Neuron
Article

A Real-World Size Organization of Object Responses in Occipitotemporal Cortex

Talia Konkle[1,*] and Aude Oliva[1,2]
[1]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02138, USA
[2]Computer Science and Artificial Intelligence Laboratory (CSAIL), Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02138, USA
*Correspondence: tkonkle@mit.edu
DOI 10.1016/j.neuron.2012.04.036

SUMMARY

While there are selective regions of occipitotemporal cortex that respond to faces, letters, and bodies, the large-scale neural organization of most object categories remains unknown. Here, we find that object representations can be differentiated along the ventral temporal cortex by their real-world size. In a functional neuroimaging experiment, observers were shown pictures of big and small real-world objects (e.g., table, bathtub; paperclip, cup), presented at the same retinal size. We observed a consistent medial-to-lateral organization of big and small object preferences in the ventral temporal cortex, mirrored along the lateral surface. Regions in the lateral-occipital, inferotemporal, and parahippocampal cortices showed strong peaks of differential real-world size selectivity and maintained these preferences over changes in retinal size and in mental imagery. These data demonstrate that the real-world size of objects can provide insight into the spatial topography of object representation.

INTRODUCTION

One of the most robust results in visual neuroscience is the systematic response of a large section of ventral temporal cortex to objects and shapes (Grill-Spector and Malach, 2004; Milner and Goodale, 1995; Ungerleider and Mishkin, 1982). To date, only a few object categories—namely faces, bodies, and letter strings—have been shown to have focal cortical regions that show strong category selectivity (Cohen et al., 2000; Downing et al., 2001; Kanwisher et al., 1997; McCarthy et al., 1997). Most other object categories such as shoes and cars do not have a clear spatially clustered region of selective cortex but instead activate a large swath of occipitotemporal cortex with distinct and reliable patterns (Carlson et al., 2003; Cox and Savoy, 2003; Haxby et al., 2001; Norman et al., 2006; O'Toole et al., 2005). A fundamental endeavor of cognitive neuroscience is to understand the nature of these object responses and how they are organized across this cortex (e.g., Kourtzi and Connor, 2011; Ungerleider and Bell, 2011).

The animate-inanimate distinction is the only known dimension that gives rise to spatially large-scale differential patterns of activity across ventral temporal cortex (e.g., Chao et al., 1999; Kriegeskorte et al., 2008; Mahon and Caramazza, 2011): this organization encompasses face- and body-selective regions (Kanwisher et al., 1997; Peelen and Downing, 2005) and scene-selective regions (Epstein and Kanwisher, 1998). For the remaining object categories, which have a more distributed response, there is currently no evidence for other factors that give rise to a large-scale organization of this object information. Interestingly, pattern analysis methods which can classify objects based on the response profile in occipitotemporal cortex do not often examine the spatial distribution of these activation profiles. Typically, these approaches assume that the distinctions between these other kinds of objects are spatially heterogeneous, reflected at a fine-scale of organization (e.g., Norman et al., 2006). However, recent evidence shows that object classification in this cortex is robust to increased spatial smoothing (Op de Beeck, 2010) and can even generalize across subjects (Shinkareva et al., 2008). This suggests that there may be a large-scale organization to these distributed object patterns that we have not yet uncovered (Op de Beeck et al., 2008; Freeman et al., 2011).

For an active observer in the natural world, objects are fundamentally physical entities. As such, an intrinsic but surprisingly overlooked property of any object is its real-world size (Konkle and Oliva, 2011). The size of objects in the world has consequences for both the nature of the objects and our experiences with them. For example, gravity and the laws of physics impose specific constrains on the shape and material properties of objects of different sizes. If an object is simply scaled up in size, the increased weight per unit surface area will cause objects with insufficient material strength to collapse, and many natural objects tend to have optimized proportions that are neither overly strong or weak for their size (Haldane, 1928; Gordan, 1981). Additionally, the physical size of objects in the world dictates how we interact with them: we pick up small objects like strawberries and paperclips, but we sit in and move around large objects like sofas and fountains. Thus different-sized objects have different action demands and typical interaction distances. Given these constraints of the physical world on the properties of objects and how we experience them, we hypothesized that object representations may be naturally differentiated by their real-world size, reflected in a coarse spatial organization across occipitotemporal cortex.

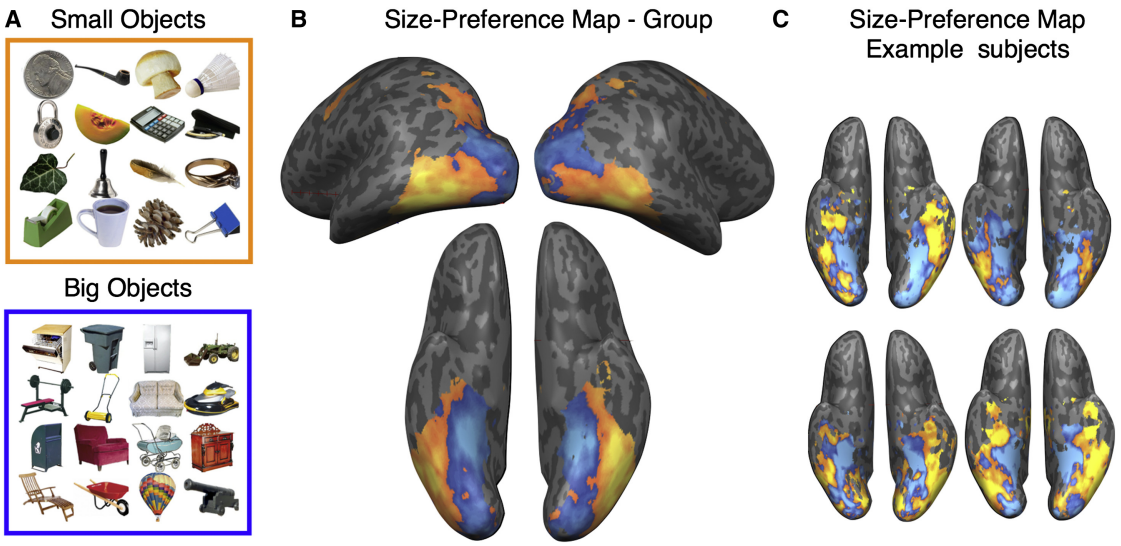


Figure 1. Size-Preference Analysis

(A) Example objects of small and big real-world sizes. Note that all images were presented at the same retinal size. The stimulus set contained 200 small and 200 big objects.

(B) Size-preference maps in the group data. Voxels with a greater response to small objects than big objects are shown in orange; Voxels with a greater response to big objects than small object are shown in blue. The data are plotted on the inflated brain of one participant. There is a medial-to-lateral arrangement of big-to-small object preferences along the ventral surface.

(C) Size-preference maps for four example subjects.

See also Figure S1.

In the current study, we compared the cortical response to big and small real-world objects. We specifically focused on the representations of everyday inanimate objects, excluding faces, bodies, animals, and classically defined tools. These everyday objects often get grouped together as "other objects" (e.g., see Hasson et al., 2003; Op de Beeck et al., 2008) and are known to have a distributed activation pattern across a large swath of ventral-temporal cortex. Here, we examined whether voxels along this cortex showed a preference for objects of big or small real-world sizes. One possibility is that big and small object preferences would be weak and heterogeneously distributed, in a "salt-and-pepper" organization that is not consistent across people. Instead, we observe that there are strong differential responses to big and small objects, and these preferences are grouped spatially in a medial-to-lateral arrangement across the ventral surface of cortex. This organization of object information is mirrored along the lateral surface, with an inferior-to-superior arrangement of small-to-big object information.

Within this organization, we find reliable spatially clustered regions that show peaks of differential selectivity to big and small objects, evident at the single-subject level. We thus characterized the responses in these new functional regions-of-interest to examine the nature of the object representations. We find that responses here are selective to real-world size despite changes in retinal size, indicating relatively high-level object-centered responses. Further, these regions respond during mental imagery of big and small objects, which is a characteristic property of other nearby category-selective regions. Finally, we find that these regions reflect information about the category of the object rather than how big the object was conceived. Broadly, these results show that real-world size is a large-scale dimension that differentiates distributed object representations in occipitotemporal cortex. We propose a potential account of this organization, in which the size of objects in the world naturally give rise to systematic biases in visual experience which are extracted in early visual areas and ultimately dictate where high-level object representations will be in anterior occipitotemporal cortex.

RESULTS

Organization of Big and Small Objects

In Experiment 1a, observers were presented with images of isolated big objects (e.g., car, piano) and isolated small objects (e.g., strawberry, safety pin), presented at the same retinal size on the screen (Figure 1A; for all stimuli see Figure S1 available online; see Experimental Procedures). The experiment consisted of one run of 8.8 min of scanning, during which 200 distinct big objects and 200 distinct small objects were presented in a standard blocked design (see Experimental Procedures). To compare the neural response of big and small objects, we conducted a size-preference map analysis and a whole-brain contrast analysis.

Size-Preference Analysis

In the first analysis, we visualized the spatial distribution of small and big object preferences across occipitotemporal cortex. Size-preference maps were computed reflecting whether the voxels had a preference for big objects (blue) or small objects (orange) within an object-responsive mask (see Experimental

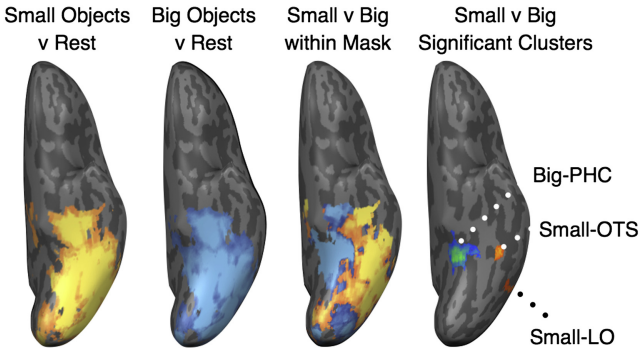


Figure 2. Single Subject Example

Responses in left ventral temporal cortex of an example subject, reflecting (1) small objects > rest, T > 2.0, (2) big objects > rest, T > 2.0, (3) the size-preference map masked by these small or big object-responsive voxels, and (4) the regions of significant differential selectivity for small versus big objects resulting from a whole-brain contrast, FDR < 0.05. See also Figure S2.

