

Tripartite Organization of the Ventral Stream by Animacy and Object Size

Talia Konkle^{1,2} and Alfonso Caramazza^{1,2}

¹Department of Psychology, Harvard University, Cambridge, Massachusetts 02138, and ²Center for Mind/Brain Sciences (CiMeC), University of Trento, 38068 Rovereto (TN), Italy

Occipito-temporal cortex is known to house visual object representations, but the organization of the neural activation patterns along this cortex is still being discovered. Here we found a systematic, large-scale structure in the neural responses related to the interaction between two major cognitive dimensions of object representation: animacy and real-world size. Neural responses were measured with functional magnetic resonance imaging while human observers viewed images of big and small animals and big and small objects. We found that real-world size drives differential responses only in the object domain, not the animate domain, yielding a tripartite distinction in the space of object representation. Specifically, cortical zones with distinct response preferences for big objects, all animals, and small objects, are arranged in a spoked organization around the occipital pole, along a single ventromedial, to lateral, to dorsomedial axis. The preference zones are duplicated on the ventral and lateral surface of the brain. Such a duplication indicates that a yet unknown higher-order division of labor separates object processing into two substreams of the ventral visual pathway. Broadly, we suggest that these large-scale neural divisions reflect the major joints in the representational structure of objects and thus place informative constraints on the nature of the underlying cognitive architecture.

Introduction

A basic empirical fact of brain organization is that the spatial organization of information is not random but has systematic structure: neurons with similar functional profiles tend to be nearby each other spatially (Durbin and Mitchison, 1990; Kaas and Catania, 2002; Rosa and Tweedale, 2005; Graziano and Aflalo, 2007; Aflalo and Graziano, 2011). For example, primary sensory cortices have a large-scale organization that follows the topography of the sensory array (e.g., somatotopy along the post-central gyrus, retinotopy along early visual areas). By extension, more distinct kinds of processing have more separation across the cortex. For example, location and object information are famously dissociated along the dorsal “where/how” pathway and the ventral “what” pathway (Ungerleider and Mishkin, 1982; Goodale and Milner, 1992).

In the domain of object processing, however, the key dimensions of object representation, and how they map across the cortex, are still being explored (Kourtzi and Conner, 2011; Ungerleider and Bell, 2011; Kravitz et al., 2013). One core distinction

is between animate and inanimate objects (Spelke et al., 1995; Caramazza and Shelton, 1998; Kuhlmeier et al., 2004; Martin, 2007; Mahon and Caramazza, 2009): fundamentally different kinds of operations apply to each domain, from social communication and theory of mind for animate entities, to manipulation, use, and function for objects. A second dimension relates to the real-world size of objects (Setti et al., 2009; Konkle and Oliva, 2011; Konkle and Oliva, 2012a). All objects are physical entities, and this intrinsically shapes our interactions with them: small objects can be carried and used as effectors, whereas big objects provide support for the body or serve as landmarks in the environment. The importance of these dimensions is also evident in the neural architecture, as both animacy and size distinctions have a large-scale organization along the ventral surface of the brain (Chao et al., 1999b; Downing et al., 2006; Martin, 2007; Bell et al., 2009; Mahon et al., 2009; Wiggett et al., 2009; Konkle and Oliva, 2012b).

Here we examined how these two dimensions of animacy and size combine to shape object responses across the ventral stream. What cortical mapping rules are possible when both dimensions are taken into account? Intuitively the dimensions of animacy and size are orthogonal; that is, there are big and small animals, just as there are big and small objects. If one were to preserve this two-dimensional representational space of objects in a projection to the two-dimensional cortical surface, this would predict that one dimension (e.g., animacy) would map to the cortex along one axis (e.g., medial to lateral) and the other dimension, real-world size, would map along an orthogonal anterior-to-posterior axis. However, previous work has shown that both animacy and size have a medial-to-lateral organization along the ventral surface of the cortex (Chao et al., 1999b; Downing et al., 2006; Martin,

Received March 4, 2013; revised April 12, 2013; accepted May 5, 2013.

Author contributions: T.K. and A.C. designed research; T.K. performed research; T.K. analyzed data; T.K. and A.C. wrote the paper.

This work was supported in part by a grant from the Fondazione Cassa di Risparmio di Trento e Rovereto and by the National Institutes of Health, National Eye Institute (Fellowship F32EY022863-01A1) and was conducted at the Laboratory for Functional Neuroimaging Center at the Center for Mind/Brain Sciences, University of Trento. We thank George Alvarez and Timothy Brady for insightful comments on this manuscript and Krista Ehinger for help with the stimulus set.

The authors declare no competing financial interests.

Correspondence should be addressed to Talia Konkle, William James Hall 913, 33 Kirkland Street, Cambridge, MA 02138. E-mail: tkonkle@fas.harvard.edu.

DOI:10.1523/JNEUROSCI.0983-13.2013

Copyright © 2013 the authors 0270-6474/13/3310235-08\$15.00/0

2007; Bell et al., 2009; Mahon et al., 2009; Konkle and Oliva, 2012b), raising a challenge for how even this simple two-dimensional representational space of objects maps onto the cortical sheet.

Materials and Methods

Participants. Fifteen healthy observers with normal or corrected-to-normal vision participated in a 2 h fMRI session (age, 18–40 years; nine females). Informed consent was obtained according to procedures approved by the Institutional Review Board at the University of Trento.

MRI acquisition. Imaging data were acquired on a BioSpin MedSpec 4T scanner (Bruker) using an eight-channel head coil. Functional data were collected using an echo-planar 2D imaging sequence (TR, 2000 ms; TE, 33 ms; flip angle, 73°; slice thickness, 3 mm; gap, 0.99 mm, with 3×3 in-plane resolution). Volumes were acquired in the axial plane parallel to the anteroposterior commissure in 34 slices, with ascending interleaved slice acquisition.

Stimuli. The stimulus set consisted of 240 unique images of big animals, small animals, big objects, and small objects (60 images per condition). These items were selected to have broad coverage over the categories. For animals, the selection was guided by the 19 orders of animals as well as the animal categorization scheme of Troyer et al. (1997). For objects, we selected human-made items from various inanimate object categories (e.g., furniture, tools, vehicles, clothing items, kitchenware, appliances, office supplies, etc.), which were devoid of brand labels and text. For real-world size considerations, all small objects and small animals could be held easily with one or two hands; all big objects could support a human (e.g., chair-sized and bigger), with big animals selected similarly. The complete image set is available for download on T.K.'s website.

Task. Observers were shown images of big animals, small animals, big objects, and small objects in a standard blocked design while undergoing functional neuroimaging. Each block was 16 s long, in which 16 images were shown for 800 ms each with a 200 ms blank, presented in isolation on a white background at $\sim 8^\circ \times \sim 8^\circ$ visual angle. Ten-second fixation periods intervened between each block. Each run had four blocks per condition (213 volumes), with six total runs yielding 24 blocks per condition. All 60 images for each condition were presented once per run (four blocks of 15 unique images). Observers were instructed to pay attention to each item and to press a button when an exact image repeated back to back, which occurred once per block. The category localizer followed the same blocked design, with face, body, scene, object, and scrambled images presented. Each localizer run had three blocks per condition, per run (200 volumes), and observers completed two runs.

