimulus was, the less likely subjects were to respond ‘‘equal’’ (compared to ‘‘upper is larger’’ and ‘‘lower is larger’’). There was also a main effect for response, $F(1, 9) = 23.7$, $p < .0001$, such that subjects were more likely to respond, ‘‘upper is larger’’ than ‘‘lower is larger’’. In addition, there was an interaction between size and response, $F(4, 32) = 17.5$, $p < .001$, such that the larger the upper stimulus was, the bigger was the difference between the responses ‘‘upper is larger’’ and ‘‘lower is larger’’. Finally, there was an interaction between stimulus and response $F(1, 8) = 84.2$, $p < .0001$, such that the difference between the responses ‘‘upper larger’’ and ‘‘lower larger’’ was greater for buildings than for objects. Due to this interaction we looked separately at the simple effects of response within each size condition, for buildings and objects. The results show that for buildings, the upper stimulus was always judged as larger than the lower stimulus, $p < .05$. However, for objects, the upper stimulus was judged as larger only when it was *actually larger* than the lower stimulus in 1% and in 2% of the size of the lower stimulus. But when the upper stimulus was equal in size to the lower stimulus, or smaller in 1% or 2%, there was no significant difference in the judgments, $p > .05$. Thus, this result suggests that there was *no distance illusion at all for ob-*