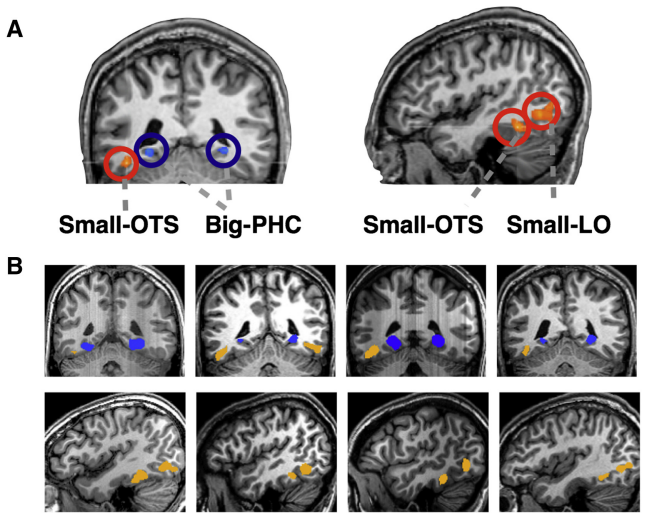
Procedures), and these are shown on an inflated cortical surface in Figure 1. We observed a striking large-scale organization along the ventral surface, evident at the group level and at the single-subject level, with big and small object preferences arranged in a medial to lateral organization across both hemispheres. Further, this organization was mirrored along the lateral surface of the cortex, with small to big object preferences arranged from inferior to superior (Figures 1B and 1C).

Importantly, these data should not be interpreted as evidence that big and small objects are represented in separate swaths of cortex. Both big and small objects activate most of this object-responsive cortex to varying degrees, illustrated in Figures 2, consistent with accounts of distributed activation profiles of these objects (e.g., Haxby et al., 2001). However, voxels with a big-object preference are consistently found along medial ventral temporal cortex, while voxels with a small-object preference were consistently found along lateral temporal cortex (Figures 2C and 2D).

Whole-Brain Contrasts

In a second analysis, we conducted a whole-brain random-effects analysis to identify any contiguous regions with a reliable preference for small objects or for big objects ($p < 0.001$, cluster threshold > 10 mm^3). Along the ventral surface of the brain, a bilateral region of the parahippocampal gyrus was significantly more active to big than to small objects (henceforth labeled as "Big-PHC"), while a left-lateralized region in the occipitotemporal sulcus extending into the inferior temporal gyrus was more active to small relative to big objects (henceforth "Small-OTS"). Along the lateral surface, a more posterior small-preference region was observed ("Small-LO" for lateral occipital), with a big-preference region in the right transverse occipital sulcus ("Big-TOS"; Figure 3).

These regions of interest were also observed reliably in single subjects (Figures 3B and 3C), even with only one run of <10 min of scanning. A left Small-OTS region was present in 9 of 12 participants (bilateral in 1), a left Small-LO region was present in all 12 participants (bilateral in half the participants), and



C

Region		n	X	Y	Z	Peak T
L Inferior Temporal	Small-OTS	9/12	-42	-46	-11	6.0
L Lateral Occipital	Small-LO	12/12	-42	-61	-2	5.9
L Parahippocampal	Big-PHC	10/12	-30	-40	-5	-6.5
R Parahippocampal	Big-PHC	10/12	30	-40	-2	-6.4
R Transverse Occipital	Big-TOS	5/12	12	-73	40	-7.0

Figure 3. Whole-Brain Analysis

(A) Whole-brain contrasts of small versus big objects. Results of a random-effects analysis, small > big contrast, (n = 12, $p < 0.002$, cluster threshold = 10), plotted on the brain of one sample subject (sagittal section, x = −42, coronal section, y = −42). The bilateral region with preference for big objects on the ventral surface is shown (Big-PHC). Two small-preference regions were found, one ventral/anterior (Small-OTS) and one lateral/posterior (Small-LO).

(B) These regions of interest are shown for 4 participants. (C) Table indicating the regions identified from the group random effects analysis. Each region's anatomical location and label are indicated, followed by the number of subjects (n) who showed this region of interest in a single-subject analysis. For each region, the Talairach coordinates of the voxel with the peak t-value (small > big contrast) are reported.

See also Figure S3 and Table S1.

a Big-PHC region was present in 10 of 12 participants (bilateral in all participants). The Big-TOS region was less reliably observed at the single-subject level with a more variable position across subjects, and it was thus not included for further analysis. These results show that big/small object selectivity is more reliable in the left hemisphere, particularly for the Small-OTS and Small-LO regions; an asymmetry opposite that of face-selective regions which show stronger representation in the right hemisphere (Kanwisher et al., 1997).

Comparing these ROIs with the size-preference analysis, it is clear that these regions are not discrete regions of selectivity among a heterogeneous mix of big and small object preferences in the surrounding cortex. Instead, these regions-of-interest reflect the peaks of significant differential activity in an otherwise

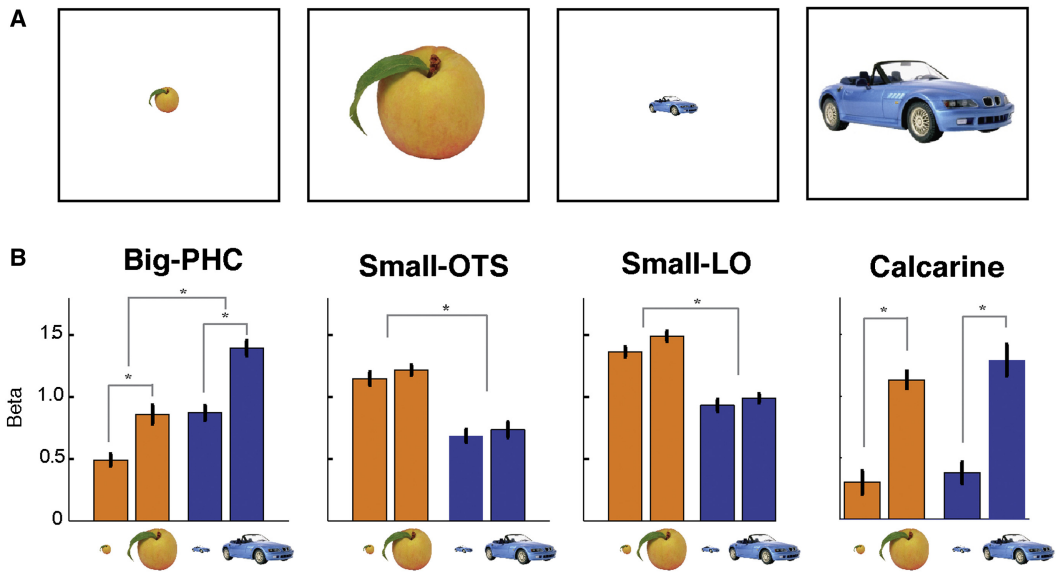


Figure 4. Retinal Size Manipulation Results

(A) Objects of small and big real-world sizes were presented at small and big retinal sizes on the screen.

(B) Activations in independently-localized Big-PHC, Small-OTS, and Small-LO, and anatomically defined early visual cortex regions (Calcarine) were measured with ROI GLMs and the beta weights for the four conditions are plotted for the left hemisphere ROIs. Error bars reflect ± 1 SEM. The Big-PHC region showed effects of both the real-world and the retinal size, while the small regions showed only preference for the real-world of objects with no modulation to retinal size. The early visual control region showed modulation by retinal size, with no effect of real-world size.

See also Table S2.

large-scale organization of big and small object preferences across this cortex. From these data, we do not mean to imply that these entire sections of cortex are devoted solely to representing big objects or small objects. Rather, whatever underlying code is being used to represent object information across this cortex, big and small objects differ strongly in some regions, and the transitions between these regions are more smooth than modular.

Effect Size Estimates

In Experiment 1a, observers were presented with one run of big and small objects. In order to estimate the effect size within these regions, 8 new participants were shown two runs of big and small objects in Experiment 1b. Regions of interest were estimated from the first run for each subject and the magnitude of activation to big and small objects was computed in these regions using data from the second run. All 8 participants showed a Small-OTS region on the left (bilateral in 3) and a Small-LO region (bilateral in all 8), and 7 of 8 showed a Big-PHC region on the left (bilateral in 6 of 8). These regions showed differential responses that were 1.5 to 1.7 times higher for objects of the preferred size relative to objects of the non-preferred size (Figure S3; Table S1; see also Figure 4), close to the effect sizes found in category-selective regions such as the FFA and EBA (Kanwisher, 2010). Mapping out the size-preferences in object-responsive areas in Experiment 1b also confirmed that these regions were peaks of selectivity in a broader map of object size preferences (see Figure S2 for ventral and dorsal maps from both experiments). These results provide an internal replication of Experiment 1a, and demonstrate that within these regions, there is a very large and robust effect of big versus small objects.

While most real-world objects activate nearly the entire ventral surface of cortex significantly more than a fixation baseline, our data indicate that the medial surface has reliably more activity to big objects while the lateral surface has reliably more activity to small objects. Importantly, the pattern-map and whole-brain analyses localize where big and small object information is processed, but they do not inform us about what properties of big and small objects drive the responses. There are a number of factors differentiating big and small objects, and this is true of the difference between faces, bodies, and scenes as well—e.g., in their shapes, in the processing demands, and in more abstract conceptual features regarding their use, importance, or natural kind. In the next experiments, we used a region-of-interest approach to examine the nature of the object representations. Specifically, we examined retinal-size tolerance and activation during mental imagery, and we examined the possibility that these regions are related to an abstract concept of size. For all subsequent experiments, the big versus small object paradigm from Experiment 1 was used as a localizer to independently define regions of interest in each participant that showed a significant difference between small and big objects response (Small-OTS, Small-LO, Big-PHC). While a clear answer to exactly what the big and small object regions and the category-selective regions are representing remains unsolved (e.g., Kourtzi and Connor, 2011; Ungerleider and Bell, 2011), these experiments probe the classic signatures of high-level object representation, serve as important controls, and take initial

steps toward understanding the nature of the representation in this cortex.

