

# Interaction envelope: Local spatial representations of objects at all scales in scene-selective regions



Wilma Alice Bainbridge<sup>a,\*</sup>, Aude Oliva<sup>b</sup>

<sup>a</sup> Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02139, USA

<sup>b</sup> Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02139, USA

## ARTICLE INFO

### Article history:

Received 28 April 2015

Accepted 25 July 2015

Available online 30 July 2015

### Keywords:

Interaction envelope

Scene-selective regions

Parahippocampal place area

Spatial object properties

fMRI

## ABSTRACT

While several cortical regions have been highlighted for their category selectivity (e.g., scene-selective regions like the parahippocampal place area, object selective regions like the lateral occipital complex), a growing trend in cognitive neuroscience has been to investigate what particular perceptual properties these regions calculate. Classical scene-selective regions have been particularly targeted in recent work as being sensitive to object size or other related properties. Here we test to which extent these regions are sensitive to spatial information of stimuli at any size. We introduce the spatial object property of “interaction envelope,” defined as the space through which a user transverses to interact with an object. In two functional magnetic resonance imaging experiments, we examined activity in a comprehensive set of perceptual regions of interest for when human participants viewed object images varying along the dimensions of interaction envelope and physical size. Importantly, we controlled for confounding perceptual and semantic object properties. We find that scene-selective regions are in fact sensitive to object interaction envelope for small, manipulable objects regardless of real-world size and task. Meanwhile, small-scale entity regions maintain selectivity to stimulus physical size. These results indicate that regions traditionally associated with scene processing may not be solely sensitive to larger object and scene information, but instead are calculating local spatial information of objects and scenes of all sizes.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

A simplified view of the human brain is that different modules exist for processing different image categories, such as objects and scenes. Object processing is the domain of the lateral occipital complex (LOC; Grill-Spector et al., 1999), while scene processing occurs in the parahippocampal place area (PPA; Epstein and Kanwisher, 1998), retrosplenial cortex (RSC; Auger et al., 2012), and occipital place area (OPA; Dilks et al., 2013). Indeed, various works have found evidence of dissociations between these two sets of regions, pointing towards separate processes (Mullin and Steeves, 2011; Wolfe et al., 2011; Ganaden et al., 2013).

However, a growing trend in perceptual neuroscience is in characterizing how these categorical lines are blurred, and what perceptual properties, beyond categories, they are sensitive to. For example, the LOC is found to be sensitive to the property of object shape, rather than to the category of “objects” itself (Kourtzi and Kanwisher, 2001; Kim et al., 2009), and recent work has found LOC sensitivity to scene-based information as well (Park et al., 2011). The story of scene-selective regions (e.g., PPA, OPA) is less clear—while it is well-accepted that they are essential for spatial navigation in scenes (Epstein, 2005), it is unclear

how this may extend to objects and other space-related object properties. Konkle and Oliva (2012) find that these regions in fact have increased activity for larger objects over smaller objects, with other work finding similar results in measures highly correlated with an object's physical size (Cate et al., 2011; Troiani et al., 2012). However, how can this selectivity to object size be consolidated with the traditional selectivity to scene information? These results could point to two distinct possibilities for the case of scene-selective regions—either they are sensitive to categorically larger entities (i.e., large objects and scenes rather than small, manipulable objects), or they are sensitive to a property of a stimulus (whether an object or scene) that is related to object size. The key may be that these regions are in fact sensitive to the navigable local space around a stimulus, including both small, manipulable objects as well as larger scenes. Indeed, PPA activity for large objects is correlated with the object's rated “spatial definition” (Mullally & Maguire, 2011). Additionally, the PPA has been found to hold information about objects relevant for navigation (Janzen and van Turennout, 2004) as well as the functional clutter in a scene, which determines actions and potential for navigation within the scene (Park et al., 2014). While these results do not address what information the PPA might encode for small, isolated, manipulable objects (if any), these point towards the possibility that the PPA is representing a spatial property of stimuli related to how one might spatially interact with the stimulus

\* Corresponding author at: MIT 32-D430, 32 Vassar St, Cambridge, MA 02139, USA.  
E-mail addresses: [wilma@mit.edu](mailto:wilma@mit.edu) (W.A. Bainbridge), [oliva@mit.edu](mailto:oliva@mit.edu) (A. Oliva).

(e.g., move through it, pick it up), rather than a categorical distinction based on the physical size of the object.

To test this hypothesis, the current study examines cortical selectivity to local spatial information of small-scale objects, separating it from pure selectivity to large entities. Importantly, as the PPA is sensitive to geometric calculations of large scenes (Epstein et al., 2003; Park and Chun, 2009), we propose that the spatial processing of small-scale objects would similarly be determined by a functionally-defined space, related to the space around which we interact with an object. We call this local spatial information the *interaction envelope* of the object, defined as the space through which a user transverses to interact with this object. We operationalize object interaction envelope as the number of hands most often used to interact with an object, as it is easily quantifiable and delineates two very different, non-overlapping volumes of interactive space. For example, a hamburger is a small object whose mode of interaction should involve its whole surface area using both hands, eliciting a larger interaction envelope than a similarly-sized coffee cup that we typically pick up by the handle with a single hand. We test to which extent regions involved in scene processing hold representations of interaction envelope, even for small-scale entities. Additionally, we propose an orthogonal experimental design that allows us to disentangle interaction envelope and physical size (Konkle and Oliva, 2012; Cate et al., 2011), while keeping other important object properties constant (Bar and Aminoff, 2003; Mullally and Maguire, 2011; Auger et al., 2012). This design ensures that any effects are not solely driven by a categorical sensitivity to large objects, or an alternate object property. We hypothesize that scene-selective regions will maintain representations of local spatial information of objects at all physical scales, while object-selective regions (LOC) will be equally sensitive to all objects. For regions selective for small entities (Konkle and Oliva, 2012), it is unclear whether they will be sensitive to physical size or interaction envelope, as they are not part of a traditional scene-processing network.

We conducted two functional magnetic resonance imaging (fMRI) studies to investigate sensitivity to interaction envelope in various regions of interest (ROIs selective to small-scale entities, large-scale entities, and control ROIs) using a crossed experimental design of small, manipulable object stimuli (Fig. 1) varying along two factors: physical size and interaction envelope. From this, we can determine if these ROIs do hold representations of small object local spatial information, rather than a solely categorical separation between large and small entities.

## 2. Materials and methods

### 2.1. Participants

Thirty right-handed adults (18 females, ages 18–38) participated in two functional magnetic resonance imaging (fMRI) studies, ( $N = 18$  for the main experiment and a separate set of  $N = 12$  for the control experiment). All participants had normal or corrected vision. In addition, 431 individuals from Amazon Mechanical Turk (AMT) participated in object norming studies. All observers, including both fMRI study and AMT study participants, consented to participation according to the guidelines of the MIT Committee on the Use of Humans as Experimental Subjects. Two participants were excluded from the analyses of the main experiment due to inability to localize at least 50% of the relevant functional regions of interest (ROIs).