Data analysis. Functional data were analyzed using Brain Voyager QX software and MATLAB. Preprocessing included slice scan-time correction, 3D motion correction, linear trend removal, temporal high-pass filtering (0.01 Hz cutoff), spatial smoothing (6 mm FWHM kernel), and transformation into Talairach (TAL) coordinates. General linear model analyses included square-wave regressors for each condition's presentation times, convolved with a gamma function to approximate the hemodynamic response. Whole-brain, random-effects group analyses were conducted with contrast t maps thresholded at $p < 0.001$ (FDR < 0.03). For analyses involving correlation (r) values, all averaging and statistics were computed over Fisher- z transformed r values.

Vector-of-ROI analysis. We designed a vector-of-ROI analysis to enable a comparison of the response magnitudes for a number of conditions along a single-dimensional path along the cortex (Fig. 1). To define the semicircular vector of ROIs, we first defined seven "spoke" vectors

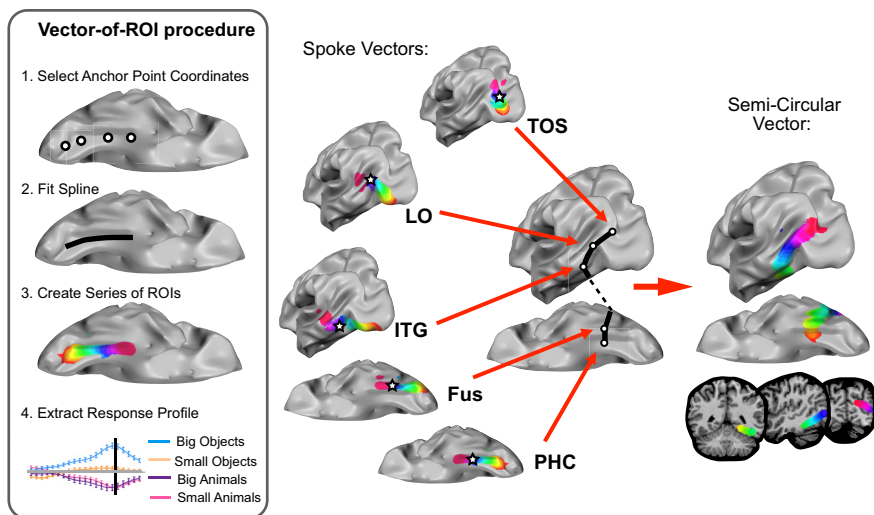


Figure 1. Vector-of-ROI schematic. Left, The procedure to define a vector of ROIs is to (1) specify a series of anchor points along a cortical path of interest (e.g., TAL coordinates along the lingual/parahippocampal gyrus from posterior to anterior), (2) fit a spline through these anchor points, (3) define a series of evenly spaced anatomical spherical ROIs along this spline, and (4) compute the response strength for all conditions in each ROI. Right, To create a semicircular vector of ROIs, spoke vectors were defined along ventral and lateral surfaces, across parahippocampal cortex (PHC), fusiform gyrus (Fus), inferior temporal gyrus (ITG), lateral occipital cortex (LO), and TOS. Data from half of the runs were used to find the maximal difference along the anteroposterior direction, and the centers of the peak ROIs along each spoke were used as new anchor points to define a semicircular vector of ROIs. The spherical ROIs were defined in the volume and assigned a color following a color gradient. When projected onto the inflated surface for visualization, these ROIs appear as a relatively continuous color gradient over the band of cortex captured by the vector of ROIs.

along occipito-temporal cortex emanating from the posterior occipital pole. To define each spoke, we (1) selected a series of anchor points that step along the cortical surface, (2) fit a spline through the anchor points of each spoke, (3) defined a series of 5 mm spherical ROIs spaced 3 mm apart along this spline, and (4) computed the response for each condition in each ROI. These spokes were along the parahippocampal gyrus, the fusiform gyrus, the inferior temporal gyrus, the lateral occipital cortex, and the medial occipital cortex toward the transverse occipital sulcus (TOS), as well as two more extreme medial spokes, one along the ventral surface and one on the lateral surface (to span from the most medial aspect of the parahippocampal gyrus and the most medial aspect of the transverse occipital sulcus). Next, we selected the ROI along each spoke with the largest differential response across the four stimulus conditions, based on data from the even runs. Finally, the center TAL coordinates of these peak ROIs along each spoke were used as anchor points to define a new "semicircular" spline along the medial-to-lateral-to-medial axis. We defined a series of 5 mm spherical ROIs spaced 3 mm apart along this spline, and within this vector-of-ROIs, β weights were extracted from a GLM over data from the odd runs. All analyses of the response patterns across this band of cortex used data that were independent of those used to select the positions of the vectors-of-ROIs.

Related analyses of the response topography typically plot activity as a function of the x , y , or z coordinate in TAL space (Mahon et al., 2009). However, these analyses are limited to the cardinal axes, which do not always walk along the anatomical path of interest, and typically select one voxel along the path even when plotting a larger regional profile would be preferable. Thus, the vector-of-ROI method can be a valuable analysis method for uncovering the large-scale structure of multiple conditions along a single anatomical axis of interest e.g., along the occipito-temporal cortex, intraparietal sulcus, superior temporal sulcus, etc. After extracting the responses from each ROI along the vector, each condition has a single-dimensional pattern of activity, and all multivoxel pattern analysis techniques can be applied (Haxby et al., 2000).

Preference maps. To compute the two-way preference maps, an object-responsive mask was computed from the contrasts of all $>$ rest with $T > 2.0$, from a group fixed-effects GLM. Next, the t map of all animals versus all objects (or all big vs all small) was multiplied by this mask and dis-

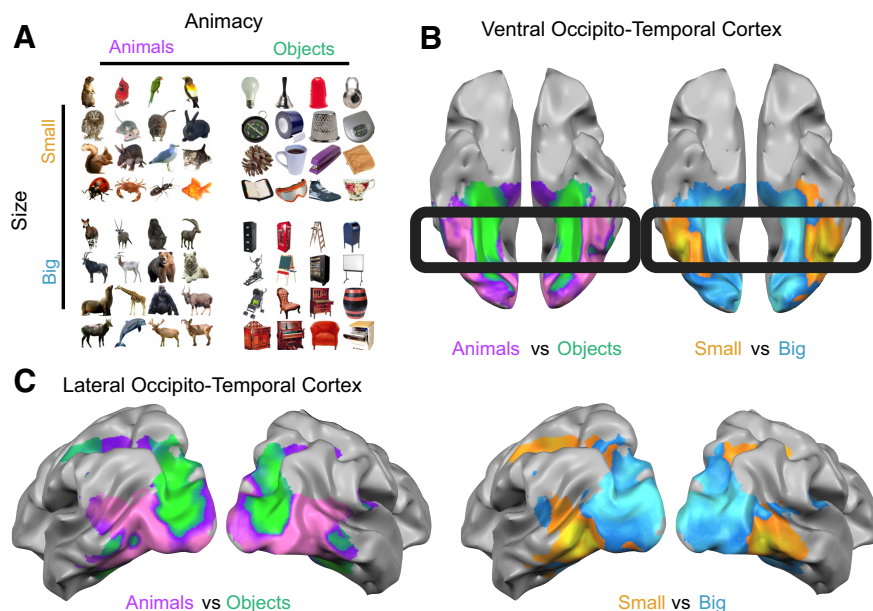


Figure 2. Stimulus conditions and main effects. **A**, Example images from each of the four conditions (animacy \times size). **B**, **C**, Preference maps of the animacy and size dimensions separately. Voxels with stronger preference for animals than objects are shown in purple, and voxels with a stronger preference for objects than animals are shown in green. Similarly, voxels with a preference for small things are shown in orange and for big things are shown in blue. **B**, Ventral occipito-temporal cortex view of animals versus objects (left) and of small versus big sizes (right). **C**, Lateral occipito-temporal cortex view of animacy organization (right) and size organization (left).