Tolerance to Retinal Size Changes

Ventral temporal cortex has object-selective responses that are tolerant to changes in retinal size, position, and viewpoint—a hallmark of high-level object representations (DiCarlo and Cox, 2007; Grill-Spector et al., 1999; Sawamura et al., 2005; Vuilleumier et al., 2002). In Experiment 2, we manipulated the retinal size at which the objects were presented, to examine the response contributions of retinal size and real-world size in these regions.

All of the regions showed more activity to objects of the preferred real-world size independent of retinal size, plotted in Figure 4 (main effect of real-world size: Small-OTS-L: $F(1,23) = 85.8$, $p < 0.001$; Small-LO-L: $F(1,31) = 317.7$, $p < 0.001$; Small-LO-R: $F(1,15) = 57.9$, $p < 0.01$; Big-PHC-L: $F(1,23) = 51.5$, $p = 0.001$; Big-PHC-R: $F(1,23) = 70.3$, $p < 0.001$; no interactions between retinal and real-world size in any of the regions: Small-OTS-L, Small-LO-L, Small-LO-R: all $Fs < 1$; Big-PHC-L: $F(1,23) = 2.3$, $p = 0.19$; Big-PHC-R: $F(1,23) = 3.8$, $p = 0.11$). As a control region, we examined the response in an anatomically-defined region of early visual cortex along the calcarine sulcus. As expected, there was more activity for retinally larger images than retinally smaller images, with no effects of real-world size (calcarine: retinal size: $F(1,27) = 22.8$, $p = 0.003$; real-world size: $F(1,27) = 2.5$, $p = 0.16$).

In the Big-PHC region, there was also a main effect of retinal size, with a stronger response to stimuli presented at retinally large compared to retinally small sizes (main effect of retinal size: Big-PHC-L: $F(1,27) = 14.8$,$p = 0.012$; Big-PHC-R: $F(1,23) = 24.4$, $p = 0.004$; no effect in Small-OTS-L: $F < 1$; Small-LO-L: $F(1,31) = 5.0$, $p = 0.06$; Small-LO-R: $F(1,15) = 1.3$, $p = 0.33$). Thus, the Big-PHC region shows higher response with more peripheral stimulation, for both big and small real-world objects. These results are consistent with other reports of peripheral biases along the collateral sulcus and parahippocampal regions (e.g., Levy et al., 2001; Levy et al., 2004; Arcaro et al., 2009). These results imply that, in this cortex, the features represented are not fully scale-invariant but are also enhanced by general peripheral input.

Critically, the results of Experiment 2 demonstrate that both big and small regions maintained their real-world size selectivity over changes in retinal size—a manipulation that varies the features presented to early areas. Thus, any uneven feature distribution stimulating early foveal versus peripheral visual cortex cannot explain away the activity in the big and small object regions. The overall pattern of results here is consistent with previous characterizations of ventral temporal cortex as "high-level object cortex": what seems to be processed or computed here is strongly related to object-centered information, above and beyond the retinotopic biases in these regions (DiCarlo and Cox, 2007; Grill-Spector et al., 1999; Sawamura et al., 2005; Vuilleumier et al., 2002).

Mental Imagery: Object Identity versus Conceived Size

One potential interpretation of the big and small regions is that the magnitude of activity in these regions is related to the size the observer thinks the object is in the world. On a pure conceived-size account of these regions, the bigger one conceives of an object, the more the object will drive activity in the big region and the less it will drive activity in the small regions, independent of the object's identity (e.g., see Cate et al., 2011). One method to dissociate an object's identity from its real-world size is to use mental imagery processes, where it is possible imagine a tiny piano that is the size of a matchbox or a giant peach that is the size of a car. A tiny piano thus becomes a hand-held object; a giant peach becomes a large object or landmark we can move around. In Experiment 3, we examined whether these regions are tied to the object category or whether the response reflects a more abstract concept of conceived size using a mental imagery task.

Names of objects were presented aurally to a new set of observers, whose task was to form a mental image of each object. In half of the blocks, observers were told to imagine isolated objects at their typical size when they heard the object names (e.g., peach, piano). In the other half of the blocks, they were told to imagine an isolated object at an atypical size: specifically, they heard the adjective "tiny" for big objects and "giant" for small objects: e.g., "tiny piano," imagined with the size of matchbox, or "giant peach," imagined with the size of car (see Experimental Procedures). Afterwards they were presented with small and big objects visually (as in Experiment 1), to independently localize the big and small regions of interest in each subject.

When participants imagined big and small objects at their typical sizes, the big and small regions showed more activity to objects with the preferred real-world size (Figure 5; Small-OTS-L: $t(7) = 2.4$, $p = 0.048$; Small-LO-L marginal: $t(7) = 1.8$, $p = 0.107$; Small-LO-R marginal: $t(6) = 2.1$, $p = 0.083$; Big-PHC-L: $t(6) = 4.0$, $p = 0.007$; Big-PHC-R: $t(7) = 3.2$, $p = 0.015$). These results are consistent with the fundamental and general finding that neural responses in object-selective cortex are similar between perception and imagery (O'Craven and Kanwisher, 2000; Reddy et al., 2010; Stokes et al., 2009). Further, these results also demonstrate that our previous results were not driven by pictoral artifacts of the stimuli: here, any perceptual features instantiated via imagery processes are meaningfully tied to object concepts and are not driven by unintentional feedforward stimulus artifacts.

When observers imagined big and small objects in the atypical-size conditions, the big and small regions did not reflect the conceived size of the object. That is, imagining a giant peach still activated the small-preference regions more than imagining a tiny piano (see Figure 4; Small-OTS-L: $t(7) = 2.6$, $p = 0.036$; Small-LO-L: $t(7) = 2.4$, $p = 0.048$; though not significantly in the right hemisphere Small-LO-R region: $t(6) = 0.8$, $p = 0.45$; Big-PHC-L and Big-PHC-R trending: both $t(7) = 1.7$, $p = 0.13$; see Table S2 for 2×2 ANOVA statistics). These results demonstrate that activity in these big and small regions does not reflect the conceived size of the imagined object—these regions are not reflecting an abstract sense of real-world size independent of the object identity. Instead, the data imply that these regions of cortex represent information that does not change when an object is imagined at a tiny or giant real-world size, such as the category or the visual form of the object. As an analogy, in V1

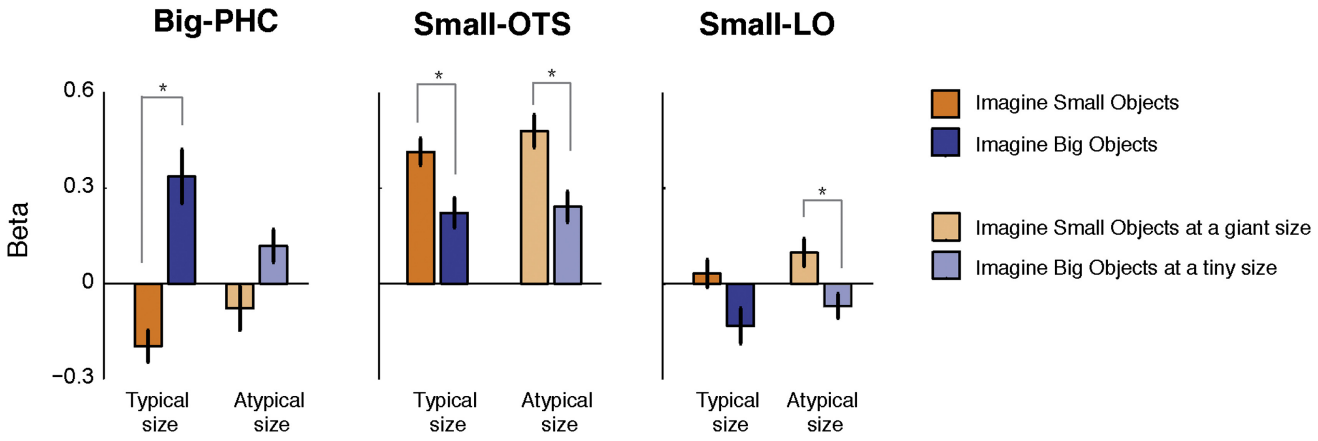


Figure 5. Mental Imagery Results

Activations in independently-localized Big-PHC, Small-OTS, and Small-LO regions in left hemisphere are shown. Orange bars show data for imagined objects with a small real-world size (e.g., strawberries) and blue bars show data for imagined objects with a big real-world size (e.g., pianos). Bars with saturated colors reflect conditions where observers imagined typically sized objects. Bars with unsaturated colors reflect conditions where the objects were imagined at atypical sizes. Error bars reflect ± 1 SEM. See also Table S3.

there is a large-scale map of eccentricity, but what is represented is not eccentricity per se but the orientation and spatial frequency of visual information at that particular eccentricity. Similarly, in these big and small object regions, what is represented is not an abstract sense of real-world size per se, but something specific about the objects that have that particular size in the world.

The Big-PHC region had a less pronounced preference for big relative to small objects when those objects were imagined at atypical sizes (marginally significant interaction: Big-PHC-L: $F(1,27) = 5.9$, $p = 0.051$; Big-PHC-R: $F(1,31) = 5.4$, $p = 0.053$). This result suggests that activity in this region may in part reflect the physical size an observer imagines the object to be (e.g., see Cate et al., 2011). However, a potentially more parsimonious account of these data is that this modulation in the big region is driven by its peripheral preference, as observed in the retinal size manipulation experiment (Figure 4). If observers were imagining giant peaches at a large retinal size and tiny pianos at a small retinal size, and the imagined retinal size affects the spatial extent of activation in early visual areas, then this would give rise to the results observed in the Big-PHC region. Consistent with this interpretation, the small regions did not have any strong modulations by retinal size, and did not show an interaction in the atypical size conditions. While there was no reliable modulations in early visual cortex above baseline in these data (Table S2), previous research supports this interpretation: bigger real-world objects are imagined at bigger retinal sizes (Konkle and Oliva, 2011), and imagining objects at bigger retinal sizes has been shown to drive more peripheral retinotopic responses in early visual areas when measured against a listening baseline (Kosslyn et al., 1995).