### 2.2. Stimuli

The full stimulus set is composed of 160 images of large-scale objects with both big physical size and a large interaction envelope, and 160 images of small-scale objects (Fig. 1). Within the small-scale set are four conditions of manipulable objects (the orthogonal  $2 \times 2$  experimental design illustrated in Fig. 1) that disentangle object *physical* size from

*envelope* size, with 40 different objects per condition. Objects of physical size 1 and 2 significantly differ in their real-world physical size; the ground truth size for each object was determined by shipping dimensions from online marketplaces such as Amazon.com. Objects of envelope size 1 and 2 significantly differ in their interaction envelope space, operationalized by whether one or two hands are necessary to handle the object, and determined by an AMT experiment.<sup>1</sup> While these stimuli have been selected to be as closely orthogonalized as possible, conditions where object physical size and envelope size are highly correlated (objects of physical size 1 and envelope size 1, and well as objects of physical size 2 and envelope size 2) could potentially dominate effects found in analyses (as, indeed, their mean sizes are more separated than the other two conditions, see Table 1). Several analyses are thus also conducted on only the orthogonalized conditions (objects of physical size 1 and envelope size 2, and objects of physical size 2 and envelope size 1) to ensure that any results are truly due to an orthogonalization of the two object properties. The four conditions were not significantly different in visual statistics, such as average RGB color, vertical symmetry, amount of white space, luminance, and distribution of spatial frequency information (Rajimehr et al., 2011).<sup>2</sup> Images were all  $350 \times 350$  pixels in size, representing  $7 \text{ deg} \times 7 \text{ deg}$  of visual angle.

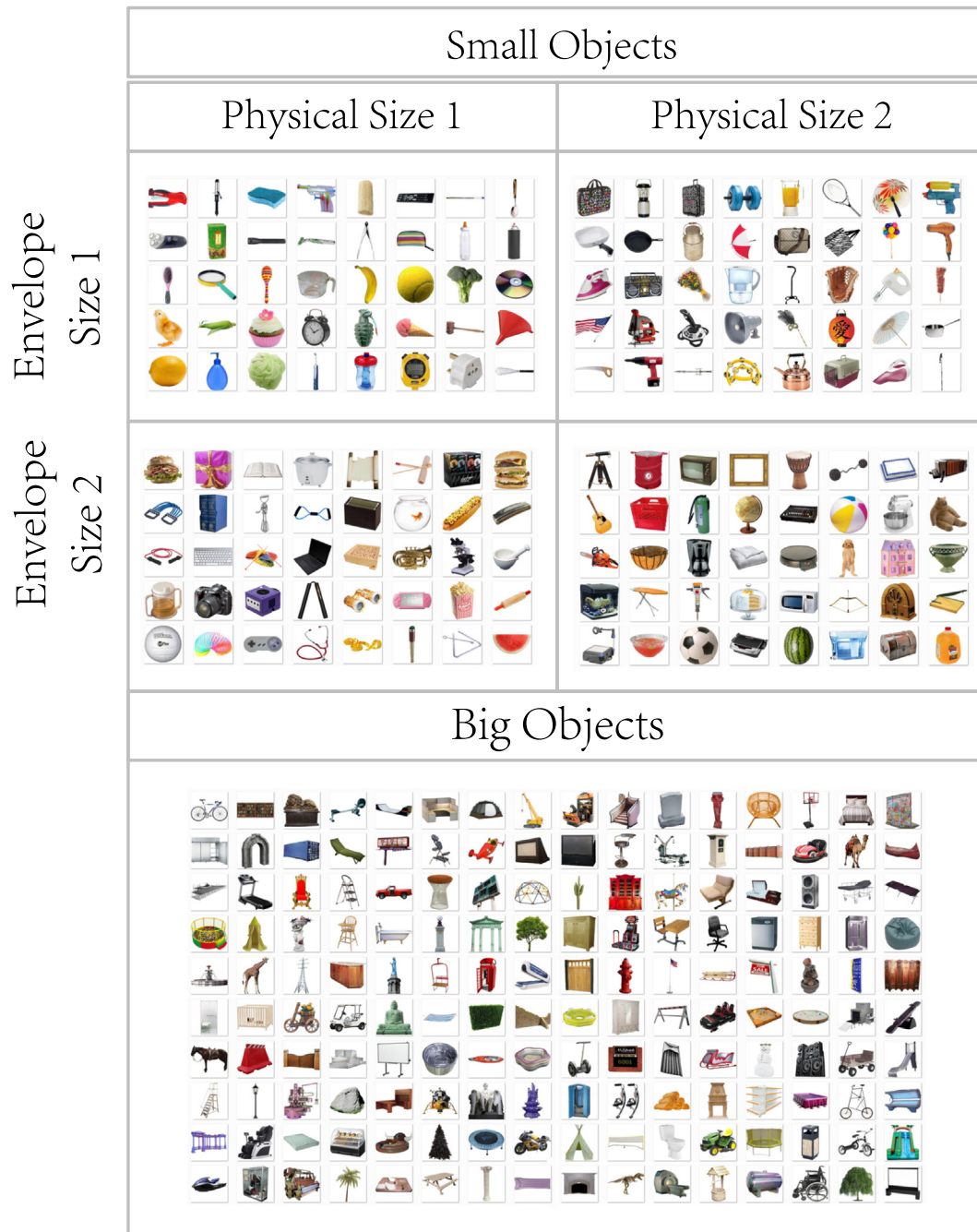
We also ran AMT studies ( $N = 291$ ) to evaluate which other object properties previously identified in modulating PPA activity correlate with our small-scale object sets selection (Table 1). Fifteen participants per object were asked to characterize the objects on various properties, following the methodology of Troiani et al. (2012). Participants could respond for multiple objects, and were all screened so that they had at least a 95% AMT approval rate beforehand. *Fixedness* was assessed by asking participants on a 5-point Likert scale how easily they could pick up and move an object (Auger et al., 2012). *Placeness* measured the degree to which people classified an object as either a place or a thing, on a binary scale (Troiani et al., 2012). *Spatial definition* was determined based on whether participants felt an object evokes a strong sense of surrounding space and is hard to imagine in isolation (Mullally and Maguire, 2011). Lastly, *context* was determined based on the degree to which there is a consistent environment an object occurs in (Bar and Aminoff, 2003). Participants stated where they would normally find each object (e.g., a fork in a kitchen), and an entropy score was calculated to indicate degree of context.<sup>3</sup> Object distance has also been correlated with PPA activity (Amit et al., 2012), and we anticipate it has been controlled for by controlling for object size.

A separate set of stimuli was used for independent functional localizers for the relevant regions of interest. These functional runs involved blocked images of: 1) isolated faces, balanced by race, gender, and facial expression, 2) scenes, with  $\frac{1}{2}$  manmade indoor images,  $\frac{1}{4}$  manmade outdoor images, and  $\frac{1}{4}$  natural outdoor images, 3) isolated hands, of various skin tones and positions, 4) isolated bodies in various positions without heads, 5) isolated objects, with balanced size and interaction envelope, and 6) scrambled images, done as a randomized mosaic of  $20 \times 20$  blocks of pixels from the object images. A separate set of object images was also used to identify the ROIs described in Konkle and Oliva (2012), using isolated images of unambiguously big (and large envelope) and small (and small envelope) objects against a white background. Objects in these localizers had no overlap with objects in the main task or the other localizers.