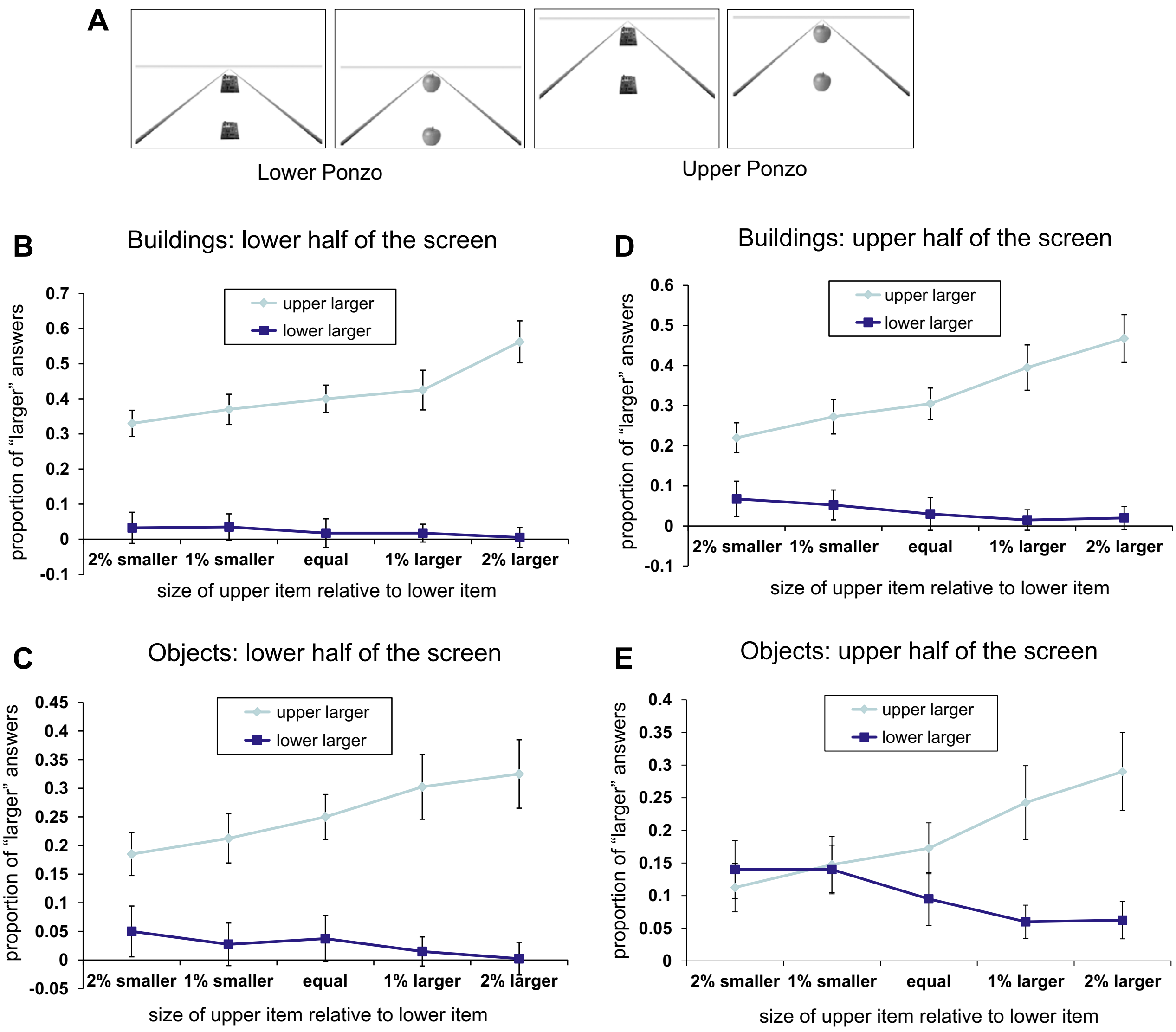

Fig. 5. A. The psychophysical experiment included two objects that were presented simultaneously or two buildings that were presented simultaneously in the upper or lower ends of two converging lines (the Ponzo illusion). The lines were located at the upper half or the lower half of the screen. Participants were asked to indicate which of the two stimuli seemed larger: the upper, the lower, or whether they seem equal. B and C. (Lower Ponzo). The proportion of "larger" answers as a function of size of the upper stimulus relative to the lower stimulus, when the Ponzo lines were located at the lower half of the screen. For both buildings (B) and objects (C), the proportion of upper stimuli that were perceived as larger was greater than the proportion of lower stimuli that were perceived as larger. D and E. (Upper Ponzo). The proportion of "larger" answers as a function of size of the upper stimulus relative to the lower stimulus. For the upper Ponzo, for buildings (D) the proportion of upper stimuli that were perceived as larger was greater than the proportion of lower stimuli that were perceived as larger. However, for objects (E) there was no difference in judgment of the upper and lower stimuli as "larger" when the upper stimulus was actually smaller or equal to the lower stimulus. Upper stimulus was judged as "larger" only when it was actually larger than the lower stimulus.

jects, because the upper objects were judged as larger only when they were actually larger than the lower ones.

3.2.2. Discussion

The results of the psychophysical experiment show that the distance illusion for half-screen Ponzo displays is effective for buildings but less effective for objects, in particular when the lines are presented in the upper part of the screen. The lack of perceived distance effect for object stimuli is consistent with previous research that suggests that the location of the horizon line in a picture has an influence on the evaluation of relative size of objects (Rogers, 1996). Specifically, Rogers found that the most accurate size evaluations of objects were given when the horizon line was placed two thirds of the way up from the bottom of the picture. It seems that the horizon line creates a better distance illusion when it is at about the same level as a subject's eyes than when it is significantly above or below eye-level. Notably, for buildings, both the upper and the lower Ponzo displays effectively generate a distance illusion. Thus, the lack of effect for the upper Ponzo for objects does not suggest an inherent problem with the distance illusion of the lower Ponzo display. Because the illusion is effective for buildings for both upper and lower Ponzo lines, we used the half-screen Ponzo display in an fMRI study and examined the fMRI data separately for buildings and objects.

3.2.3. fMRI experiment

Because the psychophysical results show better illusion for building than object stimuli we report fMRI data separately for building and object stimuli.

3.2.3.1. Building stimuli. Results for buildings were consistent with our perceived distance hypothesis. The stimulus display and the results are presented in Fig. 6A–C. In the LO, responses were higher to perceived proximal than to perceived distal stimuli (*Ms* = 1.11