DISCUSSION

Most categories of objects do not have a spatially contiguous and highly selective cortical representation, but instead activate a swath of ventral and lateral temporal cortex to varying degrees (Carlson et al., 2003; Cox and Savoy, 2003; Haxby et al., 2001; Norman et al., 2006; O'Toole et al., 2005). Here, we show that within this cortex there are large-scale differential responses to big and small real-world objects. Big versus small object preferences are arranged in a medial-to-lateral organization in ventral temporal cortex in both the left and right hemispheres, and this is mirrored along the lateral surface. Within this large-scale organization, several regions show strong differential activity that survive strict whole-brain contrasts, both at the single subject level and at the group level. A bilateral region in the parahippocampal gyrus was preferentially active to big versus small objects (Big-PHC), while an adjacent region in left occipital temporal sulcus was more active to small versus big objects (Small-OTS), with an additional small-preference region in more posterior lateral occipital cortex (Small-LO). While both big and small objects drove these regions above baseline, the differential activity between objects of different sizes was on the order of 1.5-1.7 times greater for objects of the preferred real-world size.

Using a region-of-interest approach, we probed the nature of the object information in these regions in subsequent experiments. We observed that (1) object responses in these regions maintain their real-world size preferences over changes in retinal size, indicating that these preferences are largely object-based rather than retinotopic; (2) these regions are activated during visual imagery, suggesting they reflect the site of stored visual knowledge about these objects; (3) these regions are not driven by whether an object is conceived of as big or small in the world, indicating that these regions are not representing an abstract concept of real-world size. Thus the real-world size preference cannot be explained by a purely low-level (retinotopic) effect, nor by a purely high-level (conceptual) effect. Instead, our data indicate that the size preferences across ventral cortex arise from information about the object category or visual form and reflect features common among small and among big objects.

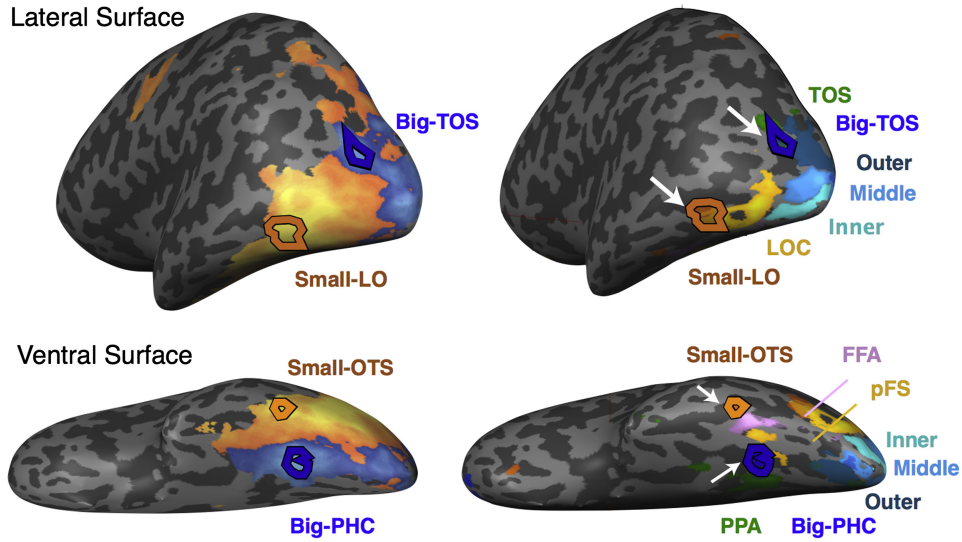


Figure 6. Relationship of These Regions to Eccentricity Bands and to Other Well-Characterized Regions

Left: Group size-preference maps, with the big and small regions of peak selectivity shown from a single representative subject. Right: Functionally-localized regions from the same single subject shown. Inner, middle, and outer eccentricity rings are shown in light, medium, and dark blue, respectively. LOC and pFS (objects > scrambled) are shown in yellow. FFA (faces > objects) is shown in pink; this participant did not have an OFA region. PPA and TOS (scenes > objects) are shown in green. The Small-OTS and Small-LO regions are shown in orange and the Big-PHC and Big-TOS regions are shown in blue, also indicated with white arrows.

Broadly, these data demonstrate that the real-world size of objects can provide insight into the spatial topography of object representations which do not have a focal category-selective response.

Relationship to Surrounding Characterized Regions

Where are the big and small object regions with respect to other well-characterized object and scene regions? Figure 6 shows the big and small object regions overlaid with face-selective, scene-selective, and general shape-selective regions, as well as inner, middle, and outer eccentricity bands (see also Table S3 and Supplemental Experimental Procedures).

Along the ventral surface Big-PHC is partially overlapped with parahippocampal place area (PPA: scenes > objects; Epstein and Kanwisher, 1998), while to our knowledge the Small-OTS region is a relatively uncharted region of cortex that is not overlapping with any other well-characterized regions. The fusiform face area (FFA: faces > objects), fusiform body area (FBA: bodies > objects), and posterior fusiform object region (pFS: objects > scrambled) fall in between the Big-PHC and Small-OTS regions, and are located along the fusiform gyrus (Peelen and Downing, 2005; Schwarzlose et al., 2008). Note that both big and small objects activate the fusiform cortex as well (Figure 2), but show the strongest differential response in more medial and more lateral cortex.

While the scene-selective PPA region is typically localized as scenes > objects (Epstein and Kanwisher, 1998), PPA is known to have a reliable above-baseline response to objects, particularly large objects such as buildings and landmarks (Aguirre et al., 1998; Diana et al., 2008; Downing et al., 2006; Epstein, 2005; Litman et al., 2009; Mullally and Maguire, 2011), and strongly contextual objects (Bar, 2004). Interestingly, strongly contextual objects tend to be larger than non-contextual objects (Mullally and Maguire, 2011). Recently this scene area was shown to respond systematically to imagined objects that define a space (Mullally and Maguire, 2011). Relevant to the current results, in their factor analysis of different object properties, Mullally and Maguire (2011) found that an object's size was highly correlated with its space-defining properties, and this dimension explained a similar amount of response variance in the PPA. Mullally and Maguire (2011) did not explore the role of real-world size outside of the PPA, so their work does not speak directly to the role of real-world size as a general organizational dimension of object-selective cortex. Nevertheless, given the proximity of the Big-PHC region to the PPA, their results are nicely convergent and consistent with the results found here regarding the response profile of medial ventral cortex to large objects, and suggest that the object information in this region may be related to some spatial properties of objects (e.g., spaces/shapes for the body).

Along the lateral surface, Small-LO is just anterior to functional area LOC, localized as objects > scrambled (Grill-Spector et al., 1999), while Big-TOS is nearby scene-selective area TOS (Epstein et al., 2005; Hasson et al., 2003). The lateral occipital cortex contains many nearby and partially-overlapped regions, such as the extrastriate body region EBA, motion area MT, the medial temporal gyrus tool region MGT-TA (Beauchamp et al., 2002; Chao et al., 1999; Downing et al., 2001; Valyear and Culham, 2010). The convergence of these regions also suggests that some abstract spatial property of objects may be represented in these regions (e.g., spaces/shapes for the hands).

Previous studies characterizing category-selective regions along the ventral and lateral surface of visual cortex have found that these regions come in pairs, e.g., faces: fusiform and

occipital face area FFA/OFA; bodies: fusiform and extrastriate body area FBA/EBA; general shape-selectivity: posterior fusiform and lateral occipital complex, pFS/LOC; and scenes: parahippocampal place area and transverse occipital sulcus, PPA/TOS (Schwarzlose et al., 2008; Taylor and Downing, 2011). Hasson et al. (2003) demonstrated that these regions are arranged in a ''mirrored'' fashion from medial-ventral regions wrapping around the lateral surface to medial-dorsal regions. Previous work has found that regions along the ventral surface have more overall visual form information while those along the lateral surface have more location-, motion-, and local-shape information (Beauchamp et al., 2002; Drucker and Aguirre, 2009; Haushofer et al., 2008; Schwarzlose et al., 2008). Consistent with this overall pattern, Small-OTS and Small-LO are paired small regions that also fall along the ventral and lateral surfaces, respectively; Big-PHC may also be paired with Big-TOS but Big-TOS was less reliably observed here and warrants further study. Given this duplication of object representations along the ventral and lateral surface, the different response properties discovered for lateral and ventral category-selective regions in general may also apply to Big-PHC, Small-OTS, and Small-LO.

Relationship to Extended Retinotopy

Object-responsive cortex anterior to early visual areas was originally thought to be nonretinotopic; however, there are now many well-documented retinotopic maps extending along dorsal and ventral streams (e.g., for reviews, see Wandell et al., 2007; Silver and Kastner, 2009). Comparing object responses with retinotopic organization in this cortex may prove to be valuable for understanding the consistent spatial arrangement of category-selective regions (e.g., Levy et al., 2001; Malach et al., 2002; Hasson et al., 2002, 2003; Sayres and Grill-Spector, 2008), as well as the big/small object regions. Here we discuss how the big and small object responses relate to the retinotopic biases in occipitotemporal cortex.

The medial ventral surface has peripheral field biases while the lateral temporal surface has central field biases, which extend directly from early visual areas V1-V4 (Levy et al., 2001; Malach et al., 2002; Hasson et al., 2003; but see Brewer et al., 2005; Arcaro et al., 2009, which suggest that there are separate foveal representations in these regions). Face- and scene-selective areas are found in cortex with foveal and peripheral biases, respectively (e.g., Levy et al., 2001; Hasson et al., 2002). Similarly, given the positions of the big/small object regions relative to the scene/face regions, there is a striking convergence between big and small object information and the eccentricity biases of high-level object areas. For example, Figure 6 illustrates that Big-PHC region is near to peripheral early visual cortex, while the Small-OTS and Small-LO preferences are closer to foveal early visual cortex, and both organizations are mirrored along the lateral surface. This convergence raises the possibility that big/small preferences may derive in part from eccentricity biases.

In their eccentricity-bias proposal of the organization of object representation, Malach and colleagues proposed a processing-based organization of cortex, positing that areas with foveal or peripheral biases carry out fine-detailed or integrative processing, respectively. On this account, any object will be represented along this cortex based on its processing-resolution needs (e.g., Malach et al., 2002). This account has met with some criticisms, however, as the concept of processing-resolution was not clearly operationalized (see also Tyler et al., 2005). For example, it is not obvious that faces require fine-detail processing and not integrative processing. Further, this proposal does not easily lead to predictions about the location of other objects until it is first determined what kind of processing resolution they require (e.g., what are the processing demands of a strawberry or a chair?). As such, their proposal does not easily predict or account for the big-small organization of this cortex. However, in the following section we suggest an alternative account of the object-size organization which shares a fundamental premise of the eccentricity-bias proposal, namely that there is a meaningful relationship between the organization of visual object responses and the large-scale eccentricity organization of early visual areas.