<sup>1</sup> In this AMT study, 100 participants (15 responses for each object) were asked to answer how many hands they would use when first interacting with an object, and we confirmed that the one versus two hands sets were significantly different in their mean responses ( $p = 1.25 \times 10^{-35}$ ).

<sup>2</sup> We compared the spectral energy of images across conditions at five increasing radii cutoffs of the images' power spectra, and selected stimuli so that there were no significant differences at any spectral range level (Torralba and Oliva, 2003).

<sup>3</sup> Entropy is calculated using the formula  $E = -\sum(p(x) \times \log(p(x)))$ , for all different answers  $x$ , where  $p(x)$  is the proportion of responses that are  $x$  (Bar and Aminoff, 2003; Troiani et al., 2012). This score ranges from  $-1.18$  (where all fifteen answers for an object are different) up to 0 (where all fifteen answers for an object are the same).



**Fig. 1.** All 320 stimuli used in the fMRI experiments, grouped by condition. The first overarching analysis of the study contrasts activity for small versus big objects, but the main manipulation can be seen in the 2-factor design of the small objects, covarying physical size with interaction envelope size.

### 2.3. MRI acquisition and preprocessing

The experiment was conducted at the Athinoula A. Martinos Imaging Center at the MIT McGovern Institute for Brain Research. Imaging data were collected on a 3 T Siemens fMRI scanner, using a 32-channel phased-array head coil. Anatomical images were acquired with a high-resolution ( $1 \times 1 \times 1$  mm voxels) T1 MPRAGE structural scan, while functional images were obtained with a gradient echo-planar T2\* sequence (33 axial slices parallel to the anterior commissure – posterior commissure line, no gap, TR = 2 s,  $3.1 \times 3.1 \times 3.1$  mm voxels). Functional data were preprocessed using BrainVoyager QX 2.6 (Brain Innovation, Formisano et al., 2006), including slice time correction, linear trend removal, trilinear motion correction,

$1/128$  Hz temporal high pass filtering, and a 4 mm spatial smoothing FWHM kernel. Functional data were aligned to the anatomical scans, which had white-matter based inhomogeneity correction and were transformed to fit the Talairach space.

### 2.4. Experimental design

A main experiment and a control experiment with two separate sets of participants were conducted with the same scanner task design. In the scanner, participants viewed sequential blocks of images and performed an orthogonal one-back task, where they pressed a button for a consecutively repeated image. The experiment consisted of eight runs, each taking 7.1 min and consisting of 16 blocks, with 20 images

**Table 1**  
Statistics for object properties across conditions.

|                                 | Intercorrelated Phys. & Env. |                           |                        | Orthogonalized Phys. & Env. |                           |                       |
|---------------------------------|------------------------------|---------------------------|------------------------|-----------------------------|---------------------------|-----------------------|
|                                 | Phys. 1, Env. 1<br>(Mean)    | Phys. 2, Env. 2<br>(Mean) | T-test<br>p-value      | Phys. 2, Env. 1<br>(Mean)   | Phys. 1, Env. 2<br>(Mean) | T-test<br>p-value     |
| Physical dimensions             |                              |                           |                        |                             |                           |                       |
| Shipping size, 3D diagonal (in) | 8.11                         | <b>29.51*</b>             | $9.67 \times 10^{-15}$ | <b>25.55*</b>               | 11.84                     | $9.67 \times 10^{-7}$ |
| Shipping weight (lbs)           | 0.40                         | <b>11.79*</b>             | $4.05 \times 10^{-4}$  | <b>3.35*</b>                | 1.19                      | $6.72 \times 10^{-4}$ |
| Subjective AMT ratings          |                              |                           |                        |                             |                           |                       |
| Size (1–7)                      | 1.88                         | <b>3.35*</b>              | $1.84 \times 10^{-14}$ | <b>2.68</b>                 | 2.49                      | 0.25                  |
| Weight (1–7)                    | 1.73                         | <b>3.19*</b>              | $6.97 \times 10^{-12}$ | <b>2.38</b>                 | 2.25                      | 0.50                  |
| Placeness (0–1)                 | 0.003                        | <b>0.063*</b>             | 0.0041                 | <b>0.058</b>                | 0.024                     | 0.159                 |
| Fixedness (0–5)                 | 1.04                         | <b>1.80*</b>              | $1.25 \times 10^{-10}$ | <b>1.25</b>                 | 1.18                      | 0.30                  |
| Context (–1.18–0)               | – <b>0.373</b>               | –0.449                    | 0.1791                 | –0.465                      | – <b>0.431</b>            | 0.55                  |
| Spatial definition (0–1)        | 0.241                        | <b>0.389*</b>             | $4.43 \times 10^{-4}$  | 0.259                       | <b>0.309</b>              | 0.11                  |

Notes—Conditions with the higher mean in each pairing are bolded. If the difference is significant, a star is also added. In the intercorrelated conditions, where both physical size and envelope size are matched, all other properties are intercorrelated, and thus significantly higher for larger objects (other than context, which is non-significant). In contrast, in the orthogonalized conditions, where physical size and envelope size are pitted against each other, there is no significant difference between conditions (though, except for spatial definition and context, the properties all trend towards increasing with larger physical size). Thus, if any effects appear due to larger interaction envelope, it is unlikely that it will be due to any of these other factors.

per block. Each image was presented for 600 ms, with 200 ms fixation between images, and 10 s fixation between blocks. Among the 16 blocks in a run, half were of the large-scale objects condition, while the other half were split amongst the four small-scale objects conditions (two blocks per condition). Image and condition order were shuffled randomly within a run, and each run contained all stimuli in the experiment. For the main experiment, while performing the one-back task, participants were asked to attend to how they might interact with each object, to encourage deeper processing of the images and to ensure participants did not solely focus on low-level visual cues. For the control experiment, participants were instead asked to attend to the physical size of each object in the real world. These two designs were used to ensure there was no top-down task-based influence on the results.

### 2.5. Regions of interest

Perceptual ROIs important to the processing of small entities, large entities, and a control set of ROIs were functionally localized for this study. The large entity category includes scene-selective regions PPA (Epstein and Kanwisher, 1998), retrosplenial cortex (RSC; Auger et al., 2012), and OPA (Dilks et al., 2013) which have been implicated in modulation by several object properties (Troiani et al., 2012), as well as regions in the PHC and TOS (Big-PHC and Big-TOS, or combined as “Big ROIs”) that show selectivity for large objects over small objects (Konkle and Oliva, 2012). The small entity category includes object-selective region lateral occipital complex (LOC; Grill-Spector et al., 1999) and regions in the occipitotemporal sulcus (OTS) and lateral occipital (LO) that show selectivity for small objects over large ones (Small-OTS and Small-LO, or “Small ROIs”). Additionally, the hand-selective area in the left lateral occipital sulcus (LOS; Bracci et al., 2010) has been included as a potential “small entity” region due to its selectivity to object effectors, or objects that extend a hand’s reach (Bracci and Peelen, 2013). Lastly, two additional ROIs not commonly associated with object or scene processing were also included as control regions: the face-selective fusiform face area (FFA; Kanwisher et al., 1997), and the body-selective extrastriate body area (EBA; Downing et al., 2001).