played on the cortical surface, and two colors were used to show the locations of voxels with a preference for either of the conditions in the two-way contrast (see Konkle and Oliva, 2012b). For the three-way preference maps, each voxel in the object-responsive mask was colored by the peak condition, and the peak strength of each voxel was computed as the peak β minus the mean of the remaining β s. To visualize the arrangement of the preference zones more clearly, we set an arbitrary lower threshold as peak β strength >0.15 (group data) and peak β strength >0.3 (single-subject data). This visualization choice serves to draw attention to regions of cortex with large differential responses and separates the zones by excluding the less differential responses; this threshold cannot change the spatial arrangement of these zones. This map analysis is a variant of a winner-analysis used in other high-level visual mapping studies (Orlov et al., 2010).

Area under the curve analysis of the category-selective overlap with preference zones. To characterize how the locations of category-selective regions are spatially organized with respect to the animacy-size preference zones, there are two main challenges. First, the classic category-selective ROIs are defined not only by their selectivity but also by their general anatomical position, typically selected manually by an experimenter. Thus, any uncertainty about which regions correspond to the classic category-selective ROIs are subject to experimenter bias and may be selected toward or away from the location of a particular preference zones. Second, the size of a category-selective ROI is subject to a statistical threshold, which is arbitrarily dependent on power, and the extent of the ROI is often constrained by an arbitrary radius, e.g., only voxels within an 8 mm radius sphere around the peak voxel. Thus, to quantify the relationship between category-selective voxels and preference zone in a way that was not subject to either experimenter bias in selection or statistical threshold, we used a receiver operating characteristic (ROC) analysis.

First, a face-selective contrast was computed as Faces $>$ [Bodies Scenes Objects]. All voxels within an object-responsive mask (the same used for the preference map analysis) were sorted by their t value. For each step, we considered the topmost face-selective t values (from percentiles of 1% to 100%). No constraints by spatial contiguity or anatomical location were used, so these topmost selective voxels were not restricted to fall in classic category-selective regions. For these voxels, we computed the proportion that fell in the target zone (e.g., animal zone) and the proportion

that fell into the nontarget zones (e.g., either the big object or small object zone). The ROC curve plots the proportion of the target zone filled relative to the nontarget zone filled, as an increasing number of face-selective voxels are considered. If the topmost face-selective voxels completely filled all of the animate zone before either filling the small object or big object zones, then the ROC curve would rise dramatically and stay at 100% [perfect precision and sensitivity, with an area under the curve (AUC) equal to 1]. If the face-selective voxels were distributed randomly with respect to the preference zones, then the expected proportion would fall along the diagonal line (chance, AUC = 0.5). If a curve falls below the chance diagonal, this means that, for example, the face-selective voxels fill the big-object zones less than expected by chance. ROC curves were computed for each category-selective contrast (for faces, bodies, and scenes) considering each of the preference zones as the target zones (small object zone, big object zone, animal zone).

We additionally defined the occipital face area (OFA), fusiform face area (FFA), extrastriate body area (EBA), fusiform body area (FBA), scene-selective TOS, and parahippocampal place area (PPA), in each participant, based on the appropriate category-selective contrast (e.g., for face-selective regions: Faces $>$ [Bodies Scenes Objects]). All

ROIs were defined by identifying the peak category-selective voxel within 20 mm of approximate coordinates of each target region derived from a meta-analysis. All significantly active voxels (FDR $<$ 0.05) within an 8 mm radius sphere of that peak voxel were defined as the target ROI. To assess the location of these ROIs with respect to the animacy-size preference zones, we computed the percentage of voxels within each ROI that fell in each zone, for each participant and each ROI, and we tested the deviation from chance using χ^2 tests, where chance was set by the relative size of the zones for each participant.

Results

We first examined the large-scale neural organization of animacy and real-world size separately, by comparing animals versus objects, collapsing across size, and by comparing big versus small entities, collapsing across animacy (Fig. 2A). In other words, how does each of these distinctions lead to a large-scale grouping of response preferences along the cortical surface? To visualize the spatial distribution of animal/object responses and small/big responses, we computed two-way preference maps (Konkle and Oliva, 2012b) (see Materials and Methods). For each voxel within a visually responsive mask, the preferred (or “winner”) condition is plotted based on the contrast comparing the two conditions. The results of the animacy-preference map and size-preference maps are shown in Figure 2.

Along the ventral surface (Fig. 2B), we observed a spatial organization of responses that is remarkably similar for both animacy and size. Specifically, object responses are adjacent to animal responses along the medial-to-lateral axis; similarly, big responses are adjacent to small responses along the same medial-to-lateral axis along the ventral surface. This result replicates previous findings and illustrates a conundrum to be reconciled: How do these two dimensions map together across this surface?

The preference-map analysis also revealed a large-scale organization of responses not only on the ventral surface but also along the lateral surface of the occipito-temporal cortex. Animal-

to-object responses wrapped from lateral-to-medial from the middle temporal gyrus to the transverse occipital sulcus (Fig. 2C, left). Similarly, small-to-big responses wrapped from lateral-to-medial as well (Fig. 2C, right). These results show that whereas the organization of the ventral surface has been the focus of most animate/inanimate mapping work (Chao et al., 1999b; Downing et al., 2006; Mahon et al., 2009), both animacy and size dimensions are consistently part of an even larger-scale organization, with alternating peaks of selectivity across both the ventral and lateral surface of occipito-temporal cortex.

Vector-of-ROI analysis

To determine how both dimensions mapped together along the ventral and lateral surface of occipito-temporal cortex, we developed a vector-of-ROI analysis procedure (Fig. 1; see Materials and Methods). A series of partially overlapping spherical regions of interest were defined along the main axis of organization. These ROIs formed a semicircular sweep across ventral occipito-temporal cortex continuing across lateral occipito-temporal cortex, encompassing the object responsive cortex just beyond early retinotopic areas [see Hasson et al. (2003) and Op de Beeck et al. (2008) for different visualizations of this band of cortex]. Within each ROI, the overall activation for each condition was estimated, and the grand mean activation across all conditions was subtracted from each condition, analogous to the procedure in more typical two-condition contrasts. This enabled a clear visualization of the relative differences in activation across the cortex for all four conditions. These data are shown in Figure 3.