Implications for the Spatial Topography of Object Representation

How might object representations come to be differentiated by real-world size in this object-responsive cortex? Here, we propose a possible account of how this organization emerged from a combination of size-dependent biases in perceptual input, and size-dependent biases in functional requirements for action. Our proposal derives from two core ideas regarding the goals of the visual system: (1) to efficiently represent systematic biases in the sensory input (e.g., along shape, retinal size, curvature, etc, e.g., Attneave, 1954; Carlson et al., 2011; Field, 1987), and (2) to facilitate action in the natural environment (Gibson, 1979; e.g., computing what effectors you use to interact with an object). Our account describes how these convergent pressures could give rise to object representations organized by real-world size in occipitotemporal cortex. Although this account is speculative and will require future work for direct supporting evidence, it nevertheless it provides a principled framework with testable predictions to guide future research.

For observers in the world, there are certain geometric constraints that we suggest give rise to a systematic covariance between an object's real-world size, shape, and experienced eccentricity. For example, although an observer can stand at any distance from an object, allowing the object to project to any retinal size, some distances of interaction may be more frequent than others. A car at a typical viewing distance of 30 feet subtends a visual angle of ~30 degrees, whereas a raisin held at an arm's length subtends a much smaller visual angle of ~1 degree, and would nearly have to touch the eye to subtend a visual angle of 30 degrees. Thus, over the course of natural viewing experience, in the lifetime or over evolutionary time, larger objects may tend to extend more peripherally on the retina than smaller objects (see also Konkle and Oliva, 2011). Additionally, we suggest that shape may be intrinsically correlated with object size based on gravitational and physical constraints of the world—e.g. smaller objects tend to be rounder and larger objects tend to be boxier (Konkle, 2011). These shape constraints manifest as systematic biases in low-level shape features such as curvature and spatial frequency content stimulating early visual areas. Based on the prominent framework

that the visual system is tuned to the natural statistics of the world (e.g., Simoncelli and Olshausen, 2001), early processing stages along the visual hierarchy may extract these low-level feature covariances in orientation, curvature, and eccentricity (e.g., Carlson et al., 2011). Due to the large-scale organization of eccentricity in early visual cortex, this could give rise to pre-cursor object representations that are naturally arrayed along the cortical sheet by real-world size. Consequently an object's real-world size would predict the location of its peak representation.

A prominent alternative account for the large-scale spatial organization of object information is the connectivity-hypothesis proposed by Mahon and Caramazza, which argues that object representation is driven by long-range network connectivity (Mahon and Caramazza, 2011; Mahon et al., 2007). On this account, manipulable objects like tools require different "downstream" action requirements than animate objects like animals, and this determines the organization of ventral stream representations. Interestingly, the real-world size of objects naturally constrains the kinds of actions and effectors that will be used when an observer interacts with an object (e.g., with the fingers, hands, arms, or full body). By incorporating the notion of real-world size into action requirements, it may be possible to extend their proposal beyond animals and tools to the large range of other biological and manmade artifacts. Thus, real-world object size may not only be related to the eccentricity and shape features of objects, but may also be a natural proxy for different classes of action and interaction types, as reflected in ventral-dorsal connectivity. While the eccentricity-bias and connectivity-driven hypothesis have often been discussed as competing alternatives, our real-world size account may unify these proposals, as here we propose both bottom-up experience-driven learning and top-down requirements for action provide convergent pressures for object knowledge to be topographically organized by real-world size.

EXPERIMENTAL PROCEDURES

Participants

Twenty-two healthy observers with normal or corrected-to-normal vision participated in one or more of the experiments in a 2 hr fMRI session (age 19–36, 13 female, 21 right-handed). Informed consent was obtained according to procedures approved by the MIT Internal Review Board.

MRI Acquisition

Imaging data were collected on a 3T Siemens fMRI Scanner at the Martinos Center at the McGovern Institute for Brain Research at MIT. Experiments 1 and 2 used a 12-channel phased-array head coil and Experiment 3 used a 32-channel phased-array head coil. Blood oxygenation level-dependent (BOLD) contrast was obtained with a gradient echo-planar T2* sequence (33 oblique axial slices acquired parallel to the anterior commissure-posterior commissure line; 64 × 64 matrix; FoV = 256 × 256 mm; 3.1 × 3.1 × 3.1 mm voxel resolution; Gap thickness = 0.62 mm; TR = 2000 ms; TE = 30 ms; flip angle = 90 degrees).

Experiment 1a and 1b: Big and Small Objects

In Experiment 1a, 12 observers completed one run of this experiment. In Experiment 1b, 8 new observers completed two runs of this experiment. Observers were shown images of big real-world objects and small real-world objects in a standard blocked design. All objects were shown at the same visual angle (9 × 9 degrees). Each block was 16 s during which 20 images were shown per block for 500 ms each with a 300 ms blank following each item. Fixation periods of 10 s intervened between each stimulus block. Ten blocks per condition were shown in a single run of 8.8 min (265 volumes). A total of 200 big and 200 small distinct object images were presented. Observers were instructed to pay attention to the objects and to press a button when a red frame appeared around an item, which happened once per block. Regions defined from contrasting small and big objects were used as ROIs in subsequent experiments.

Experiment 2: Retinal Size Manipulation

Eight observers were shown blocks of big and small objects at big and small retinal sizes. The big and small objects stimuli were the same as in Experiment 1, and the retinal sizes were 11 × 11 degrees visual angle and 4 × 4 degrees visual angle for the big and small visual sizes, respectively. The blocked design and stimuli were the same as in Experiment 1: each block was 16 s during which 20 images were shown for 500 ms each with a 300 ms blank following each item. Blocks were separated by fixation periods of 10 s. There were four conditions (2 real-world sizes × 2 retinal sizes), presented in a pseudorandom order, such that all conditions appeared in a shuffled order 5 times per run (8.8 min, 265 volumes). Two runs were conducted in this experiment, yielding 10 blocks per condition. Observers were instructed to pay attention to the objects and to press a button when a red frame appeared around an item, which happened once per block.

Experiment 3: Mental Imagery

The names of different objects were presented aurally to 8 naive observers, and observers were instructed form a mental image of each object. Observer's eyes were closed for the entire duration of each run. In 16 s blocks, observers heard 5 object names (3.2 s per object), followed by the word "blank" signifying the beginning of each 10 s blank interval. Runs always began with a 10 s blank interval. In the typical size conditions, blocks of small object names (e.g., peach) and big object names (e.g., lawn chair) were presented. In the atypical size conditions, observers imagined these small objects at giant sizes (e.g., hearing the words "giant peach") and the big objects at tiny sizes (e.g., hearing the words "tiny lawn chair"). There were 30 small objects and 30 big objects, divided into two sets. Each run used the stimuli from one set and contained 3 blocks of each condition, lasting for 5.4 min (161 volumes). Six runs were conducted in the experiment, three for each object set, yielding 12 total blocks per condition. All imagery runs were conducted first, prior to the presentation of any experiments with visual stimuli, including the Big versus Small Object localizer.

Sounds were presented through Sensimetric MRI Compatible Insert Earphones (www.sens.com/s14/index.htm). To set the volume levels in the scanner, a functional run was started and the volume of the stimuli was slowly increased until the participant pressed a button indicating they could hear the stimuli clearly.

Before the experiment, observers were given detailed instructions that they should imagine only isolated objects, and that "giant" versions of small objects should be imagined "as having the same size as a car or piano" while tiny versions of large objects should be imagined "as having the same size as a matchbox or something that could fit in your hand." Observers then were given a short practice run outside the scanner in which they heard the names of small objects, big objects, tiny versions of big objects, and giant versions of small objects, following the same timing as in the experimental runs. None of these practice object stimuli were used in the main experiment.

Data Analysis

Functional data were preprocessed using Brain Voyager QX software (Brain Innovation, Maastricht, Netherlands). Preprocessing included slice scan-time correction, 3D motion correction, linear trend removal, temporal high-pass filtering (0.01 Hz cutoff), spatial smoothing (4 mm FWHM kernel), and transformation into Talairach coordinates. For the ROI overlap computations, analyses were performed on unsmoothed functional data in ACPC space (no Talairach transform).

Statistical analyses were based on the general linear model. All GLM analyses included regressors for each experimental condition, defined as square-wave regressors for each stimulus presentation time convolved with

a gamma-function to approximate the idealized hemodynamic response. A whole-brain, random-effects group average analysis was conducted on data from the Big versus Small Object Experiment (E1). A contrast was performed at an uncorrected threshold of $p < 0.001$ (with an additional cluster threshold of 10 mm^3 applied) to test for regions more active to small versus big objects and vice-versa.

To obtain size-preference maps for each subject, an object-responsive mask was computed by taking all voxels posterior to $Y = -19$ (to isolate the occipital-temporal lobes) that were active in either the Small > Rest or the Big > Rest contrast at $T > 2.0$. The preference map shows the t values of the small object versus big object contrast, within this object-responsive mask. To compute the group size-preference map, the time series of each subject was concatenated and a fixed-effects GLM analysis was run on the group data (see Hasson et al., 2003; Levy et al., 2001), and the same procedure as in the single subject case was subsequently followed.

To obtain regions-of-interest from the Big and Small Object experiment, whole-brain GLMs were conducted for each individual. The Small-OTS and Small-LO regions were defined from contrasts of Small > Big, and the Big-PHC regions were defined from the opposite contrast of Big > Small. All ROIs were taken from t maps corrected at FDR < 0.05, with a cluster threshold of 10 mm^3 (10 contiguous voxels). In some cases, the FDR threshold was made more conservative, e.g., when the Small-OTS and Small-LO regions, which each have distinct peaks, were connected by voxels with lower t values. If any of the targeted ROIs were not present at FDR < 0.05, the threshold was lowered to FDR < 0.1. If no clear ROI was present at that threshold, then that ROI was not defined for that participant. ROIs were defined as the set of contiguous voxels that were significantly activated around the peak voxel identified from within a restricted part of cortex based on the anatomical position.