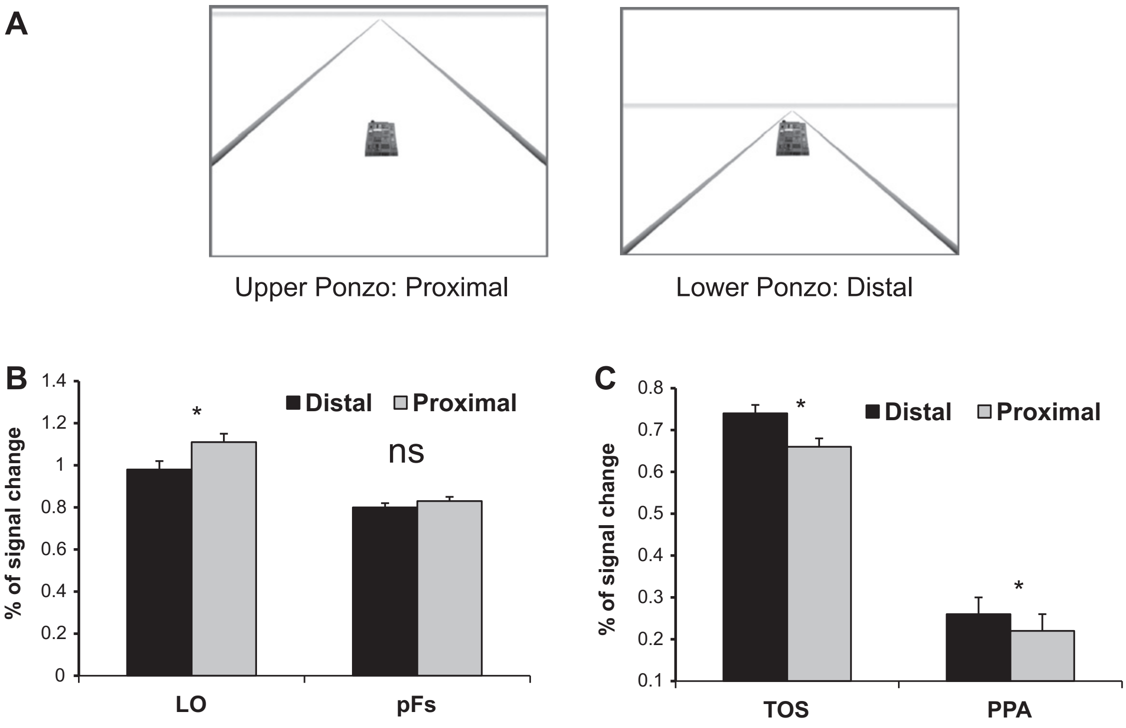

Fig. 6. A. In the fMRI experiment building stimuli were presented at the center of the screen. The two converging lines (the Ponzo illusion) were located either at the upper half (i.e. stimuli are proximal) or the lower half of the screen (i.e. stimuli are distal). Participants fixated at the stimuli and performed a 1-back task. B. The lateral occipital (LO) object selective regions showed a higher response to stimuli presented in the proximal (upper Ponzo) than distal (lower Ponzo). The posterior fusiform (pFs) object area showed no distance bias. C. Scene selective regions (PPA and TOS) showed a higher response to stimuli presented in the distal (lower Ponzo) than proximal (upper Ponzo) location.

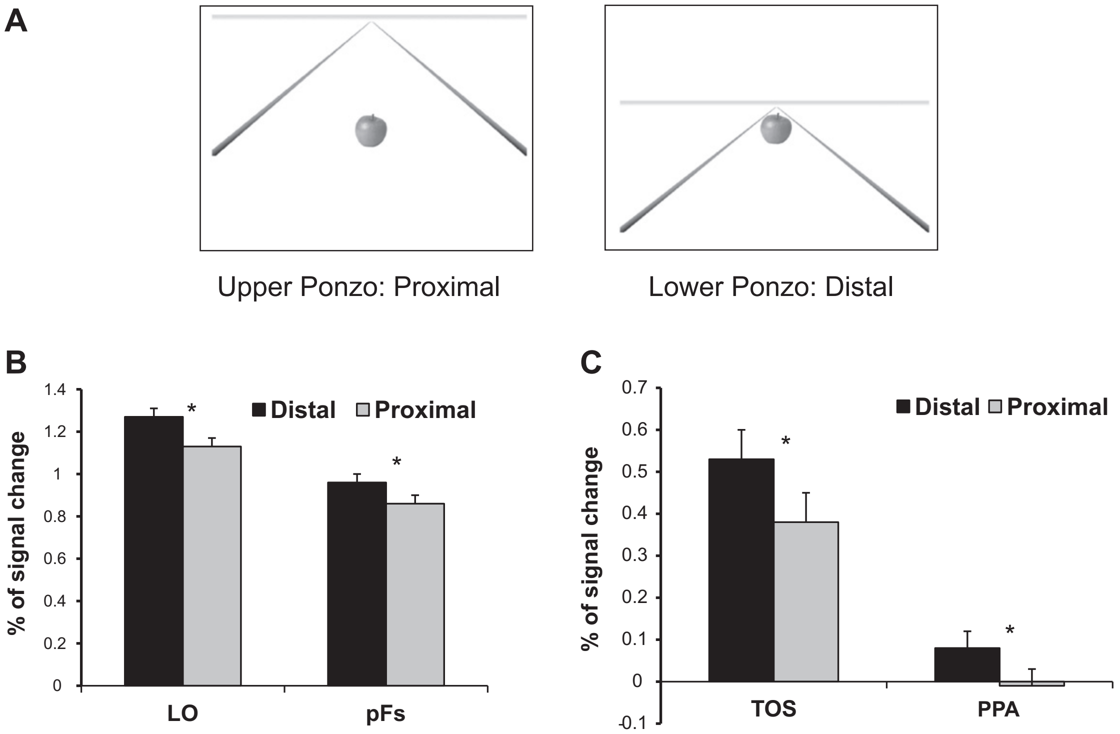

Fig. 7. A. In the fMRI experiment object stimuli were presented at the center of the screen. The two converging lines (the Ponzo illusion) were located either at the upper half (i.e. stimuli are proximal) or the lower half of the screen (i.e. stimuli are distal). Participants fixated at the stimuli and performed a 1-back task. B. The lateral occipital (LO) and posterior fusiform (pFs) object selective regions also showed a higher response to stimuli presented in the distal (lower Ponzo) than proximal (upper Ponzo) which is consistent with the lack of distance illusion for proximal objects presented within upper Ponzo lines. C. Scene selective regions (PPA and TOS) showed a higher response to stimuli presented in the distal (lower Ponzo) than proximal (upper Ponzo).

and 0.98, respectively, $t(13) = 2.69$, $p < .01$). In contrast, in the PPA responses were higher to perceived distal than perceived proximal stimuli (*Ms* = 0.26 and 0.22, respectively, $t(11) = 1.98$, $p < .05$). Similarly to the PPA, in the TOS responses were higher to perceived distal than perceived proximal stimuli (*Ms* = 0.74 and 0.66, respectively, $t(11) = 1.95$, $p < .05$). Finally, in pFs responses were similar to perceived proximal and perceived distal stimuli (*Ms* = 0.9 and 0.89, respectively, $t(13) = 1$, $p = .16$).