These ROIs were defined for each participant using independent localizers, following the guidelines of other published works (see Section 2.2). A first functional localizer run involved blocked images of six stimulus types: 1) isolated faces, 2) scenes, 3) isolated hands, 4) isolated bodies, 5) isolated objects, and 6) scrambled objects. As participants viewed these blocks of images, they performed a one-back task. This localizer run took 10.6 min and consisted of 24 blocks total, or four each of the six stimulus types, arranged in random order. Like the main experiment, images were presented at  $350 \times 350$  pixels, and

were shown for 600 ms, with 200 ms fixation between images and 10 s fixation between blocks. A second localizer was composed of images of big objects (e.g., a pool table) and small objects (e.g., a dart) following the methodology of Konkle and Oliva (2012), in order to examine interaction envelope effects in regions previously identified for size selectivity. Participants viewed blocks of images and pressed a button when a red frame appeared around the image. The run took 8.83 min, and consisted of ten blocks each of the two object sizes. Images were  $500 \times 500$  pixels in size, and were presented for 500 ms, with 300 ms fixation between images and 10 s fixation between blocks. Of the sixteen participants in the main experiment, twelve completed two runs of this size localizer. Due to time constraints, two were only able to complete one run, and two were not able to complete any runs—resulting in fewer participants with these ROIs than other ROIs.

The ROIs were defined using BrainVoyager’s ROI tool with the following contrasts: PPA, RSC, and OPA as scenes > objects; FFA as faces > scenes; LOC as objects > scrambled; EBA as body > scrambled; and LOS as hand > scrambled minus overlap with the EBA (Bracci et al., 2010). Big-PHC and Big-TOS were defined as big objects > small objects; and Small-OTS and Small-LO as small objects > big objects. ROIs were chosen as clusters at a level of false discovery rate (FDR) < 0.05 or 0.1 (if there was no activity at the 0.05 level). The Big-PHC and Big-TOS were found to have the same patterns of activity, and so have been combined in the results section into the “Big ROIs,” and similarly the Small-OTS and Small-LO also have the same patterns of activity, and thus have been combined into the “Small ROIs”. These Big and Small ROIs were found to spatially overlap with some of the other ROIs; this is important to note, as it means that effects found in these overlapping ROIs may not necessarily be independent. On average, 3.5% (SD = 6.1%, N = 9) of the Small-OTS was found to overlap with the LOS, 18.9% (SD = 32.6, N = 9) of the Small-LO was found to overlap with the LOC, 17.0% (SD = 22.1%, N = 13) of the Big-TOS was found to overlap with the OPA, and 60.8% (SD = 34.0%, N = 13) of the Big-PHC was found to overlap with the PPA. As one can see, there was high variance in degree of overlap between these ROIs.

Each ROI was examined unilaterally, and when there was no significant difference in activity found between the two hemispheres, the ROIs were combined by taking the mean percent signal change between the left and right ROIs. Two regions were examined unilaterally: the Small ROIs, which have only been shown to have left-lateralized significant activity for smaller objects (Konkle and Oliva, 2012), and the LOS, which has also only shown left-lateralized activity for hand stimuli (Bracci et al., 2010). However, all results also generalize when these ROIs are examined bilaterally. Both unilateral and bilateral results for all ROIs are included in the *Supplementary Material*. Additionally,

summaries of all of the ROIs' mean centroids and frequencies in participants can be found in the *Supplementary Material*.

### 3. Results

#### 3.1. Establishing the large versus small entity ROIs

First, cortical activity for a broad range of perceptual ROIs were compared for viewing small-scale objects (e.g., a pencil), versus large-scale objects (e.g., a car), to confirm selectivity of different regions based on large differences in scale and to ensure replicability of previous object size-based effects (Konkle and Oliva, 2012) with the current methodology. The ROIs can be grouped into three different categories: 1) ROIs mainly associated with large entities (scenes, big objects), 2) ROIs mainly associated with small entities (small objects, hands), and 3) other control regions of interest.

Results of this ROI analysis can be seen in Fig. 2. As expected, large-scale entity ROIs (PPA, OPA, the Big ROIs, and RSC) all show the same pattern of having significantly more blood-oxygen-level dependent (BOLD) response upon viewing large-scale objects versus small-scale objects (PPA:  $t(15) = 11.03$ ,  $p = 1.36 \times 10^{-8}$ , RSC:  $t(14) = 5.60$ ,  $p = 6.58 \times 10^{-5}$ , OPA:  $t(14) = 8.10$ ,  $p = 1.18 \times 10^{-6}$ , Big ROIs:  $t(12) = 7.68$ ,  $p = 5.73 \times 10^{-6}$ ).

Amongst the small-scale entity ROIs, indeed the Small ROIs and LOS show a significant preference for small-scale over large-scale objects (Unilaterally, left Small ROIs:  $t(12) = 4.60$ ,  $p = 6.09 \times 10^{-4}$ , left LOS:  $t(13) = 3.76$ ,  $p = 2.40 \times 10^{-3}$ ; bilaterally, Small ROIs:  $t(12) = 4.43$ ,  $p = 8.14 \times 10^{-4}$ , LOS:  $t(13) = 3.55$ ,  $p = 3.60 \times 10^{-3}$ ). The LOS has been shown to have selectivity not only to pictures of hands, but also to objects that extend a hand's reach (Bracci and Peelen, 2013), and thus these results provide evidence that the LOS may be additionally involved in object representations.

Lastly, several ROIs show no difference between activities for small versus large-scale objects. While LOC is an object-selective region, it does not show a preference for objects of any particular scale ( $p = 0.297$ ), supporting its role as a general object-processing region tuned to object shape (Grill-Spector et al., 1999). Additionally, two control regions not associated with object processing – the FFA ( $p = 0.277$ ) and EBA ( $p = 0.524$ ) – show no selectivity to object scale.

These results confirm previous literature in ROI sensitivity to different scale stimuli (Troiani et al., 2012; Konkle and Oliva, 2012) and also establish the ROI preferences for the next critical analysis. We hypothesize that the large-scale entity ROIs will be sensitive to object interaction envelope even for small-scale objects. We do not expect this effect in LOC, FFA, or EBA because they show no sensitivity to the contrast of small-scale and large-scale objects, different in both size and interaction space. The small-scale entity ROIs could potentially hold representation of an object's physical size, its interaction envelope, or both.

#### 3.2. Object interaction envelope sensitivity in large-scale ROIs irrespective of physical size

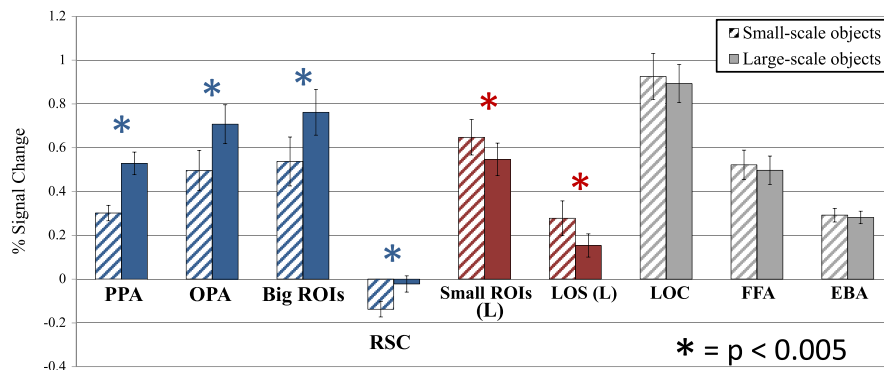
##### 3.2.1. ROI analyses

These same ROIs were then examined for significant effects of object interaction envelope size and physical size within the small-scale objects. We conducted 2-factor repeated-measures ANOVAs and also calculated eta-squared as effect size for each ROI. Fig. 3 shows the average percent signal change for each condition within each ROI, along with significant differences and effect sizes.