For all ROI analyses, all ROIs were defined from the Big versus Small object experiment (independent dataset), and the response of these regions to different experimental conditions was assessed in subsequent experiments. For each subject and each ROI, GLMs were run on the average time series of the voxels in the ROI to obtain regression coefficients (betas) for the experimental conditions. For the subsequent experiments with 2 × 2 designs (Experiment 2: retinal size manipulation; Experiment 3: mental imagery), to evaluate the effects of each factor across observers, repeated-measures ANOVAs were run on the betas across observers for each ROI.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, three figures, and three tables, and can be found with this article online at doi:10.1016/j.neuron.2012.04.036.

ACKNOWLEDGMENTS

This work was funded by a National Science Foundation Graduate Fellowship (to Talia Konkle), and a National Eye Institute grant EY020484 (to Aude Oliva), and was conducted at the Athinoula A. Martinos Imaging Center at McGovern Institute for Brain Research, MIT. We thank George Alvarez, Timothy Brady, Mark Williams, Daniel Dilks, and the reviewers for thoughtful comments on the manuscript.

Accepted: April 24, 2012
Published: June 20, 2012

REFERENCES

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

Arcaro, M.J., McMains, S.A., Singer, B.D., and Kastner, S. (2009). Retinotopic organization of human ventral visual cortex. J. Neurosci. *29*, 10638–10652.

Attneave, F. (1954). Some informational aspects of visual perception. Psychol. Rev. *61*, 183–193.

Bar, M. (2004). Visual objects in context. Nat. Rev. Neurosci. *5*, 617–629.

Beauchamp, M.S., Lee, K.E., Haxby, J.V., and Martin, A. (2002). Parallel visual motion processing streams for manipulable objects and human movements. Neuron *34*, 149–159.

Brewer, A.A., Liu, J., Wade, A.R., and Wandell, B.A. (2005). Visual field maps and stimulus selectivity in human ventral occipital cortex. Nat. Neurosci. *8*, 1102–1109.

Carlson, T.A., Schrater, P., and He, S. (2003). Patterns of activity in the categorical representations of objects. J. Cogn. Neurosci. *15*, 704–717.

Carlson, E.T., Rasquinha, R.J., Zhang, K., and Connor, C.E. (2011). A sparse object coding scheme in area V4. Curr. Biol. *21*, 288–293.

Cate, A.D., Goodale, M.A., and Köhler, S. (2011). The role of apparent size in building- and object-specific regions of ventral visual cortex. Brain Res. *1388*, 109–122.

Chao, L.L., Haxby, J.V., and Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. Nat. Neurosci. *2*, 913–919.

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M.A., and Michel, F. (2000). The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. Brain *123*, 291–307.

Cox, D.D., and Savoy, R.L. (2003). Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. Neuroimage *19*, 261–270.

Diana, R.A., Yonelinas, A.P., and Ranganath, C. (2008). High-resolution multi-voxel pattern analysis of category selectivity in the medial temporal lobes. Hippocampus *18*, 536–541.

DiCarlo, J.J., and Cox, D.D. (2007). Untangling invariant object recognition. Trends Cogn. Sci. (Regul. Ed.) *11*, 333–341.

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

Downing, P.E., Chan, A.W.-Y., Peelen, M.V., Dodds, C.M., and Kanwisher, N. (2006). Domain specificity in visual cortex. Cereb. Cortex *16*, 1453–1461.

Drucker, D.M., and Aguirre, G.K. (2009). Different spatial scales of shape similarity representation in lateral and ventral LOC. Cereb. Cortex *19*, 2269–2280.

Epstein, R. (2005). The cortical basis of visual scene processing. Vis. Cogn. *12*, 954–978.

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

Epstein, R.A., Higgins, J.S., and Thompson-Schill, S.L. (2005). Learning places from views: variation in scene processing as a function of experience and navigational ability. J. Cogn. Neurosci. *17*, 73–83.

Field, D.J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. J. Opt. Soc. Am. A *4*, 2379–2394.

Freeman, J., Brouwer, G.J., Heeger, D.J., and Merriam, E.P. (2011). Orientation decoding depends on maps, not columns. J. Neurosci. *31*, 4792–4804.

Gibson, J.J. (1979). The Ecological Approach to Visual Perception (Boston: Houghton Mifflin).

Gordan, J.E. (1981). Structures: Or Why Things Don't Fall Down (Cambridge, MA: Da Capo Press).

Grill-Spector, K., and Malach, R. (2004). The human visual cortex. Annu. Rev. Neurosci. *27*, 649–677.

Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzchak, Y., and Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. Neuron *24*, 187–203.

Haldane, J.B.S. (1928). On Being the Right Size (Oxford: The Oxford Book of Modern Science Writing).

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

Hasson, U., Harel, M., Levy, I., and Malach, R. (2003). Large-scale mirror-symmetry organization of human occipitotemporal object areas. Neuron *37*, 1027–1041.

Haushofer, J., Livingstone, M.S., and Kanwisher, N. (2008). Multivariate patterns in object-selective cortex dissociate perceptual and physical shape similarity. PLoS Biol. *6*, e187.

Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., and Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science *293*, 2425–2430.

Kanwisher, N. (2010). Functional specificity in the human brain: a window into the functional architecture of the mind. Proc. Natl. Acad. Sci. USA *107*, 11163–11170.

Kanwisher, N., McDermott, J., and Chun, M.M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. J. Neurosci. *17*, 4302–4311.

Konkle, T. (2011). The role of real-world size in object representation. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA.

Konkle, T., and Oliva, A. (2011). Canonical visual size for real-world objects. J. Exp. Psychol. Hum. Percept. Perform. *37*, 23–37.

Kosslyn, S.M., Thompson, W.L., Kim, I.J., and Alpert, N.M. (1995). Topographical representations of mental images in primary visual cortex. Nature *378*, 496–498.

Kourtzi, Z., and Connor, C.E. (2011). Neural representations for object perception: structure, category, and adaptive coding. Annu. Rev. Neurosci. *34*, 45–67.

Kriegeskorte, N., Mur, M., Ruff, D.A., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K., and Bandettini, P.A. (2008). Matching categorical object representations in inferior temporal cortex of man and monkey. Neuron *60*, 1126–1141.

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

Levy, I., Hasson, U., Harel, M., and Malach, R. (2004). Functional analysis of the periphery effect in human building related areas. Hum. Brain Mapp. *22*, 15–26.

Litman, L., Awipi, T., and Davachi, L. (2009). Category-specificity in the human medial temporal lobe cortex. Hippocampus *19*, 308–319.

Mahon, B.Z., and Caramazza, A. (2011). What drives the organization of object knowledge in the brain? Trends Cogn. Sci. (Regul. Ed.) *15*, 97–103.

Mahon, B.Z., Milleville, S.C., Negri, G.A.L., Rumiati, R.I., Caramazza, A., and Martin, A. (2007). Action-related properties shape object representations in the ventral stream. Neuron *55*, 507–520.

Malach, R., Levy, I., and Hasson, U. (2002). The topography of high-order human object areas. Trends Cogn. Sci. (Regul. Ed.) *6*, 176–184.

McCarthy, G., Puce, A., Gore, J.C., and Allison, T. (1997). Face-specific processing in the human fusiform gyrus. J. Cogn. Neurosci. *9*, 605–610.

Milner, A.D., and Goodale, M.A. (1995). The Visual Brain in Action (Oxford: Oxford University Press).

Mullally, S.L., and Maguire, E.A. (2011). A new role for the parahippocampal cortex in representing space. J. Neurosci. *31*, 7441–7449.

Norman, K.A., Polyn, S.M., Detre, G.J., and Haxby, J.V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. Trends Cogn. Sci. (Regul. Ed.) *10*, 424–430.

O'Craven, K.M., and Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stiimulus-specific brain regions. J. Cogn. Neurosci. *12*, 1013–1023.

O'Toole, A.J., Jiang, F., Abdi, H., and Haxby, J.V. (2005). Partially distributed representations of objects and faces in ventral temporal cortex. J. Cogn. Neurosci. *17*, 580–590.

Op de Beeck, H.P. (2010). Against hyperacuity in brain reading: spatial smoothing does not hurt multivariate fMRI analyses? Neuroimage *49*, 1943–1948.

Op de Beeck, H.P., Haushofer, J., and Kanwisher, N.G. (2008). Interpreting fMRI data: maps, modules and dimensions. Nat. Rev. Neurosci. *9*, 123–135.

Peelen, M.V., and Downing, P.E. (2005). Selectivity for the human body in the fusiform gyrus. J. Neurophysiol. *93*, 603–608.

Reddy, L., Tsuchiya, N., and Serre, T. (2010). Reading the mind's eye: decoding category information during mental imagery. Neuroimage *50*, 818–825.

Sawamura, H., Georgieva, S., Vogels, R., Vanduffel, W., and Orban, G.A. (2005). Using functional magnetic resonance imaging to assess adaptation and size invariance of shape processing by humans and monkeys. J. Neurosci. *25*, 4294–4306.

Sayres, R., and Grill-Spector, K. (2008). Relating retinotopic and object-selective responses in human lateral occipital cortex. J. Neurophysiol. *100*, 249–267.

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

Shinkareva, S.V., Mason, R.A., Malave, V.L., Wang, W., Mitchell, T.M., and Just, M.A. (2008). Using FMRI brain activation to identify cognitive states associated with perception of tools and dwellings. PLoS ONE *3*, e1394.

Silver, M.A., and Kastner, S. (2009). Topographic maps in human frontal and parietal cortex. Trends Cogn. Sci. (Regul. Ed.) *13*, 488–495.

Simoncelli, E.P., and Olshausen, B.A. (2001). Natural image statistics and neural representation. Annu. Rev. Neurosci. *24*, 1193–1216.

Stokes, M., Thompson, R., Cusack, R., and Duncan, J. (2009). Top-down activation of shape-specific population codes in visual cortex during mental imagery. J. Neurosci. *29*, 1565–1572.

Taylor, J.C., and Downing, P.E. (2011). Division of labor between lateral and ventral extrastriate representations of faces, bodies, and objects. J. Cogn. Neurosci. *23*, 4122–4137.

Tyler, C.W., Likova, L.T., Chen, C.C., Kontsevich, L.L., Schira, M.M., and Wade, A.R. (2005). Extended concepts of occipital retinotopy. Curr. Med. Imaging Rev. *1*, 319–329.