The first key result is that the size distinction applies primarily within the object domain. This can be observed by comparing the response profiles across the four conditions: the responses to big animals and small animals (Fig. 3, purple/pink lines) are very similar, whereas the response profiles for big objects and small objects (Fig. 3, blue/orange lines) are more different. To quantify this, we took the unnormalized β weights across this vector-of-ROIs to be a large-scale pattern (rather than a “multivoxel pattern”) and computed correlations between pairs of conditions following standard pattern analysis methods. This analysis confirmed that the similarity between small and big animals was significantly greater than the similarity between small and big objects: $t_{(14)} = 10.39$, $p < 0.001$ (pairwise large-scale pattern correlations: SmallAnimals-BigAnimals: $r = 0.98$; SEM, 0.01; SmallObjects-BigObjects: $r = 0.89$; SEM, 0.02; SmallAnimals-SmallObjects: $r = 0.88$; SEM, 0.06; BigAnimals-SmallObjects: $r = 0.85$; SEM, 0.07; SmallAnimals-BigObjects: $r = 0.67$; SEM, 0.07; BigAnimals-BigObjects: $r = 0.64$; SEM, 0.08).

The second key result is highlighted by considering the peaks of activity across this band of cortex (Fig. 3). Focusing on the conditions that have the highest relative β , there are five distinct peaks. At the most medial extremes, both the parahippocampal cortex (ventromedial) and transverse occipital sulcus regions (dorsomedial) have a strong preference for big objects (Fig. 3, blue line). Adjacent to these peaks, in the fusiform and lateral occipital cortex, responses show an animacy preference (Fig. 3, pink/purple lines), with no difference between big and small animals responses. Finally, at the center of the map, there is a region

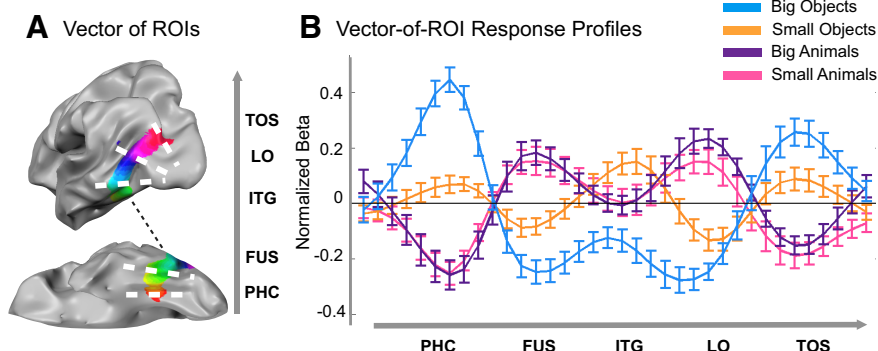


Figure 3. Vector of ROIs. **A**, A series of partially overlapping spherical ROIs were defined in the volume and are shown projected on an inflated cortical surface. Labels denote approximate anatomical positions and not functionally defined regions: parahippocampal cortex (PHC), fusiform gyrus (Fus), inferior temporal gyrus (ITG), lateral occipital cortex (LO), and TOS. **B**, Relative activation is plotted for each condition as a function of position along the cortex. The y-axis shows the normalized β s, and the x-axis indicates the position of the ROI across the cortex from medioventral, to lateral, to mediadorsal. Error bars reflect ± 1 SEM across subjects.

of cortex with a preference for small objects around the inferior temporal gyrus (Fig. 3, orange line). These five peaks form a mirror-symmetric organization across this large band of cortex and have an approximate period of 36 mm/cycle (as estimated based on the distance between the ROI centers in the volume). Thus, overall we observed that there is a set of large-scale “zones” of response preferences that tile occipito-temporal cortex in a clear mirrored macro-organization.

Although we subsequently refer to each preference zone by its peak condition (big objects, animals, small objects), it is important to note that there is systematic structure in the nonpreferred conditions (Fig. 3). In the big object zones, for example, small objects are the next most active condition (both object conditions have higher responses than the animal conditions), whereas in the small object zone, animals are the next most active condition (splitting the animate/inanimate boundary). Thus, the zone labels should be taken as a guide to territories with distinctive response profiles, rather than as a strong statement about fully modular divisions of object cortex.

Cortical arrangement of preference zones

One consequence of the interaction between animacy and size dimensions is that no two-way contrast (e.g., animals vs objects or big vs small entities) can capture the underlying organization. Thus, to visualize where these peak zones are on the cortex, we computed a three-way preference map, in which each voxel was colored by the peak condition (big objects, all animals, small objects) within an object-responsive mask (see Materials and Methods). These three-way winner maps are shown both for the group (Fig. 4A) and for three individual subjects (Fig. 4B). To highlight the geometric arrangement of the preference zones, we restricted the voxels to those with the strongest differential response across conditions for visualization purposes (arbitrary β differential > 0.3 , e.g., computed as the peak β minus the mean of the remaining conditions β s).

These results reveal that the preference zones have large contiguous expanses that have a systematic preference for big objects, animals, or small objects. These preference zones have an apparent spoked organization around the occipital pole (Fig. 4A, posterior view) and maintain their response preferences along the posterior-to-anterior axis. Furthermore, these zones have a mirrored organization, with small objects at the center, surrounded by animal zones on either side, surrounded by big object zones.

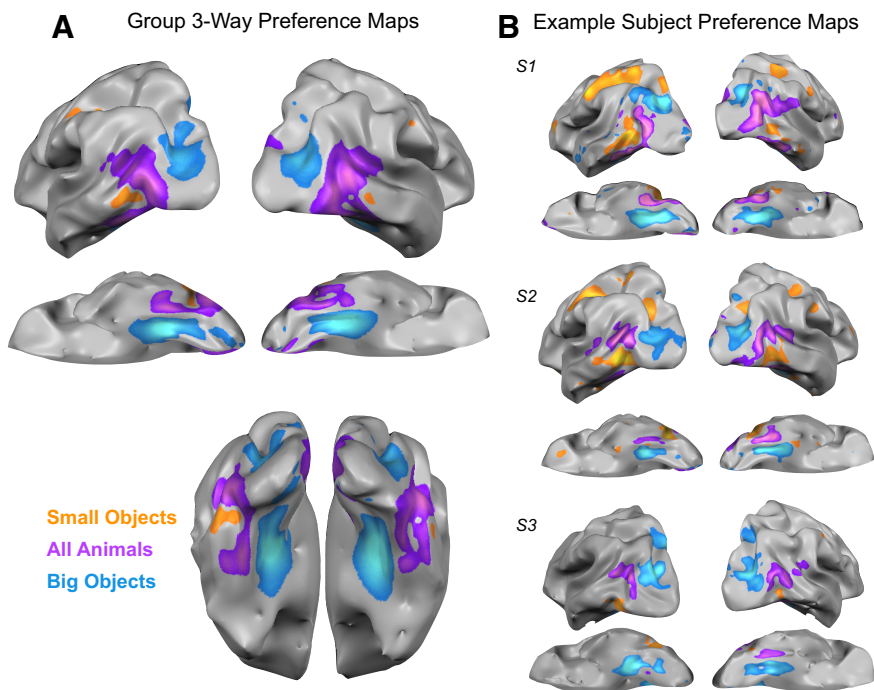


Figure 4. Three-way preference maps. **A**, Three-way preference maps from the group data. Blue zones show regions with preferences for big objects, purple zones show regions with preferences for all animals (collapsing across size), and orange zones show regions with preferences for small objects. The ventral and lateral views of both hemispheres are shown above. The posteroventral view shown below highlights the spoked organization around the occipital pole. **B**, Maps of three example subjects.