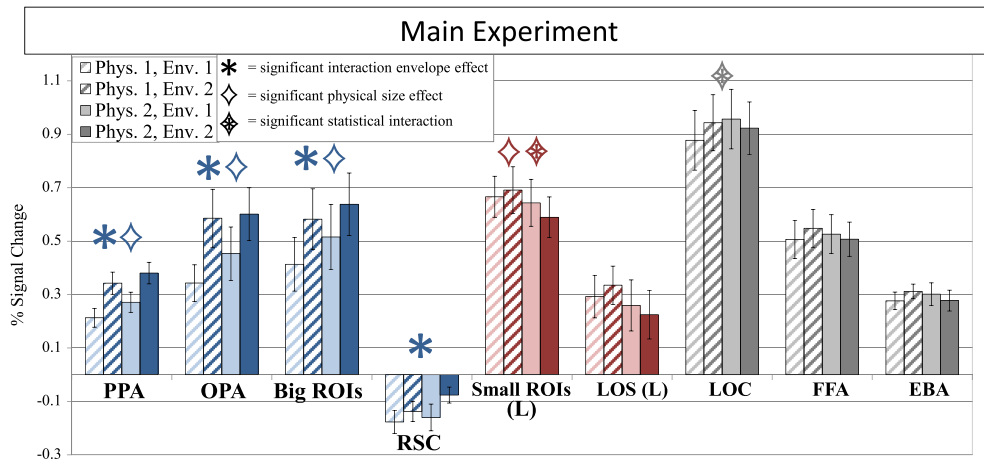
The large-scale entity regions all show a significant preference for the larger interaction envelope factor (PPA:  $F(1,15) = 65.45$ ,  $p = 7.50 \times 10^{-7}$ ,  $\eta^2 = 0.134$ ; RSC:  $F(1,14) = 7.90$ ,  $p = 0.014$ ,  $\eta^2 = 0.038$ ; OPA:  $F(1,14) = 35.06$ ,  $p = 3.73 \times 10^{-5}$ ,  $\eta^2 = 0.069$ ; Big ROIs:  $F(1,12) = 81.22$ ,  $p = 1.09 \times 10^{-6}$ ,  $\eta^2 = 0.033$ ). The PPA, OPA, and Big ROIs (i.e., Big-PHC and Big-TOS combined) also show a significant effect of size, but of smaller effect size (PPA:  $F(1,15) = 12.44$ ,  $p = 0.003$ ,  $\eta^2 = 0.022$ ; OPA:  $F(1,14) = 11.35$ ,  $p = 0.005$ ,  $\eta^2 = 0.007$ ; Big ROIs:  $F(1,12) = 18.57$ ,  $p = 0.001$ ,  $\eta^2 = 0.01$ ). There are no significant statistical interactions in these regions (PPA:  $p = 0.661$ ; RSC:  $p = 0.425$ ; OPA:  $p = 0.109$ ; Big ROIs:  $p = 0.412$ ). This indicates that these scene-selective and large-object regions are indeed sensitive to the local spatial information of small-scale objects in addition to physical size.

On the other hand, within the small-scale entity regions, the Small ROIs (i.e., Small-OTS and Small-LO combined) show a significant preference for smaller physical size (Left:  $F(1,9) = 13.49$ ,  $p = 0.005$ ,  $\eta^2 = 0.015$ ; Bilateral:  $F(1,11) = 6.87$ ,  $p = 0.024$ ,  $\eta^2 = 0.004$ ). There is also a significant effect of the statistical interaction with interaction envelope and physical size, although of lower effect size (Left:  $F(1,9) = 7.76$ ,  $p = 0.021$ ,  $\eta^2 = 0.006$ ; Bilateral:  $F(1,11) = 5.31$ ,  $p = 0.04$ ,  $\eta^2 = 0.013$ ). The LOS (left or bilateral) shows no significant effect of either factor (for the left LOS, envelope size:  $p = 0.871$ ; physical size:  $p = 0.178$ ), but its pattern follows the trend of a small physical size preference. Lastly, for the remaining ROIs (LOC, FFA, EBA), there is no significant effect of either factor (physical size or envelope size—all  $p > 0.20$ ), although there is a significant statistical interaction in LOC ( $F(1,14) = 5.93$ ,  $p = 0.029$ ,  $\eta^2 = 0.004$ ). These results indicate that small-scale entity regions are instead sensitive to object physical size than interaction envelope.

These interaction envelope effects in the large-scale entity regions could potentially be driven by the two conditions where object physical size and interaction envelope are intercorrelated (i.e., physical size 1 and envelope size 1; physical size 2 and envelope size 2). To test this, separate independent t-tests were performed between those conditions, and the conditions where the object properties are fully orthogonalized (i.e., physical size 1 and envelope size 2; physical size 2 and envelope size 1). First, the intercorrelated conditions show significant effects of interaction envelope in the same regions as found with the



**Fig. 2.** ROI analysis of percent signal change for viewing objects of small scale versus large scale. Regions are grouped based on whether they are more selective for larger entities (in blue: PPA, OPA, Big ROIs, RSC), smaller entities (in red: left Small ROIs, left LOS), or neither (LOC, FFA, EBA). Asterisks indicate significant differences in a paired t-test (all  $p < 0.005$ ), and error bars indicate standard error of the mean.