3.2.3.1.1. LO versus PPA. Consistently with the distance hypothesis, analysis of the 12 participants who had activations in both the LO and the PPA showed a significant interaction between Region and Distance, $F(1, 11) = 13.23$, $p < .01$.

3.2.3.1.2. LO versus TOS. Consistently with the distance hypothesis, analysis of the 12 participants who had activations in both the LO and the TOS showed a significant interaction between Region and Distance, $F(1, 11) = 20.81$, $p < .01$.

3.2.3.1.3. pFs versus PPA. Analysis of 12 participants who had activations in both regions showed marginal interaction between Region and Distance, $F(1, 11) = 3.95$, $p = .072$.

3.2.3.1.4. pFs versus TOS. Analysis of 12 participants who had activations in both regions showed marginal interaction between Region and Distance, $F(1, 11) = 4$, $p = .069$.

3.2.3.2. Object stimuli. The stimulus display and the results are presented in Fig. 7A–C. Consistent with the hypothesis which suggests that scene regions would be biased toward perceived distal stimuli, in the PPA responses were higher to perceived distal than perceived proximal stimuli (*Ms* = .08 and −.01, respectively, $t(11) = 2.2$, $p < .05$). Similarly to the PPA, in the TOS responses were higher to perceived distal than perceived proximal stimuli (*Ms* = 0.51 and 0.35, respectively, $t(11) = 2.4$, $p < .05$). In contrast to our hypothesis, which suggests that object regions would be biased toward perceived proximal stimuli, in the LO responses were higher to perceived distal than perceived proximal objects, (Ms = 1.27 and 1.13, respectively, $t(11) = 3.2$, $p < .01$.) Finally, in pFs responses were higher for perceived distal than perceived proximal stimuli (*Ms* = 0.96 and 0.86, respectively, $t(13) = 2.2$, $p < .05$).

3.2.3.2.1. LO versus PPA. Analysis of 12 participants who showed activations of both regions showed that the interaction between Region and Distance was not significant, $F < 1$. Responses of both regions were higher for perceived distal than to perceived proximal stimuli $F(1, 11) = 6.29$, $p < .05$.

3.2.3.2.2. LO versus TOS. Analysis of 12 participants who showed activations of both regions showed that the interaction between Region and Distance was not significant, $F < 1$. Responses of both regions were higher for perceived distal than to perceived proximal stimuli $F(1, 11) = 8.1$, $p < .05$.

3.2.3.2.3. pFs versus PPA and TOS. Similar to LO, the interaction of Region and Distance was not significant for pFs with each of the scene-selective areas (PPA and TOS). For both analyses of pFs and PPA, and pFs and TOS, there was a main effect for distance (PPA: $F(1, 11) = 11.2$, $p < .01$, TOS $F(1, 13) = 7.8$, $p < .05$):, such that activation was stronger for perceived distal than perceived proximal targets .

3.2.4. Discussion of Experiment 2 (half screen Ponzo)

In the second fMRI experiment we found that for buildings there was a bias toward perceived distal stimuli in the scene-selective regions and a bias toward perceived proximal stimuli in the objects regions. Thus, the effect of the building is consistent with the perceived distance hypothesis. For objects, results were similar to those for buildings and consistent with our hypothesis (i.e., stronger activation for distal than for proximal stimuli) in scene areas. This effect is in line with the effectiveness of the lower Ponzo lines in generating a distance illusion for objects that we found in the psychophysical experiment. However, the bias toward perceived proximal stimuli with upper Ponzo lines was not found for objects in LO. Notably, these findings are in line with the results of the psychophysical experiment that tested the effectiveness of the distance illusion and revealed that the half-screen Ponzo manipulation was not effective for objects when the Ponzo lines appeared at the upper part of the screen (i.e. objects are proximal). The pattern of activation for objects in the LO in Experiment 2 might therefore be explained by the lack of distance illusion for proximal objects presented with upper Ponzo. This lack of distance illusion for object stimuli possibly facilitated the effect of location of the Ponzo display itself. That is, there is a competition between the distance effect (which supposedly biases the preference toward lower stimuli that are perceived as proximal in the LO) and the location effect (which supposedly biases the preference toward upper stimuli in the LO). Since there is a competition between distance and location, it is possible that the lack of the former facilitates the latter, resulting in a preference for lower Ponzo over upper Ponzo display, rather then for perceived proximal than perceived distal objects in LO. Further research is needed with a more effective illusion that could generate a reliable perceived distance effect for objects in fixed locations.

3.3. Analysis of behavioral performance

For both fMRI experiments (full-screen Ponzo and half-screen Ponzo), we found that performance was at ceiling, such that 99% of the responses were correct in Experiment 1, and 98% of the responses were correct in Experiment 2. To test whether performance may have varied with Stimulus Location (Proximal, Distal) or Category (Objects, Buildings), we ran repeated-measures ANOVA with Location and Category as factors, and accuracy as the dependent measure. There were neither significant effects of Stimulus Location (Proximal, Distal) nor Stimulus Category (Objects, Buildings), nor an interaction between the two (all three *P* values >0.50) on accuracy.