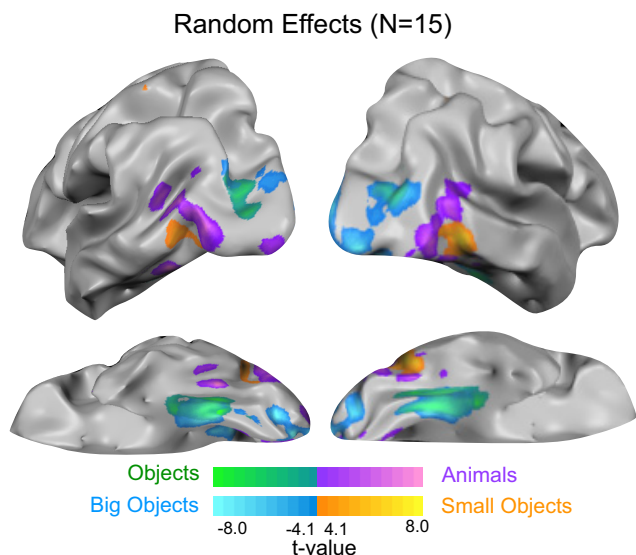


Figure 5. Whole-brain random-effects analysis. Whole-brain random-effects analysis of two targeted contrasts. Blue, Big objects; orange, small objects; purple, all animals; green, all objects. The *t* maps are semitransparent, highlighting that the medial object responses (green) and big object responses (blue) are similarly located. These whole-brain contrasts reveal that, across participants, the preference zones have a reliable geometric layout and robust differential responses.

These zones are also apparent in a random-effects, whole-brain analysis. Two-way contrasts targeting the tripartite division show the animal zones, based on the animals versus objects contrasts, with interleaved inanimate zones separated by size, based on the big object versus small object contrasts (Fig. 5). This analysis demonstrates the relatively robust response differences across these categories in occipito-temporal cortex

and confirms the general reliability of the location of these zones across participants.

In some individuals, viewing objects elicited responses along the various parts of the intraparietal sulcus (Fig. 4*B*; see also Fig. 7). However, these responses were weak in magnitude and were not reliably organized across participants in the present data (e.g., no intraparietal sulcus regions were revealed in the random-effects analysis). Further consideration of these dorsal stream responses, and how they relate to the mirrored organization of the ventral stream, will require a different task that drives stronger and more reliable activation patterns along the dorsal stream.

Mesomap structure: face, body, and scene ROIs

Within the occipito-temporal cortex, there is a replicable mosaic of category-selective regions for faces, bodies, and scenes (Kanwisher et al., 1997; Epstein and Kanwisher, 1998; Downing et al., 2001). Whereas these regions have typically been characterized separately, comparison of their anatomical positions and response properties provides another clue to their underlying functions (Taylor and Downing, 2011; Weiner and Grill-Spector, 2013). To this end, we next mapped the locations of these highly selective regions within this larger-scale organization of animal and object responses. Situating this macro- and meso-scale structure together is important for constructing an integrated schema of object organization along the ventral stream.

Given that face- and body-selective regions show high responses to both humans and nonhumans (Chao et al., 1999a; Tong et al., 2000; Connolly et al., 2012; Looser et al., 2013), we expect that face- and body-selective regions would overlap with animal zones. Similarly, given that scene-selective regions also show high responses to big objects (Aguirre et al., 1998; Mullally and Maguire, 2011; Harel et al., 2013; Konkle and Oliva, 2012b), we expect these regions to overlap with big object zones.

A receiver operating characteristic analysis confirmed these predictions: across all subjects, face- and body-selective voxels fell predominantly within the animal zones, whereas scene-selective regions fell within the big object zones (AUCs significantly greater than chance, all $t > 10.0$; all $p < 0.001$; Fig. 6; see Materials and Methods). In all participants, body-selective voxels that were not in the animal zones were more likely to fall within the small object zone than the big object zone ($t_{(14)} = 10.40$; $p < 0.001$). The mean AUC and SE across participants for each category and target zone are as follows: face-selective: animal zone AUC = 0.73 (SEM = 0.02), small object zone AUC = 0.46 (SEM = 0.02), big object zone AUC = 0.28 (SEM = 0.02); body-selective: animal zone AUC = 0.66 (SEM = 0.02), small object zone AUC = 0.55 (SEM = 0.02), big object zone AUC = 0.27 (SEM = 0.02); scene-selective: big object zone AUC = 0.80 (SEM = 0.01), small object zone AUC = 0.39 (SEM = 0.03), animal zone AUC = 0.30 (SEM = 0.02).

We additionally defined these ROIs for each participant following traditional procedures and quantified the overlap be-

tween the ROIs and each of the preference zones. Figure 7 shows the locations of the category-selective regions for a single participant, as well as the locations with respect to their animacy-size zones. The ROI analysis confirmed the ROC analysis [FFA/OFA: animal zone = 97% (SEM = 1.0), small object zone = 2.7% (SEM = 0.9), big object zone = 0.6% (SEM = 0.3); EBA/FBA: animal zone = 83% (SEM = 2.9), small object zone = 15% (SEM = 2.5), big object zone = 1.8% (SEM = 1.1); PPA/TOS: big object zone = 94% (SEM = 1.8), small object zone = 3.3% (SEM = 1.4), animal zone = 3% (SEM = 1.3); χ^2 tests, $p < 0.001$ in all participants and all face, body, and scene conditions].

Importantly, the category-selective mosaic alone (Fig. 7A) does not easily predict the anatomical arrangement and elongated shape of the animacy-size zones (Fig. 7B), nor does it predict the functional organization that inanimate objects are separated by size, whereas animals are not. Together, however, this analysis suggests that these broader and narrower distinctions among objects are similarly reflected by larger and smaller cortical parcellations: at a macro-scale, there are large cortical territories with differential responses along the core dimensions of animacy and object size; at a meso-scale, there is further organization within these territories and domains, where faces, bodies, and scenes, have highly selective responses that are meaningfully related to the response preferences in surrounding cortex.