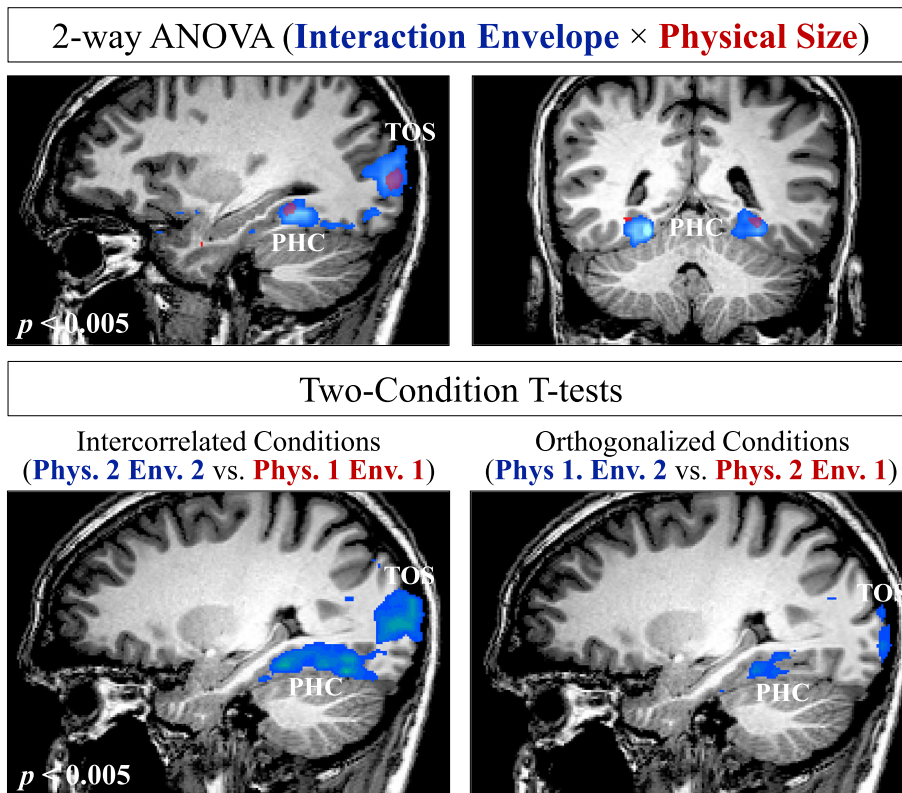
**Fig. 3.** A chart of percent signal change for the ROI analysis in the Main Experiment. ROI graphs are grouped as in Fig. 2, with ROIs with selectivity to large entities in blue, ROIs for smaller entities in red, and other ROIs in gray. Error bars indicate standard error of the mean. Stars indicate a significant effect of interaction envelope, diamonds indicate a significant effect of physical size, and the star within the diamond indicates a significant effect of the statistical interaction of the factors. Significance was determined at a level of  $p < 0.05$ . For each ROI, the highlighted symbol is the effect with the higher effect size.

main experiment, in the PPA ( $t(15) = 7.44, p = 2.09 \times 10^{-6}$ ), the RSC ( $t(15) = 3.33, p = 0.005$ ), the OPA ( $t(15) = 5.29, p = 9.02 \times 10^{-5}$ ), and the Big ROIs ( $t(7) = 2.86, p = 0.024$ ). However, the fully orthogonalized conditions also show significant effects of interaction envelope in key regions of the PPA ( $t(15) = 4.12, p = 9.03 \times 10^{-4}$ ), the OPA ( $t(15) = 4.53, p = 4.0 \times 10^{-4}$ ), and the Big ROIs ( $t(7) = 2.74, p = 0.029$ ). There is no significant effect for the RSC ( $p = 0.514$ ). These results demonstrate that the PPA, OPA, and the Big ROIs in particular

appear to be regions that maintain information about interaction envelope, regardless of object physical size.

3.2.2. Whole-brain analyses

The results that emerge in the ROI analyses can be further visualized in whole brain analyses of the data (see Fig. 4). In a 2-way repeated-measures ANOVA analysis for interaction envelope and physical size, a significant effect of interaction envelope can be found bilaterally in the



**Fig. 4.** A random-effects whole brain analysis ( $N = 16$ ). (Top) F-maps (a sagittal and coronal view) for a repeated-measures 2-way ANOVA, showing voxels with a significant main effect ( $p < 0.005$  uncorrected) for interaction envelope (blue) and physical size (red). (Bottom) T-maps for general linear models on the conditions. The intercorrelated condition map (left) shows the contrast of significant activation ( $p < 0.005$  uncorrected) for objects of both small physical size and envelope (red—no voxels emerge) versus objects of both large physical size and envelope (blue). The orthogonalized conditions map (right) shows the contrast of significant activation of the other two conditions where size and interaction envelope are not correlated—small physical size and large envelope size (blue) versus large physical size and small envelope size (red—no voxels emerge).

posterior PHC and TOS, encompassing areas where the PPA, OPA, Big-PHC, and Big-TOS regions would fall for individual participants. While in Fig. 4 this effect is visualized at a level of  $p < 0.005$  uncorrected, it still also exists bilaterally at a FDR  $< 0.05$  level. Similarly, a significant effect of physical size can also be found in posterior PHC and TOS, however with smaller significant regions of activation, and no effects remain at a FDR  $< 0.05$  level.

To see the degree of this response that remains when interaction envelope and physical size are orthogonalized, further t-tests were conducted (Fig. 4). Contrasting the two conditions where the two factors are intercorrelated (physical size 1, envelope size 1, and physical size 2, envelope size 2), as expected we find posterior PHC and TOS activity only for the condition of higher physical and interaction envelope size. However, the critical analysis is what patterns emerge when the two factors are orthogonalized (physical size 1, envelope size 2 versus physical size 2, envelope size 1). Will significant patterns appear for the condition with higher physical size, or for the condition with higher envelope size? We find that ultimately there is significant activation only for the condition with higher envelope size (physical size 1, envelope size 2), again bilaterally in the key regions of the posterior PHC and TOS. These results indicate that these traditional scene-selective regions maintain representations of object interaction envelope, regardless of object physical size.

### 3.3. Object interaction envelope sensitivity regardless of task

Several of these results replicate in the control study, where participants are asked to attend to each object's physical size (see Fig. 5). Indeed, one possible interpretation of the main experiment's results could be that attending to object interaction envelope in the task artificially inflates responses in scene-selective regions, as task-dependent activity for object processing has been found in other work (Harel et al., 2014). However, in this control study, PPA, OPA, and the Big ROIs still show a significant effect of interaction envelope size (PPA:  $F(1,11) = 7.52, p = 0.019, \eta^2 = 0.036$ ; OPA:  $F(1,11) = 41.94, p = 4.57 \times 10^{-5}, \eta^2 = 0.042$ ; Big ROIs:  $F(1,4) = 28.49, p = 0.006, \eta^2 = 0.025$ ), though RSC is no longer significant ( $p = 0.621$ ). Similarly, PPA and the Big ROIs still show a significant effect of physical size with lower effect size (PPA:  $F(1,11) = 6.66, p = 0.026, \eta^2 = 0.026$ ; Big ROIs:  $F(1,4) = 14.92, p = 0.018, \eta^2 = 0.006$ ), and OPA is no longer significant for a size effect ( $p = 0.096$ ). The EBA now shows a significant preference for smaller interaction envelope ( $F(1,11) = 5.291, p = 0.042, \eta^2 = 0.010$ ). Small ROIs still show a significant preference for small physical size (Left:  $F(1,8) = 10.49, p = 0.012, \eta^2 = 0.011$ ;

Bilateral:  $F(1,8) = 8.34, p = 0.020, \eta^2 = 0.009$ ). A significant effect of the statistical interaction between the factors can be found in the OPA ( $F(1,11) = 5.23, p = 0.043, \eta^2 = 0.014$ ) and the Big ROIs ( $F(1,4) = 7.87, p = 0.049, \eta^2 = 0.003$ ).

If the two studies are combined into a 3-way independent measures ANOVA (experiment  $\times$  interaction envelope  $\times$  physical size), the PPA ( $F(1,104) = 9.07, p = 0.003$ ) and the OPA ( $F(1,100) = 5.68, p = 0.019$ ) still emerge as having a significant effect of interaction envelope (refer to the *Supplementary Material*). No regions emerge for a main effect of physical size (all  $p > 0.10$ ), although it must be noted that independent measures ANOVAs are more stringent than the repeated-measures ANOVAs used when analyzing within studies. Additionally, no regions emerge as significant for the interaction of the experiment and interaction envelope (all  $p > 0.28$ ), nor for the interaction of the experiment and physical size (all  $p > 0.71$ ). This lends further evidence that task is not modulating ROI sensitivity to interaction envelope.