Ungerleider, L.G., and Bell, A.H. (2011). Uncovering the visual "alphabet": Advances in our understanding of object perception. Vision Res. *51*, 729–799.

Ungerleider, L.G., and Mishkin, M. (1982). Two cortical visual systems. In Analysis of Visual Behavior, D.J. Ingle, M.A. Goodale, and R.J.W. Mansfeld, eds. (Cambridge, MA: MIT Press), pp. 549–586.

Valyear, K.F., and Culham, J.C. (2010). Observing learned object-specific functional grasps preferentially activates the ventral stream. J. Cogn. Neurosci. *22*, 970–984.

Vuilleumier, P., Henson, R.N., Driver, J., and Dolan, R.J. (2002). Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. Nat. Neurosci. *5*, 491–499.

Wandell, B.A., Dumoulin, S.O., and Brewer, A.A. (2007). Visual field maps in human cortex. Neuron *56*, 366–383.

Neuron, Volume *74*

Supplemental Information

A Real-World Size Organization of Object Responses in Occipitotemporal Cortex

Talia Konkle and Aude Oliva

SUPPLEMENTAL DATA

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

SUPPLEMENTAL REFERENCES

Figure S1: related to Figure 1: Big and Small object stimuli

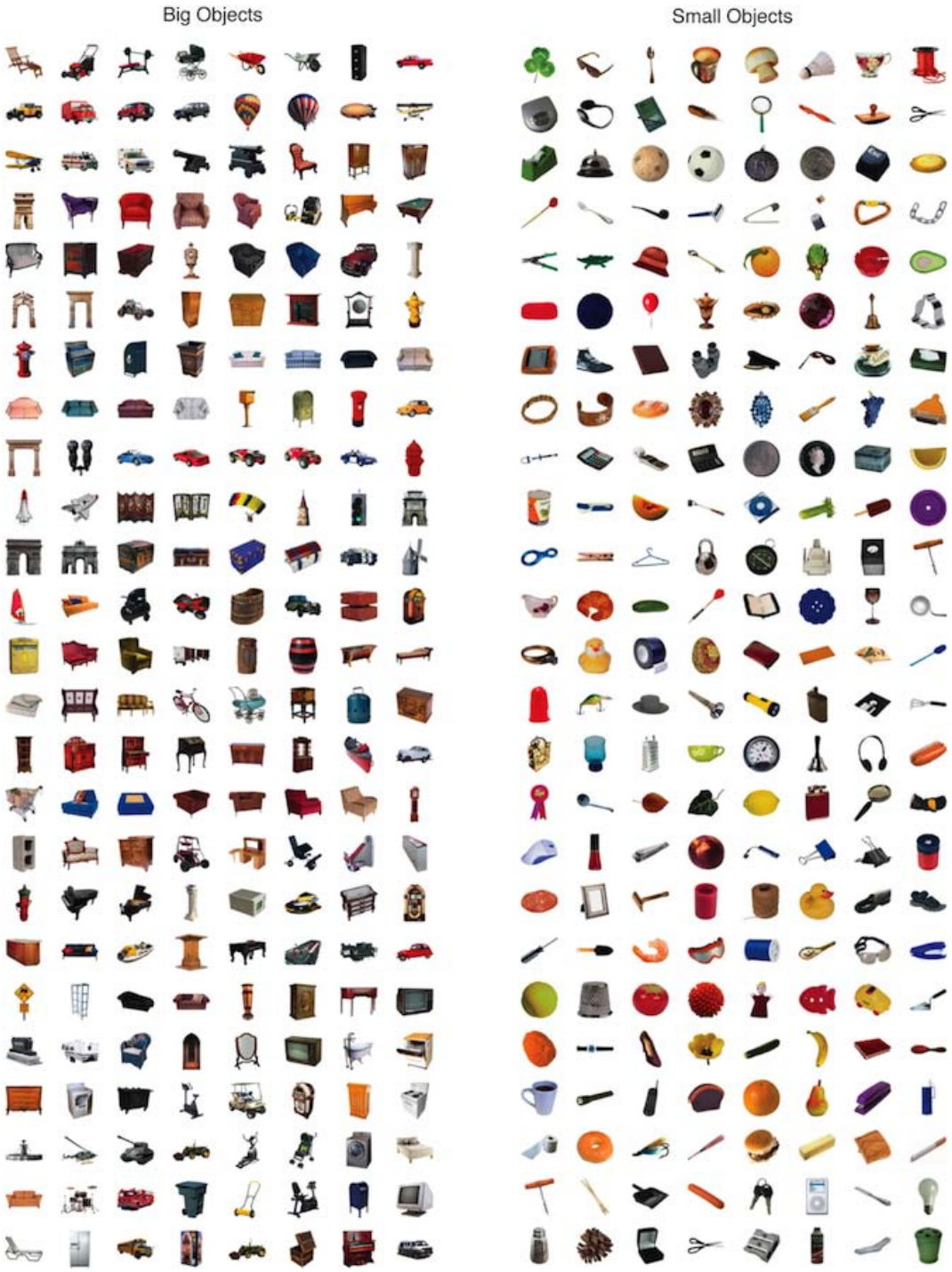


Figure S1. Stimulus set. Thumbnails of the 200 small and 200 big objects used in Experiment 1. The stimulus set can be downloaded from the first author's website.

Figure S2: related to Figure 1: Dorsal and Ventral views of group preference maps

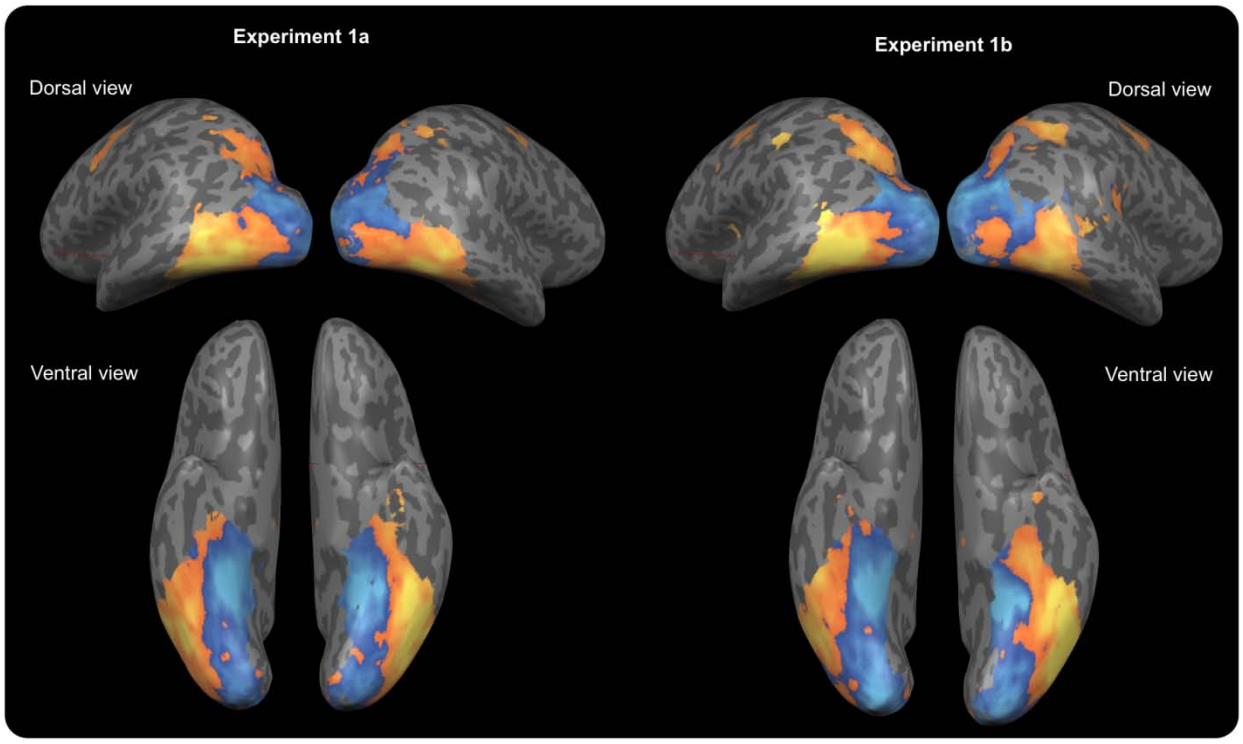


Figure S2. Dorsal and ventral views of group preference maps. The preference maps were computed from a fixed-effect group analysis, masked by voxels significantly more active for big or small objects versus rest. Voxels with a preference for small objects are shown in orange and voxels with a preference for big objects are shown in blue. Left shows results for Experiment 1a (n=12) and right shows results for replication Experiment 1b (n=8).

Figure S3: related to Figure 3: Effect size in big and small preference regions

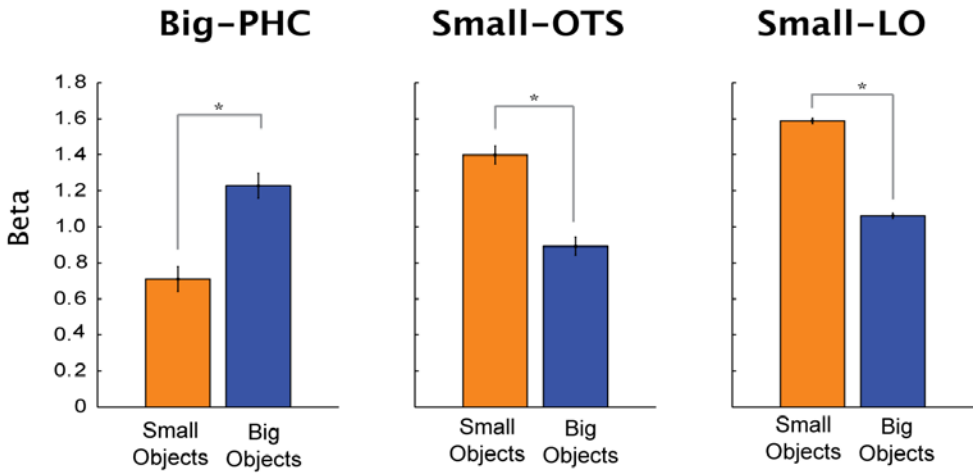


Figure S3. Effect size in big and small preference regions. In Experiment 1b, we presented two runs of the big small object localizer to 8 new participants. Regions of interest were defined from the first run of the big small object localizer, and the magnitude of response in each region was estimated using data from the second run of the same experiment. Beta weights are shown for the big- and small-preference regions in the left-hemisphere. Error bars reflect ± 1 within-subject S.E.M.