3.4. Fixation check: analysis of eye position

To insure that stimuli in our experiments were indeed presented in the fovea, we also conducted an eye tracking experiment outside the scanner and asked participants to perform the same 1-back task that they performed in the two fMRI experiments. We measured both behavioral performance on the 1-back task and the proportion of time participants fixated on the stimulus relative to other locations outside the stimulus across the different conditions.

The overall percent of looking at the stimuli versus outside the stimuli was 91.5% in the first experiment (full screen Ponzo), and 96% in the second experiment (half screen Ponzo). To test whether eye position varied with stimulus location (proximal, distal) or category (objects, buildings), we ran repeated-measures ANOVA with Location and Category as factors, and the proportion of looking at the stimulus versus other areas as the dependent measure. In Experiment 1, participants looked at the perceived proximal buildings for 96% (SD = 2%) of the time, and perceived distal buildings for 96% (SD = 2%) of the time. In addition, they looked at the perceived proximal objects for 95% (SD = 2%) of the time and perceived distal objects for 96% (SD = 2%) of the time. In Experiment 2, participants looked at perceived proximal building for 91% (SD = 11%) of the time and perceived distal buildings for 88% (SD = 12%) of the time. In addition, they looked at perceived proximal objects for 95% (SD = 3%) of the time, and perceived distal objects for 92% (SD = 9%) of the time. For both Experiment 1 and Experiment 2, we found neither significant effects of stimulus location (proximal, distal) or stimulus category (objects, buildings), nor an interaction between the two (all three *P* values >0.5).

4. General discussion

The current study examined the hypothesis that the ventral visual cortex plays a role in representing perceived distance information. Our findings provide support for this hypothesis. Experiment 1 showed that the scene-selective regions, PPA and TOS, are biased toward perceived distal information, whereas the object-selective region, LO, is biased toward perceived proximal information. Experiment 2 further showed that for building stimuli this bias was independent of the location of the stimulus. These findings are consistent with recent studies, which show that location information is represented in the ventral visual cortex (Schwarzlose et al., 2008), and suggest that these location biases may reflect (partially or fully) the representation of the perceived distance of the stimulus from the observer.

Schwarzlose et al. (2008) reported that scene-selective regions are biased toward stimuli presented in the upper visual field, whereas object-selective regions are biased toward stimuli presented in the lower visual field. As discussed earlier, the vertical location of a stimulus may function as a depth cue (e.g., Bruno & Cutting, 1988; Cutting & Vishton, 1995; Epstein, 1966; Gibson, 1950, p. 180; Yonas, Elleiff, & Arterberry, 2002). In fact, Schwarzlose et al. (2008) suggested that the location bias of object category might be the result of experience, such that the location biases of different categories reflect the locations where these stimuli typically appear in daily life. Since proximal items tend to appear in lower locations and distal items in upper locations (e.g., Bruno & Cutting, 1988; Cutting & Vishton, 1995; Epstein, 1966; Gibson, 1950, p. 180; Yonas, Elleiff, & Arterberry, 2002), the location effect found by Schwarzlose et al. (2008) may be consistent with a perceived distance effect. Indeed, Experiment 2 tested the effect of distance while holding the location fixed, by presenting the stimulus at the center of the screen while changing the location of the Ponzo lines. Although the distance illusion with this display was weaker for objects, it was still effective for buildings. Consistent with the distance hypothesis, we found a bias for distal buildings in scene-selective regions and for proximal buildings in object-selective region. Thus, the current findings go beyond the data reported by Schwarzlose and colleagues in that it reveals an egocentric distance bias that is independent of the stimulus location.

Furthermore, our findings are consistent with recent findings, which suggest that scene regions (PPA and TOS) are biased toward stimuli presented in the periphery whereas the object region (LOC) is biased toward stimuli presented at the fovea (Hasson et al., 2002; Levy et al., 2001). Foveal vision is typically associated with recognition of proximal stimuli whereas peripheral vision may be associated with recognition of distal stimuli. More broadly, our findings suggest that these previous finding that revealed effects of location in a two-dimensional frontoparallel plane (upper versus lower, e.g., Levy et al., 2001; Schwarzlose et al., 2008) may reflect the representation of a third dimension in the ventral visual stream. Finally, our findings are consistent with a recent report, which suggest that activity in high- level visual areas (such as the PPA, and LOC) reflects the perceived location of the stimulus more than its physical location (Fischer, Spotswood, & Whitney, 2009). Experiment 2 demonstrated this point by showing that the activation in the PPA, TOS and LO reflected the perceived locations of the stimulus (near, far), despite the fact that the physical location of the stimulus remained constant across all conditions.