Discussion

The aim of the current study was to uncover the neural organization of object responses arising from two major dimensions of object representation, animacy and real-world size. We developed a new vector-of-ROI analysis to reveal the structure of responses along the major axis of variation and visualized the geometric arrangement of these zones on the cortical surface. Considering animacy and size together revealed a mirrored macro-organization of object responses across the entire occipito-temporal cortex. These factors do not map in a two-dimensional arrangement across the cortex but instead show an interleaved organization along a single ventromedial, to lateral, to dorsomedial axis. Real-world size drives differential responses only in the object domain, not the animate domain, yielding a tripartite distinction in the space of object representation. Finally, there is a duplication of response selectivities along the ventral and lateral surface, suggesting a major division of labor separates object processing into two substreams of the ventral visual pathway.

Inferences from spatial topography

What do these results reveal about the nature of object representations within these zones of cortex? One window into this question is to consider the structure of the responses in the nonpreferred conditions. In the small object zone, animals were

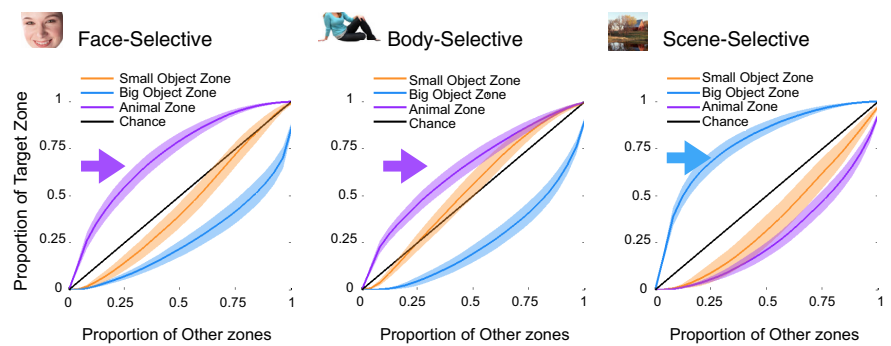


Figure 6. AUC analysis of the category-selective overlap with the preference zones. This analysis was conducted for faces (left), bodies (middle), and scenes (right). Shown are receiver operating characteristic curves, which show how each of the preference zones fill as an increasing number of voxels are included, starting from the most category selective. Purple arrows on the face and body subplots highlight that the face- and body-selective fall predominantly within the animal zones; the blue arrow on the scene subplot highlights that the scene-selective voxels fall predominantly within the big object zones. The shaded area around each line reflects ± 1 SEM across subjects.

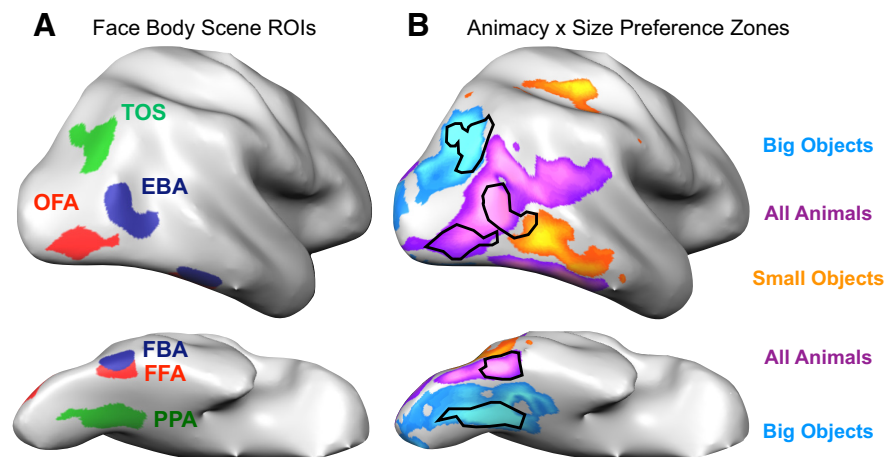


Figure 7. Comparison with classic category-selective areas. **A**, Face-, scene-, and body-selective ROIs for an example subject are highlighted: face-selective areas are in red and include the OFA and FFA. Body-selective areas are in blue and include the EBA and FBA. Scene-selective areas are in green and include the TOS and PPA. **B**, The three-way preference map for this participant is shown, with big object zones in blue, animate zones in purple, and small object zones in orange. Black outlines show the locations of the category-selective regions from **A**.

the next most active condition, not big objects; thus, this region does not have a strong animate/inanimate divide. Relatedly, in the category-selective analysis, body-selective voxels also partially overlapped with this small-object zone. Thus, one possibility is that the small object zone is important for coordinating object-agent interactions information (Beauchamp et al., 2002; Bracci et al., 2012). In contrast, the big object zones do preserve the animate/inanimate divide, with both big and small objects driving stronger responses than animals. Thus, the nature of the information here is likely importantly different from the computations in the small object zone and may reflect how well the object defines a space or marks a navigational junction (Janzen and van Turenout, 2004; Epstein, 2008; Mullally and Maguire, 2011). Critically, to examine the nature of the representations following this approach, future work is required to measure the responses to a number of conditions to triangulate what combination of factors best account for the response profile in targeted regions of cortex (Mullally and Maguire, 2011; see also Huth et al., 2012).

As a complementary approach, we suggest that the spatial organization itself can also be an informative window into the underlying representational structure. On the assumption that

similar representations are near each other on the cortical sheet (Jacobs and Jordan, 1992; Kaas and Catania, 2002; Rosa and Tweedale, 2005; Graziano and Aflalo, 2007), we can make inferences about the underlying representational space from the spatial arrangement of responses (Aflalo and Graziano, 2011). Formally, cortical mapping can be operationalized as a form of dimensionality reduction: a simple two-dimensional space can project directly along the two-dimensional cortical sheet; a higher-dimensional space will require a more complex mapping, and thus the ultimate projection depends on the weight and number of factors in the underlying representational space (Kohonen, 1982; Durbin and Mitchison, 1990). For example, within this framework, a duplication of a response selectivity happens for a reason: there is some other dimension along which the responses or processing demands differ, implying another factor in the representational space (Aflalo and Graziano, 2011). Within this framework, we can consider what the tripartite distinction and the duplication observed here imply about the structure of the representational space of objects.

Size applies only within the inanimate domain

We found that size is not a major factor distinguishing the neural response profile of different animals, whereas objects showed a large-scale separation by size. Within a dimension-reduction framework, this tripartite organization suggests that the processing of inanimate object information shows a division of labor based on real-world size, whereas the processing of animal information does not. Why might this be the case? In other words, what key properties of small objects, big objects, and animals are distinctive, and what neural mechanism might underlie this tripartite division?

One possibility is that this tripartite distinction reflects different functional behavioral roles: the size of an inanimate object causally influences how we interact with it (with our hands or whole body), whereas for animals, our primary interactions are not related to real-world size. For example, the danger posed by an animal is not necessarily size-related, nor is the task of inferring intentions and goals. On this functional account, cortical organization may be driven by distinct long-range connections to downstream processes (Mahon and Caramazza, 2011), for example, connecting big object zones to navigational networks (Epstein, 2008), small object zones to dorsal stream reaching regions (Valyear et al., 2007; Bracci et al., 2012), and animate zones to goal-inference or other social regions (Caramazza and Shelton, 1998; Frith, 2007). Alternatively, one could interpret these divisions in terms of the statistical structure in visual/shape properties of these categories. To illustrate with a simplified example, big objects may be more rectilinear, small objects may be more rounded, and animals may have distinct part-relationships that are similar for big and small animals but are distinct from both small and large objects. Such form-based representations may emerge via experience-dependent tuning mechanisms that detect such shape regularities and drive the functional clustering of object responses (Kohonen, 1982; Polk and Farah, 1995; Srihasam et al., 2012).