We additionally computed percent signal change in these regions by first removing the grand mean of the time course and then computing an ROI GLM, where the beta weights can now be interpreted as percent signal change. The effect size estimates of these regions did not change (mean percent signal change: Big-PHC-L: small objects 27%, big objects 47%; Big-PHC-R: Small objects 32%, big objects 56%; Small-LO-L: small objects 96%, big objects 66%; Small-LO-R: small objects 90%, big objects 68%; Small-OTS-L: small objects 73%, big objects 48%).

Table S1: related to Figure 3: Effect Size estimates in big and small preference regions

	Small Objects Mean Beta (sem)	**Big Objects** Mean Beta (sem)	Small Objects vs. Big Objects	Non-preferred vs. Baseline
BigPHC-L	.71 (.23)	1.23 (.18)	t(6)=3.54, p<0.05	t(6)=3.11, p<0.05
BigPHC-R	.73 (.19)	1.31 (.15)	t(7)=4.46, p<0.005	t(7)=3.85, p<0.05
Small-OTS-L	1.4 (.18)	.89 (.11)	t(6)=4.61, p<0.005	t(6)=7.83, p<0.001
Small-LO-L	1.59 (.14)	1.06 (.14)	t(7)=16.63, p<0.001	t(7)=7.57, p<0.001
Small-LO-R	1.52 (.18)	1.13 (.16)	t(6)=5.77, p<0.001	t(6)=7.15, p<0.001

Table S1. Effect Size estimates in big and small preference regions. Average beta weights for big and small objects in independently localized big and small regions in the left hemisphere. Statistics are reported for t-tests compared the objects of the preferred size vs. non-preferred size, as well as the non-preferred size relative to baseline. Each region showed a reliable differential response to objects of big and small known sizes, while also showing an above-baseline response to objects with the non-preferred size.

Table S2: related to Figure 4: Mental Imagery Experiment Statistics

	Big/Small Object Main effect	Typical/Atypical size Main Effect	Interaction
Big-PHC-L	**F(1,27)=12.5, p=0.012 ***	F(1,27)=0.6, p=0.460	**F(1,27)=5.9, p=0.051 ~**
Big-PHC-R	**F(1,31)=8.9, p=0.020 ***	F(1,31)=0.0, p=0.977	**F(1,31)=5.4, p=0.053 ~**
Small-OTS-L	**F(1,31)=6.9, p=0.034 ***	F(1,31)=1.2, p=0.304	F(1,31)=0.7, p=0.421
Small-LO-L	**F(1,31)=5.8, p=0.047 ***	F(1,31)=2.3, p=0.174	F(1,31)=0.0, p=0.978
Small-LO-R	F(1,27)=3.9, p=0.097	F(1,27)=0.1, p=0.805	F(1,27)=3.3, p=0.118
EarlyV - Inner	F(1,31)=1.0, p=0.340	F(1,31)=2.0, p=0.204	F(1,31)=1.3, p=0.287
EarlyV - Middle	F(1,31)=0.1, p=0.725	F(1,31)=0.1, p=0.738	F(1,31)=0.5, p=0.512
EarlyV - Outer	F(1,31)=0.2, p=0.679	F(1,31)=0.1, p=0.740	F(1,31)=0.5, p=0.508

Table S2. Results of the 2x2 ANOVA for the mental imagery experiment. In Experiment 3, participants imagined big and small objects at normal and atypical sizes. The table shows the results of a 2x2 repeated measures ANOVA on the beta weights for each ROI. All big and small object regions showed a main effect of the real-world size of the object, and no significant effect of typical/atypical size, with a marginal interaction in the Big-PHC region. Early visual cortex in inner, middle, and outer eccentricity bands were not modulated significantly above baseline and showed no differences in response across the imagery conditions.

Table S3: related to Figure 5: Reliability and Overlap of ROIs

	B-PHC1 B-PHC2	PPA1 PPA2	B-PHC1 PPA2	PPA1 B-PHC2	Sm-LO1 Sm-LO2	LOC1 LOC2	Sm-LO1 LOC2	LOC1 Sm-LO2
Sub1	0.91	0.81	0.21	0.01	0.95	0.87	0.52	0.80
Sub2	-	0.98	-	-	1.00	0.93	0.00	0.27
Sub3	0.81	0.86	0.38	0.95	0.75	0.57	0.26	0.46
Sub4	-	0.92	0.19	-	0.32	0.59	0.36	0.19
Sub5	0.98	0.93	-	0.91	-	0.69	0.61	-
Sub6	0.89	0.93	0.68	0.97	0.99	0.78	0.14	-
Sub7	0.93	0.36	0.84	0.44	0.86	0.58	0.27	0.35
Mean	**0.90**	**0.83**	**0.46**	**0.66**	**0.81**	**0.72**	**0.31**	**0.41**
SEM	0.03	0.08	0.56	0.11	0.11	0.06	0.08	0.11

Table S3: Reliability and Overlap of ROIs. Average % containment between a region with itself across runs and with the comparison region across runs. For example, B-PHC1 indicates the big-preference region (big>small) defined from the first functional run; B-PHC2 indicates the big-preference region defined from the second functional run. Sm-LO1 and Sm-LO2 indicate the Small-LO region (small>big) defined in the first and second runs. The calculation procedure and summary of results is described in the Supplementary Experimental Procedures.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Reliability and Overlap of ROIs

The Small and Big ROIs showed proximity to and overlap with the well-known functionally localized regions of the lateral occipital complex (LOC: objects>scrambled) and the parahippocampal place area (PPA: scenes>objects). To characterize the overlap between these areas, we first gathered data from a new set of 7 observers on two runs of the Big and Small ROI localizer and two runs of a PPA and LOC localizer described below. We then computed overlap between two ROIs, e.g. LOC and Small-LO, and compared it to the overlap of the ROI to itself across runs. Overlap was characterized using a procedure modified from previous methods (Scholz et al., 2009).

PPA and LOC were localized from a standard localizer experiment in which stimulus blocks of scenes, objects, faces, and scrambled objects were shown, with each block lasting 16s during which 20 images were shown for 500ms each with a 300 ms blank (images shown at 9x9 degrees visual angle). Fixation periods of 10s preceded and followed each stimulus block. The conditions were presented in a pseudorandom order, such that all 4 conditions appeared in a shuffled order 4 times per run. A run was 7.1 min (213 volumes). Observers were instructed to press a button when a red frame appeared around an item, which happened once per block.

Overlap between two target regions was computed over a range of t-value thresholds then averaged over a range of t-values, allowing for regions to be different sizes. For any two regions being compared, the range of t-values started at the maximum t-value of the two region's FDR<0.05 threshold, and increased by steps of 0.02 to the lowest of the two peak t-values. Additionally, we required a minimum of 10 voxels and a maximum of 500 voxels from both ROIs at any given threshold. The analysis proceeded by getting the contiguous set of voxels around the peak voxel within an anatomically defined mask that were above the specific threshold. At each threshold, degree of overlap was quantified as the percent of voxels of the smaller region that were contained in the larger region, for left hemisphere ROIs only. This measure can be conceptualized as what percent the smaller region is contained in the bigger region without relying on a specific arbitrary t-threshold.

The overlap analysis showed that the LOC region across two runs was 72% (SEM=6%) contained with itself, and the Small-LO region was 81% contained with itself (SEM=11%). Given these numbers as reference of within-ROI reliability, the Small-LO was on average 35% contained in the LOC region (SEM=6%). Thus while there is some overlap between LOC and Small-LO, the regions which show a preference for small objects are not capturing the same region as is localized with objects>scrambled.

The PPA region across two runs was 85% contained with itself (SEM=8%) and the Big region across two runs was 90% contained with itself (SEM = 3%). Comparing these two regions together, the PPA and Big regions were on average 58% contained (SEM=11%) (See Supplementary Table 3). On average there was relatively more overlap between PPA and Big regions then there was with LOC and Small-LO regions.

Inverting the overlap measure: Computing Non-Overlap of ROIs

Neither the Big-PHC and PPA regions, nor the Small-LO and LOC regions were fully overlapped. To estimate the percent of the region that was not overlapped, one can approximate this by subtracting

100% from the %-contained measure. However, this subtraction assumes that the regions being compared were roughly a similar number of voxels, which was not assumed by the %-containment measure.

Thus we conducted an additional, targeted analysis, in which we computed the percent of the Big-PHC region that was not overlapping with the PPA. That is, are there voxels that show a big>small significant difference that do not show a significant scene>object difference? In this analysis, rather than flexibly using the region with the fewest number of voxels as the denominator of the percentage calculation, we used the number of voxels in the Big-PHC region as the denominator in the percentage calculation. We examined this directly by following the same procedure, but computing the percent of voxels in the Big-PHC region not contained in the PPA. This yielded a 51% non-overlap (SEM 11%). In other words, 51% of the voxels with a significant Big>Small contrast did not also have a significant Scene>Object contrast at that threshold, averaged over a range of thresholds for each subject, and averaged across subjects.

We conducted a similar analysis for the Small-LO and LOC regions. That is, are there voxels that show a significant small>big difference that do not show a significant object>scrambled difference? We followed the same procedure, and computed the average percent of voxels in the Small-LO region that were not contained in the LOC region, yielding 70% (SEM 7%).

Note that these numbers are approximately equal to 100% - %contained measure (PPA/BigPHC: 100%-58% = 42%; compared to 51%; LOC/SmallLO: 100%-35%=65%, compared to 70%). These numbers would be exactly equal if the two regions were the same number of voxels at each statistical threshold, which was roughly true on average across subjects. The value of the %-containment measure is that it allows for easy comparison with the within ROI reliability.

SUPPLEMENTAL REFERENCES

Scholz, J., Triantafyllou, C., Whitfield-Gabrieli, S., Brown, E.N. and Saxe, R. (2009). Distinct regions of right temporo-parietal junction are selective for theory of mind and exogenous attention. PLoS One 4, e4869.