4.1. Size versus distance

Our findings are also consistent with Cate, Goodale and Kohler's (2011) research, which tested the representation of perceived size in object and scene regions. Cate et al. (2011) used the Ponzo illusion in order to generate a size illusion. Specifically, subjects were presented with stimuli of cameras and garages that appeared either in proximal location within the Ponzo illusion and thus appeared small, or in a distal location and thus appeared large. It was found that scene regions (PPA and TOS) were biased toward the large (distal) targets (both cameras and garages), whereas the object region (LO) was biased toward small (proximal) garages but not cameras. This finding is consistent with ours, since apparent size and distance are inherent to each other.

However, while Cate et al. (2011) provided evidence that converges with our findings, their study did not include additional measures – which are included in our study – to test for alternative explanations for their results. First, Cate et al. (2011) did not conduct a behavioral manipulation check in order to confirm the effectiveness of their perceived distance manipulation. Thus, it is hard to know whether their visual display indeed generated a distance/size bias. Notably, the lack of the predicted effect in the LO for objects they reported (Cate et al., 2011) could be explained by the fact the manipulation was not effective for proximal objects, as was shown in our study. In addition, in Cate et al. (2011) study, eye movements were not monitored. Since in their perceived distal/large scene there are two focal objects (i.e., a hand which is presented separately from an item), while in the perceived proximal/small scene there is only a single focal object (the hand holds the item), it is possible that in the perceived distal/large condition people focused more on the periphery than in the perceived near/proximal condition, whereas in the perceived proximal/small condition they focused only at the center. In other words, the perceived distance/size effect might have been confounded with a center/periphery effect. In contrast, in the current study we controlled for eye movements and demonstrated that the participants fixated on the target stimuli under all conditions and to the same extent, thus showing that center/periphery effect could not account for the results. Finally and more broadly, the theoretical account and motivation for the study in Cate et al. (2011) was to explore perceived size, rather than perceived distance. Indeed, as we noted earlier, perceived distance and perceived size are inherently related. Specifically, when presented with two items identical in size, the distal item is perceived to be larger than proximal item (e.g., Murray et al., 2006; Smith, 1958).

Could sensitivity to perceived size in itself account for the results? Cate et al. (2011) do not provide data to support this claim. However, another recent study tested the representation of size information in scene and objects regions (Konkle & Oliva, 2010). In this study, a region in the parahippocampal gyrus was preferentially active to big (e.g., a car) relative to small (e.g., a strawberry) objects, whereas the object area, LOC was preferentially active to small relative to big objects. Importantly, these regions were tolerant to the physical size but sensitive to the real world size. Thus, these findings are inconsistent with Cate et al. (2011) conclusion and suggest that size by itself may not account for the effect of perceived distance. Furthermore, given the strong relationship between size and distance, the real-world size effects that were found by Konkle and Oliva (2010) may be also explained in terms of the perceived distance of the stimulus from the observer. That is, real-life, large objects (such as a house) might signal that the item is distal, and therefore activate the scene regions, whereas small objects (such as a strawberry) might signal that the item is close and therefore activate the object region. Finally, note that contrary to the rationale presented in Cate et al. (2011) and Konkle & Oliva's (2010), the PPA and TOS are not selective only to large stimuli like buildings but also to open landscapes such as beaches or deserts (e.g., Epstein & Kanwisher, 1998). Size is not an inherent attribute of scene stimuli. Rather, it is an inherent attribute of buildings, which are one sub-category of the preferred stimuli of these scene-selective areas. In contrast, distance is common to all types of preferred stimuli of the TOS and PPA, including the bias for large real life stimuli. This idea is also consistent with the role of the scene regions in navigation (Epstein, 2008), and with the idea that in order to interact with an object (e.g., touch it) one needs first and foremost to know whether the object is located in a reachable distance (i.e., proximal) or not (i.e., distal). Thus, a major contribution of the current study is to provide "theoretical glue" that integrates the findings about center/periphery and object categories, using the concept of distance.

4.2. Distance as a binary versus parametric variable

An interesting question is whether the effect of perceived distance is binary or parametric. In the current study perceived distance was manipulated as a binary variable, with spatially "near" and "far" conditions. However, by no means do we claim that distance is a binary dimension. In fact, it is more likely that representation of distance in object and scene brain regions is continuous. Thus, we would expect that the response of the PPA increases linearly for more perceived distal stimuli, whereas the response of the LO will increases linearly for more perceived proximal stimuli. Future studies may apply a parametric design to assess this hypothesis.

4.3. Why is perceived distance represented in the ventral visual stream?

Why is perceived distance represented in the ventral visual stream, which is devoted to the process of object recognition? Why does a region that prefers buildings and mountains show a bias towards distal information? Why does a region that prefers apples and pots (and other small tangible entities) show a bias towards proximal information? In real life, buildings (and other scenes, like mountains) typically appear at a distance, whereas apples and pots (and other objects) typically appear in proximity. It is possible, then, that the representation of distance information in the ventral visual stream assists the process of object identification. In other words, part of knowing "what" is knowing "where." Perceived distance serves as a first gross discrimination that eventually leads to the final identification of the target.