It is important to recognize these accounts are not mutually exclusive. For example, the ventral surface may be related to form processing, whereas the lateral surface may be related to functional processing (Martin, 2007). Alternatively, although form and function can be intuitively dissociated, they may not be directly related to the major joints in the neural architecture. If form is intrinsically correlated with function, then the brain might naturally leverage the covariation between these properties, such that both visual and functional features are jointly responsible for selectivity along each zone. One interesting

possibility is that these are differentially weighted from posterior to anterior, with emphasis on visual/shape feature in posterior occipital cortex, driven by local organizing mechanisms, channeling to more abstract functional features in anterior cortex, determined by long-range network architecture.

Division of the “what”-pathway

The distinction between animate and inanimate objects is repeatedly identified as a strong predictor of variance in neural similarity structure (Kiani et al., 2007; Kriegeskorte et al., 2008; Haxby et al., 2011; Huth et al., 2012). However, even though it is the strongest modulator of responses, the present results show that it is not the largest grouping factor. It could have been the case that all animal responses were grouped along the ventral surface and all object responses were grouped along the lateral surface (perhaps segregated by size). Instead, we see a clear duplication of response selectivity, with a set of regions along the lateral surface and a matching set of regions along the ventral surface, namely, a major division within the ventral stream. Within a dimensionality-reduction framework, this duplication suggests that the proximity of a different relationship is being maximized on the lateral and ventral surfaces.

To gain insight into this major division, we can examine how the ventral and lateral pathways differ in processing object information. The entire lateral surface is relatively more sensitive to human and object motion and may thus be a pathway for coordinating interaction with objects and agents in the world (Beauchamp et al., 2002). In a review comparing the response properties of the paired category-selective regions, the lateral surface regions (OFA, EBA, TOS) tend to be more “primitive” and part based, whereas the ventral surface regions (FFA, FBA, PPA) tend to be more integrated and invariant to visual transformations (Taylor and Downing, 2011). Finally, the first observation of the duplication of category-selective areas proposed that this organization is inherited from adjacent retinotopic cortex, with the lateral surface extending from lower visual field representations and the ventral surface extending from upper visual field representations (Hasson et al., 2003; see also Kravitz et al., 2013). These results set the stage for future research to uncover the different computational goals subserved by these lateral and ventral substreams.

Conclusion

The large-scale organization shown here raises questions about the nature of object responses across occipito-temporal cortex, the computational and behavioral goals supported by these regions, and the roles of experience and network architecture in driving the spatial organization. We suggest that the topography of responses can be informative to the structure of the representational space, where large-scale neural divisions are meaningfully related to core factors in the underlying representational space. Broadly, we suggest that the response properties and the computational goals of object-responsive cortex can be meaningfully described at multiple spatial scales and that doing so will enable a deeper understanding of the principles of object representation.

References

- Aflalo TN, Graziano MS (2011) Organization of the macaque extrastriate visual cortex re-examined using the principle of spatial continuity of function. *J Neurophysiol* 105:305–320. [CrossRef Medline](#)
- Aguirre GK, Zarahn E, D’Esposito M (1998) An area within human ventral cortex sensitive to “building” stimuli: evidence and implications. *Neuron* 21:373–383. [CrossRef Medline](#)
- Beauchamp MS, Lee KE, Haxby JV, Martin A (2002) Parallel visual motion