Overall, these results are consistent with the main study. The large-entity regions (except for the RSC) still show a robust significant effect of object interaction envelope, despite a task emphasizing physical size. Meanwhile, the small-entity regions still show a significant effect of object physical size in the control study.

## 4. Discussion

### 4.1. Object interaction envelope sensitivity in scene-selective regions

Using a carefully controlled stimulus set in a visual fMRI task, we found that scene-selective regions (PPA, OPA, Big ROIs) are indeed selective to object interaction envelope, even for small, manipulable objects of equal real-world size. These results hold true even when participants are asked to focus on the real-world size of the objects, indicating that these effects are likely to be implicit and automatic, rather than top-down and task-dependent. While these regions also show effects of object physical size, these effects have a smaller statistical effect size and do not remain when interaction envelope is directly compared to physical size in a whole brain analysis. Amongst these regions, effects in the PPA and OPA remain in all analyses, even when completely orthogonalizing the stimulus conditions, comparing across the main and control experiments, and when analyzing whole-brain contrasts for interaction envelope. In contrast, small-scale entity processing regions (Small ROIs and LOS) show selectivity to objects of smaller real-world size, but not smaller object interaction envelope. Additionally, while other perceptual ROIs (LOC, FFA, EBA) show some effects that may warrant future investigation (e.g., LOC sensitivity to the statistical

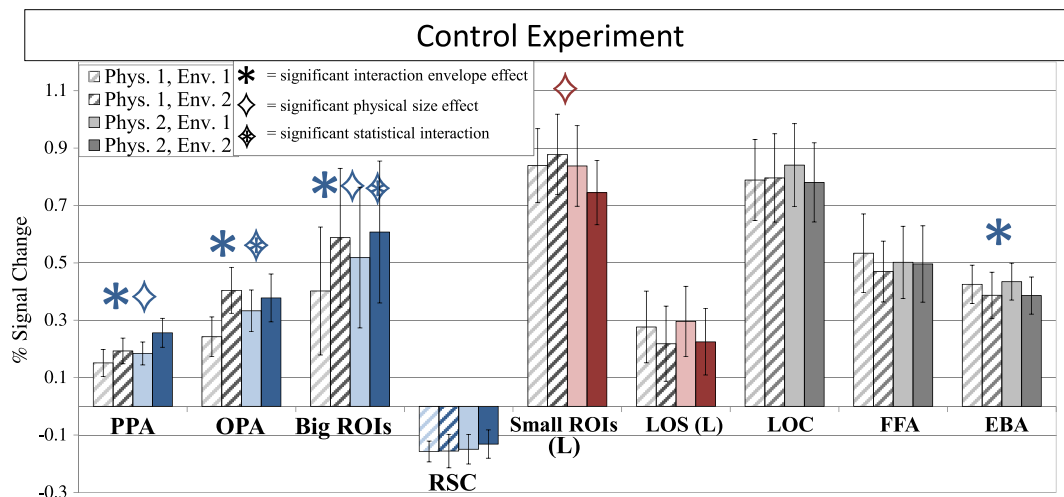


Fig. 5. A chart of percent signal change for each ROI in the Control Experiment, where participants are asked to attend to the size of each object while performing a 1-back task. This graph is read in the same way as Fig. 2.

interaction of the two factors), they ultimately show no effects that are preserved across both studies and all analyses. Overall, these results highlight different calculations on image properties occurring in separate perceptual cortical regions, with the PPA and OPA sensitive to interactive spatial information for images ranging from scenes down to small objects, and the small-scale processing regions instead sensitive to object physical size. Sensitivity to object interaction envelope is specific to ROIs traditionally associated with selectivity to scenes, and these regions may be maintaining geometric calculations about local, functionally-defined space around objects of all scales as well as scenes.

#### 4.2. Object interaction envelope and spatial definition

This large-entity region sensitivity to object interaction envelope aligns with the hypothesis presented by Mullally and Maguire (2011), that along with processing information about scenes and layouts (Epstein and Kanwisher, 1998; Epstein et al., 2003; Troiani et al., 2012), the PPA also has representations of local space around individual objects. While Mullally and Maguire define this local space as “spatial definition,” which is loosely defined and directly related to the volume of the objects, the current results provide an operationalization of spatial definition as a measure of the space through which we interact with objects (i.e., interaction envelope). These results also complement Park et al.'s (2014) recent work in scene processing that finds that scene size and clutter – both properties with a direct correspondence to interaction envelope with the scene – are encoded in the PHC. In fact, the selectivity of the PPA to scenes may not be the result of a strong dissociation between scenes and objects, but instead a continuum of interaction envelope, where smaller objects may present relatively limited interactions, larger objects may present more possibilities with increasing movement in a three-dimensional space, and scenes may provide the upper end with a wide range of interactive possibilities. These results with object interaction envelope may also explain results found with other object properties, as previous studies have used properties highly intercorrelated with all other object properties (Troiani et al., 2012), while the current study has studied object interaction envelope in isolation.

The current study parameterizes object interaction envelope by whether the object can be interacted with on a small scale (i.e., with one hand) or a larger scale (i.e., with two hands). Further work would be to determine what calculations make up this perceptual property of interaction envelope. For example, possible calculations that could be incorporated into this property could include the amount of surface area of contact, points of contact, complexity of the interaction, and percentage of bodily involvement. Additionally, interaction envelope could also be applied beyond objects to the realm of scenes, to see if scene interactivity or navigability can be quantified in a similar way.

#### 4.3. Object interaction envelope: action or perception?

These results also could hold possible implications for the action and perception two streams hypothesis. While lesion data and neuroimaging data have indicated that the ventral (perception) and dorsal (action) streams may act independently (Haxby et al., 1991; Goodale and Milner, 1992), other behavioral work has found evidence of action influencing perception (Wohlschläger, 2000; Hommel et al., 2001) and perception influencing action (Grèzes et al., 2003). Object interaction envelope could potentially represent an action-based perceptual property, as it is related to the way in which someone interacts with or navigates around an object or scene. However, according to this current study, this property is implicitly encoded by the brain during a perceptual task, in perceptual regions (PPA and OPA). It could then be possible that this represents the encoding of an action-based property in the ventral visual stream. Indeed, Milner and Goodale's (2008) revisiting of the two-streams hypothesis supports the possibility of some abstract

action planning occurring in the ventral stream, while the dorsal stream handles bottom-up motor control. However, it is possible that PPA and OPA sensitivity to interaction envelope is a more traditional perceptual property (e.g., the size of interaction envelope), and the interactivity is determined upstream of the ventral visual stream. More neuroimaging studies are needed to fully understand what properties or calculations make a property “ventral” (i.e., perception) versus “dorsal” (i.e., action). For example, the task in the current study is a perceptual task—participants are presented the stimuli for only a brief amount of time (600 ms) and are asked to identify two repeated images in a row. Future work could study more motor-based tasks, such as mental imagery action tasks, to examine how object interaction envelope relates to action-related processing in the dorsal stream.