The following examples reflect this logic. The identification of scenes ("Is this building my house or my friend's house?"; "Is this mountain Mount Fuji?") requires one to stand and observe from a distance, otherwise one is unable to see what it is. Taking into account our distance from the target precedes identifying the exact identity of the house as yours, or of the mountain as Mount Fuji. Similarly, the identification of objects ("Do these sunglasses belong to my wife or to my daughter?"; "Is the item at the top of the fridge mushrooms or strawberries?") requires standing in proximity to it, otherwise one is unable to see what it is. Taking into account our proximity from the object precedes identifying its exact identity.

4.4. Distance as an organizing principle in the brain

The idea that distance is related to representation of content is consistent with a social-cognitive theory: Construal level theory (CLT; Liberman & Trope, 2008; Trope & Liberman, 2010). According to CLT, perceived distance (whether spatial, temporal, or social) changes the way people represent objects and events (Amit, Algom, & Trope, 2009). People represent distal events in an abstract (schematic and global) manner and proximal events in a concrete (detailed and specific) manner. For example, it has been shown that temporal distance facilitates identifying global patterns, whereas temporal proximity facilitates identifying constituent elements (Forster, Liberman, & Friedman, 2004; see also Wakslak, Trope, Liberman, & Alony, 2006). Because objects are usually specific parts of more global scenes, CLT would predict that scenes would be associated with perceived distality and objects with perceived proximity. Thus, similarly to the current study, CLT suggests that perceived distance is a powerful factor that shapes the way people represent and process information. Indeed, recent cognitive neuroscience research provides converging evidence for this argument with respect to various distance dimensions (e.g., spatial, temporal, and social; Addis, Wong, & Schacter, 2008; Mitchell, Macrae, & Banaji, 2006; Mobbs et al., 2007).

4.5. The pFs object area and distance representation

Despite the fact that the object-selective posterior fusiform area, pFs, shows a clear preference for objects, it does not show a preference for stimuli presented in the lower visual field, as does the LO object-selective region (e.g., Schwarzlose et al., 2008). Consistent with Schwarzlose et al. (2008), the current study showed no clear preference for perceived proximal or distal information in this region. Therefore, we conclude that this region shows no perceived distance bias. The differences between the LO and the pFs in the perceived distance bias raise interesting questions regarding possible differences in the properties and functions of LO and pFs, and the function of perceived distance in each region in supporting identification. It is possible, for example, that the LO is more dependent on distance cues in order to identify a stimulus than is the pFs. Further research is needed in order to test this prediction.

5. General conclusion

The current study shows that the ventral visual stream plays a role in representing distance information, such that scene-selective regions (PPA and TOS) are biased toward perceived distal information whereas the object-selective region (LO) is biased toward perceived proximal information. These findings extend previous reports on the representation of location in the ventral visual stream and suggest that previously reported retinotopic biases might reflect the representation of perceived distance information. The representation of perceived distance information in cortical regions that are devoted to object recognition may reflect a strong link between the ability to identify an object and its perceived distance from the observer, such that each object category is associated with the distance that is ideal for its successful identification.

Acknowledgments

We would like to thank Yonantan Douek and Vadim Axelrod for valuable help with fMRI data analysis, Peter Huy Wong for assisting with data collection and stimulus preparation of the psychophysical studies, Jonathan Oron for help with collection of eye tracking data, and Tejaswinhi Srinivas for proofreading. We also like to thank Russell Epstein for valuable comments. This work was funded by a grant from the Adams Super Center for Brain Research to G.Y.

References

Addis, D. R., Wong, A. T., & Schacter, D. L. (2008). Constructive episodic simulation: temporal distance and detail of past and future events modulate hippocampal engagement. *Hipocampus, 18*, 227–237.

Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1998). An area within human ventral cortex sensitive to "building" stimuli: Evidence and implications. *Neuron, 21*, 373–383.

Amit, E., Algom, D., & Trope, Y. (2009). Distance-dependent processing of pictures and words. *Journal of Experimental Psychology: General, 138*, 400–415.

Brainard, D. H. (1997). Psychophysics software for use with MATLAB. *Spatial Vision, 10*, 433–436.

Bruno, N., & Cutting, J. E. (1988). Minimodularity and the perception of layout. *Journal Experimental Psychology: General, 117*, 161–170.

Cate, A. D., Goodale, M. A., Kohler, S., 2011. The role of apparent size in building and object specific regions of ventral visual cortex. To appear in: *Brain Research*.

Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M. A., et al. (2000). The visual word form area. *Brain, 123*, 291–307.

Cutting, J. E., & Vishton, P. M. (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In W. Epstein & S. Rogers (Eds.), *Handbook of perception and cognition, 5; perception of space and motion* (pp. 69–117). San Diego, CV: Academic Press.

Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science, 293*, 2470–2473.

Epstein, W. (1966). Perceived depth as a function of relative height under three background conditions. *Journal of Experimental Psychology, 72*, 335–338.

Epstein, R. A. (2005). The cortical basis of visual scene processing. *Visual Cognition, 12*, 954–978.