- processing streams for manipulable objects and human movements. *Neuron* 34:149–159. [CrossRef Medline](#)
- Bell AH, Hadj-Bouziane F, Frihauf JB, Tootell RB, Ungerleider LG (2009) Object representations in the temporal cortex of monkeys and humans as revealed by functional magnetic resonance imaging. *J Neurophysiol* 101:688–700. [Medline](#)
- Bracci S, Cavina-Pratesi C, Ietswaart M, Caramazza A, Peelen MV (2012) Closely overlapping responses to tools and hands in left lateral occipito-temporal cortex. *J Neurophysiol* 107:1443–1456. [CrossRef Medline](#)
- Caramazza A, Shelton JR (1998) Domain-specific knowledge systems in the brain: the animate-inanimate distinction. *J Cogn Neurosci* 10:1–34. [CrossRef Medline](#)
- Chao LL, Martin A, Haxby JV (1999a) Are face-responsive regions selective only for faces? *Neuroreport* 10:2945–2950. [CrossRef Medline](#)
- Chao LL, Haxby JV, Martin A (1999b) Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nat Neurosci* 2:913–919. [CrossRef Medline](#)
- Connolly AC, Guntupalli JS, Gors J, Hanke M, Halchenko YO, Wu YC, Abdi H, Haxby JV (2012) Representation of biological classes in the human brain. *J Neurosci* 32:2608–2618. [CrossRef Medline](#)
- Downing PE, Jiang Y, Shuman M, Kanwisher N (2001) A cortical area selective for visual processing of the human body. *Science* 293:2470–2473. [CrossRef Medline](#)
- Downing PE, Chan AW, Peelen MV, Dodds CM, Kanwisher N (2006) Domain specificity in visual cortex. *Cereb Cortex* 16:1453–1461. [CrossRef Medline](#)
- Durbin R, Mitchison G (1990) A dimension reduction framework for understanding cortical maps. *Nature* 343:644–647. [CrossRef Medline](#)
- Epstein RA (2008) Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends Cogn Sci* 12:388–396. [CrossRef Medline](#)
- Epstein R, Kanwisher N (1998) A cortical representation of the local visual environment. *Nature* 392:598–601. [CrossRef Medline](#)
- Frith CD (2007) The social brain? *Philos Trans R Soc Lond B Biol Sci* 362:671–678. [CrossRef Medline](#)
- Goodale MA, Milner AD (1992) Separate visual pathways for perception and action. *Trends Neurosci* 15:20–25. [CrossRef Medline](#)
- Graziano MS, Aflalo TN (2007) Mapping behavioral repertoire onto the cortex. *Neuron* 56:239–251. [CrossRef Medline](#)
- Harel A, Kravitz DJ, Baker CI (2013) Deconstructing visual scenes in cortex: gradients of object and spatial layout information. *Cereb Cortex* 23:947–957. [CrossRef Medline](#)
- Hasson U, Harel M, Levy I, Malach R (2003) Large-scale mirror-symmetry organization of human occipito-temporal object areas. *Neuron* 37:1027–1041. [CrossRef Medline](#)
- Haxby JV, Hoffman EA, Gobbini MI (2000) The distributed human neural system for face perception. *Trends Cogn Sci* 4:223–233. [CrossRef Medline](#)
- Haxby JV, Guntupalli JS, Connolly AC, Halchenko YO, Conroy BR, Gobbini MI, Hanke M, Ramadge PJ (2011) A common, high-dimensional model of the representational space in human ventral temporal cortex. *Neuron* 72:404–416. [CrossRef Medline](#)
- Huth AG, Nishimoto S, Vu AT, Gallant JL (2012) A continuous semantic space describes the representation of thousands of object and action categories across the human brain. *Neuron* 76:1210–1224. [CrossRef Medline](#)
- Jacobs RA, Jordan MI (1992) Computational consequences of a bias toward short connections. *J Cogn Neurosci* 4:323–336. [CrossRef](#)
- Janzen G, van Turenhout M (2004) Selective neural representation of objects relevant for navigation. *Nat Neurosci* 7:673–677. [CrossRef Medline](#)
- Kaas JH, Catania KC (2002) How do features of sensory representations develop? *Bioessays* 24:334–343. [CrossRef Medline](#)
- Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci* 17:4302–4311. [Medline](#)
- Kiani R, Esteky H, Mirpour K, Tanaka K (2007) Object category structure in response patterns of neuronal population in monkey inferior temporal cortex. *J Neurophysiol* 97:4296–4309. [CrossRef Medline](#)
- Kohonen T (1982) Self-organized formation of topologically correct feature maps. *Biol Cybern* 43:59–69. [CrossRef](#)
- Konkle T, Oliva A (2011) Canonical visual size for real-world objects. *J Exp Psychol Hum Percept Perform* 7:23–37. [CrossRef Medline](#)
- Konkle T, Oliva A (2012a) A familiar-size Stroop effect: real-world size is an automatic property of object representation. *J Exp Psychol Hum Percept Perform* 38:561–569. [CrossRef Medline](#)
- Konkle T, Oliva A (2012b) A real-world size organization of object responses in occipitotemporal cortex. *Neuron* 74:1114–1124. [CrossRef Medline](#)
- Kourtzi Z, Connor CE (2011) Neural representations for object perception: structure, category, and adaptive coding. *Annu Rev Neurosci* 34:45–67. [CrossRef Medline](#)
- Kravitz DJ, Saleem KS, Baker CI, Ungerleider LG, Mishkin M (2013) The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends Cogn Sci* 17:26–49. [CrossRef Medline](#)
- Kriegeskorte N, Mur M, Ruff DA, Kiani R, Bodurka J, Esteky H, Tanaka K, Bandettini PA (2008) Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron* 60:1126–1141. [CrossRef Medline](#)
- Kuhlmeier VA, Bloom P, Wynn K (2004) Do 5-month-old infants see humans as material objects? *Cognition* 94:95–103. [CrossRef Medline](#)
- Looser CE, Guntupalli S, Wheatley T (2013) Multivoxel patterns in face-sensitive temporal regions reveal an encoding schema based on detecting life in a face. *Soc Cogn Affect Neurosci*. Advance online publication. Retrieved August 27, 2013. doi:10.1093/scan/nss078. [CrossRef Medline](#)
- Mahon BZ, Caramazza A (2009) Concepts and categories: a cognitive neuropsychological perspective. *Ann Rev Psychol* 60:27–51. [CrossRef Medline](#)
- Mahon BZ, Caramazza A (2011) What drives the organization of object knowledge in the brain? *Trends Cogn Sci* 15:97–103. [CrossRef Medline](#)
- Mahon BZ, Anzellotti S, Schwarzbach J, Zampini M, Caramazza A (2009) Category-specific organization in the human brain does not require visual experience. *Neuron* 63:397–405. [CrossRef Medline](#)
- Martin A (2007) The representation of object concepts in the brain. *Annu Rev Psychol* 58:25–45. [CrossRef Medline](#)
- Mullally SL, Maguire EA (2011) A new role for the parahippocampal cortex in representing space. *J Neurosci* 31:7441–7449. [CrossRef Medline](#)
- Op de Beeck HP, Haushofer J, Kanwisher NG (2008) Interpreting fMRI data: maps, modules and dimensions. *Nat Rev Neurosci* 9:123–135. [CrossRef Medline](#)
- Orlov T, Makin TR, Zohary E (2010) Topographic representation of the human body in the occipitotemporal cortex. *Neuron* 68:586–600. [CrossRef Medline](#)
- Polk TA, Farah MJ (1995) Brain localization for arbitrary stimulus categories: a simple account based on Hebbian learning. *Proc Natl Acad Sci U S A* 92:12370–12373. [CrossRef Medline](#)
- Rosa MG, Tweeddale R (2005) Brain maps, great and small: lessons from comparative studies of primate visual cortical organization. *Philos Trans R Soc Lond B Biol Sci* 360:665–691. [CrossRef Medline](#)
- Setti A, Caramelli N, Borghi AM (2009) Conceptual information about size of objects in nouns. *Eur J Cogn Psychol* 21:1022–1044. [CrossRef](#)
- Spelke ES, Phillips A, Woodward AL (1995) Infants' knowledge of object motion and human action. In: *Causal cognition: a multidisciplinary debate*, Chap 3 (Sperber D, Premack D, Premack AJ, eds), pp 44–78. Oxford: Oxford UP.
- Srihasam K, Mandeville JB, Morocz IA, Sullivan KJ, Livingstone MS (2012) Behavioral and anatomical consequences of early versus late symbol training in macaques. *Neuron* 73:608–619. [CrossRef Medline](#)
- Taylor JC, Downing PE (2011) Division of labor between lateral and ventral extrastriate representations of faces, bodies, and objects. *J Cogn Neurosci* 23:4122–4137. [CrossRef Medline](#)
- Tong F, Nakayama K, Moscovitch M, Weinrib O, Kanwisher N (2000) Response properties of the human fusiform face area. *Cogn Neuropsychol* 17:257–280. [CrossRef Medline](#)
- Troyer AK, Moscovitch M, Winocur G (1997) Clustering and switching as two components of verbal fluency: evidence from younger and older healthy adults. *Neuropsychology* 11:138–146. [CrossRef Medline](#)
- Ungerleider LG, Bell AH (2011) Uncovering the visual “alphabet”: advances in our understanding of object perception. *Vision Res* 51:782–799. [CrossRef Medline](#)
- Ungerleider LG, Mishkin M (1982) Two cortical visual systems. In: *Analysis of visual behavior* (Ingle DJ, Goodale MA, Mansfield RJW, eds), pp 549–586. Cambridge, MA: MIT.
- Valyear KF, Cavina-Pratesi C, Stiglick AJ, Culham JC (2007) Does tool-related fMRI activity within the intraparietal sulcus reflect the plan to grasp? *Neuroimage* 36 [Suppl 2]:T94–T108. [CrossRef Medline](#)

Weiner KS, Grill-Spector K (2013) Neural representations of faces and limbs neighbor in human high-level visual cortex: evidence for a new organization principle. *Psychol Res* 77:74–97. [CrossRef Medline](#)

Wiggett AJ, Pritchard IC, Downing PE (2009) Animate and inanimate objects in human visual cortex: evidence for task-independent category effects. *Neuropsychologia* 47:3111–3117. [CrossRef Medline](#)