This study serves as a step forward in understanding the properties that govern object representations in high-level ventral visual cortex, supporting the framework of an interconnected perceptual network, where different regions are co-opted for different geometric calculations (i.e., shape, size, patterns, spatial interaction) for processing a wide range of stimuli, ranging from small, isolated objects up to large, complex scenes. While small-entity regions are sensitive to object physical size, traditionally scene-selective regions are additionally maintaining representations of spatial properties of any object, big or small.

#### Acknowledgments

The authors would like to thank Radoslaw M. Cichy, Daniel D. Dilks, Seyed-Mahdi Khaligh-Razavi, and Talia Konkle for their comments on this work. This work was supported by the US National Eye Institute (grant number EY020484). W.A.B. is supported by the Department of Defense, through the National Defense Science & Engineering Graduate Fellowship (NDSEG) Program.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.neuroimage.2015.07.066>.

#### References

- Amit, E., Trope, Y., Mehudar, E., Yovel, G., 2012. Activation of ventral visual cortex supports distance representation. *Brain Cogn.* 80, 201–213.
- Auger, S.D., Mullally, S.L., Maguire, E.A., 2012. Retrosplenial cortex codes for permanent landmarks. *PLoS ONE* e43620.
- Bar, M., Aminoff, E., 2003. Cortical analysis of visual context. *Neuron* 38, 347–358.
- Bracci, S., Peelen, M., 2013. Body and object effectors: the organization of object representations in high-level visual cortex reflects body-object interactions. *J. Neurosci.* 33, 18247–18258.
- Bracci, S., Ietswaart, M., Peelen, M.V., Cavina-Pratesi, C., 2010. Dissociable neural responses to hands and non-hand body parts in the extrastriate visual cortex. *J. Neurophysiol.* 103, 3389–3397.
- Cate, A.D., Goodale, M.A., Köhler, S., 2011. The role of apparent size in building- and object-specific regions of ventral visual cortex. *Brain Res.* 1388, 109–122.
- Dilks, D.D., Julian, J.B., Paunov, A.M., Kanwisher, N., 2013. The occipital place area (OPA) is causally and selectively involved in scene perception. *J. Neurosci.* 33, 1331–1336.
- 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, R., 2005. The cortical basis of visual scene processing. *Vis. Cogn.* 12, 954–978.
- Epstein, R., Kanwisher, N., 1998. A cortical representation of the local visual environment. *Nature* 392, 598–601.
- Epstein, R., Graham, K.S., Downing, P.E., 2003. Viewpoint-specific scene representations in human parahippocampal cortex. *Neuron* 37, 865–876.
- Formisano, E., Di Salle, F., Goebel, R., 2006. Fundamentals of data analysis methods in fMRI. In: Landini, L., Positano, V., Santarelli, M.F. (Eds.), *Advanced Image Processing in Magnetic Resonance Imaging*. CRC Press, USA, pp. 481–503.
- Ganaden, R.E., Mullin, C.R., Steeves, J.K., 2013. Transcranial magnetic stimulation to the transverse occipital sulcus affects scene but not object processing. *J. Cogn. Neurosci.* 25, 961–968.
- Goodale, M.A., Milner, A.D., 1992. Separate visual pathways for perception and action. *Trends Neurosci.* 15, 20–25.
- Grèzes, J., Tucker, M., Armony, J., Ellis, R., Passingham, R.E., 2003. Objects automatically potentiate action: an fMRI study of implicit processing. *Eur. J. Neurosci.* 17, 2735–2740.



- Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzhak, Y., Malach, R., 1999. Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* 24, 187–203.
- Harel, A., Kravitz, D.J., Baker, C., 2014. Task context impacts visual object processing differentially across the cortex. *PNAS* 111, E962–E971.
- Haxby, J.V., Grady, C.L., Horowitz, B., Ungerleider, L.G., Mishkin, M., Carson, R.E., Hershovitch, P., Schapiro, M.B., Rapoport, S.I., 1991. Dissociation of object and spatial visual processing pathways in human extrastriate cortex. *Proc. Natl. Acad. Sci. U. S. A.* 88, 1621–1625.
- Hommel, B., Müsseler, J., Aschersleben, G., Prinz, W., 2001. The Theory of Event Coding (TEC): a framework for perception and action planning. *Behav. Brain Sci.* 24, 849–937.
- Janzen, G., van Turenout, M., 2004. Selective neural representation of objects relevant for navigation. *Nat. Neurosci.* 7, 673–677.
- Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* 17, 4302–4311.
- Kim, J.G., Biederman, I., Lescroart, M.D., Hayworth, K.J., 2009. Adaptation to objects in the lateral occipital complex (LOC): shape or semantics? *Vis. Res.* 49, 2297–2305.
- Konkle, T., Oliva, A., 2012. A real-world size organization of object responses in occipitotemporal cortex. *Neuron* 74, 1114–1124.
- Kourtzi, A., Kanwisher, N., 2001. Representation of perceived object shape by the human lateral occipital complex. *Science* 293, 1506–1509.
- Milner, A.D., Goodale, M.A., 2008. Two visual systems re-viewed. *Neuropsychologia* 46, 774–785.
- Mullally, S.L., Maguire, E.A., 2011. A new role for the parahippocampal cortex in representing space. *J. Neurosci.* 31, 7441–7449.
- Mullin, C.R., Steeves, J.K., 2011. TMS to the lateral occipital cortex disrupts object processing but facilitates scene processing. *J. Cogn. Neurosci.* 23, 4174–4184.
- Park, S., Chun, M.M., 2009. Different roles of the parahippocampal place area (PPA) and retrosplenial cortex (RSC) in scene perception. *Neuroimage* 2009, 1747–1756.
- Park, S., Brady, T.F., Greene, M.R., Oliva, A., 2011. Disentangling scene content from spatial boundary: complementary roles for the PPA and LOC in representing real-world scenes. *J. Neurosci.* 34, 1333–1340.
- Park, S., Konkle, T., Oliva, A., 2014. Parametric coding of the size and clutter of natural scenes in the human brain. *Cereb. Cortex* <http://dx.doi.org/10.1093/cercor/bht418>.
- Rajimehr, R., Devaney, K.J., Bilenko, N.Y., Young, J.C., Tootell, R.B.H., 2011. The “parahippocampal place area” responds preferentially to high spatial frequencies in humans and monkeys. *PLoS Biol.* 9, e1000608.
- Torralba, A., Oliva, A., 2003. Statistics of natural image categories. *Network* 14, 391–412.
- Troiani, V., Stigliani, A., Smith, M.E., Epstein, R.A., 2012. Multiple object properties drive scene-selective regions. *Cereb. Cortex* <http://dx.doi.org/10.1093/cercor/bhs364>.
- Wohlschläger, A., 2000. Visual motion priming by invisible actions. *Vis. Res.* 40, 925–930.
- Wolfe, J.M., Võ, M.L.-H., Evans, K.K., Greene, M.R., 2011. Visual search in scenes involves selective and nonselective pathways. *Trends Cogn. Sci.* 15, 77–84.