Epstein, R. A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends in Cognitive Science, 12*, 388–396.

Epstein, R., Harris, A., Stanley, D., & Kanwisher, N. (1999). The parahippocampal place area: Recognition, navigation, or encoding? *Neuron, 23*, 115–125.

Epstein, R. A., Higgins, J. S., Jablonski, K., & Feiler, A. M. (2007). Visual scene processing in familiar and unfamiliar environments. *Journal of Neurophysiology, 97*, 3670–3683.

Epstein, R. A., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature, 392*, 598–601.

Fischer, J., Spotswood, N., & Whitney, D. (2009). The emergence of perceived position in the visual system. *Journal of Cognitive Neuroscience*, 1–18.

Forster, J., Liberman, N., & Friedman, R. S. (2004). Temporal construal effects on abstract and concrete thinking: Consequences for insight and creative cognition. *Journal of Personality and Social Psychology, 87*, 177–189.

Gibson, J. J. (1950). *The perception of the visual world*. Oxford, England: Houghton Mifflin.

Goodale, M. A., Milner, A. D., Jakobson, L. S., & Carey, D. P. (1991). A neurological dissociation between perceiving objects and grasping them. *Nature, 349*, 154–156.

Habib, M., & Sirigu, A. (1987). Pure topographical disorientation: A definition and anatomical basis. *Cortex, 23*, 73–85.

Hasson, U., Levy, I., Behrmann, M., Hendler, T., & Malach, R. (2002). Eccentricity bias as an organizing principle for human high-order object areas. *Neuron, 34*, 479–490.

Kanwisher, N., Chun, M., McDermott, J., & Ledden, P. (1996). Functional imaging of human visual recognition. *Cognitive Brain Research, 5*, 55–67.

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: A cortical region specialized for the perception of faces. *Philosophical Transactions, 361*, 2109–2128.

Konkle, T., & Oliva, A. (2010). Examining how the real-world size of objects is represented in ventral visual cortex. *Visual Science Society Meeting*. Naples, Florida.

Levy, I., Hasson, U., Avidan, G., Hendler, T., & Malach, R. (2001). Center-periphery organization of human object areas. *Nature Neuroscience, 4*, 533–539.

Liberman, N., & Trope, Y. (2008). The psychology of transcending the here and now. *Science, 322*, 1201–1205.

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., et al. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of National Academy of Science, 92*, 8135–8139.

Mazer, J. A., & Gallant, J. L. (2000). Object recognition: Seeing us seeing shapes. *Current Biology, 10*, R668–R670.

Mishkin, M., Ungeleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences, 6*, 414–417.

McNeil, J. E., & Warrington, E. K. (1993). Prosopagnosia – a face-specific disorder. *Quarterly Journal of Experimental Psychology-A, 46*, 1–10.

Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron, 50*, 655–663.

Mobbs, D., Petrovic, P., Marchant, J. L., Hassabis, D., Weiskopf, N., Symour, B., et al. (2007). When fear is near: Threat imminence elicits prefrontal-periaqueductal gray shifts in humans. *Science, 317*, 1079–1083.

Moscovitch, M., Winocur, G., & Behrmann, M. (1997). What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. *Journal of Cognitive Neuroscience, 9*, 555–604.

Murray, S. O., Boyaci, H., & Kersten, D. (2006). The representation of perceived angular size in human primary visual cortex. *Nature Neuroscience, 9*, 429–434.

Op de Beeck, H. P., Haushofer, J., & Kanwisher, N. (2008). Interpreting fMRI data: Maps, modules and dimensions. *Nature Review Neuroscience, 9*, 123–135.

Rogers, S. (1996). The horizon-ratio relation as information for relative size in pictures. *Perception & Psychophysics, 58*, 142–152.

Saxe, R., Brett, M., & Kanwisher, N. (2006). Divide and conquer: A defense of functional. *Neuroimage, 30*, 1088–1096.

Schwarzlose, R. F., Swisher, J. D., Dang, S., & Kanwisher, N. (2008). The distribution of category and location information across object-selective regions in human visual cortex. *Proceedings of National Academy of Science USA, 105*, 4447–4452.

Smith, O. W. (1958). Judgments of size and distance in photographs. *American Journal of Psychology, 71*, 529–538.

Wakslak, C. J., Trope, Y., Liberman, N., & Alony, R. (2006). Seeing the forst when entry is unlikely: Probability and the mental representation of events. *Journal of Experimental Psychology: General, 135*, 641–653.

Trope, Y., & Liberman, N. (2010). Construal level theory of psychological distance. *Psychological Review, 117*, 440–463.

Yi Li, C., & Guo, K. (1995). Measurements of geometric illusions, illusory contours and stereo-depth at luminance and colour contrast. *Vision Research, 35*(12), 1713–1720.

Yonas, A., Elleiff, C. A., & Arterberry, M. E. (2002). Emergence of sensitivity to pictorial depth cues: Charting development in individual infants. *Infants Behavior & Development, 25*, 495–514.