

Dissociations and Associations between Shape and Category Representations in the Two Visual Pathways

 Stefania Bracci and Hans Op de Beeck

Laboratory of Biological Psychology, KU Leuven, 3000, Belgium

The dorsal and ventral visual pathways represent both visual and conceptual object properties. Yet the relative contribution of these two factors in the representational content of visual areas is unclear. Indeed, research investigating brain category representations rarely dissociate visual and semantic properties of objects. We present a human event-related fMRI study with a two-factorial stimulus set with 54 images that explicitly dissociates shape from category to investigate their independent contribution as well as their interactions through representational similarity analyses. Results reveal a contribution from each dimension in both streams, with a transition from shape to category along the posterior-to-anterior anatomical axis. The nature of category representations differs in the two pathways: ventral areas represent object animacy and dorsal areas represent object action properties. Furthermore, information about shape evolved from low-level pixel-based to high-level perceived shape following a posterior-to-anterior gradient similar to the shape-to-category emergence. To conclude, results show that representations of shape and category independently coexist, but at the same time they are closely related throughout the visual hierarchy.

Key words: cortical organization; object category; object shape; representational similarity analysis; visual cortex

Significance Statement

Research investigating visual cortex conceptual category representations rarely takes into account visual properties of objects. In this report, we explicitly dissociate shape from category and investigate independent contributions and interactions of these two highly correlated dimensions.

Introduction

Visual information is processed throughout a series of hierarchical stages in at least two pathways: a ventral stream for object recognition and a dorsal stream for the visual guidance of actions (Goodale and Milner, 1992; Kravitz et al., 2013). At different stages along the ventral visual pathway, neurons are tuned to object contours and curvatures, position, and 3D object configurations (Kobatake and Tanaka, 1994; Brincat and Connor, 2004; Yamane et al., 2008). Higher up, in both human and monkey inferotemporal cortex, a large-scale division for animate and inanimate entities and further subdivisions within the animate domain for faces and bodies has been shown (Kiani et al., 2007; Kriegeskorte et al., 2008b; Bell et al., 2009). Although a large body

of evidence has shown that the ventral stream plays a critical role in representing both shape and semantic information (e.g., Grill-Spector et al., 1998), it is not clear what is the relative contribution of these two factors, and the nature of their interaction, in visual cortex representations.

Recent work in monkeys (Rajimehr et al., 2011; Baldassi et al., 2013; Yue et al., 2014) as well as in humans (Nasr et al., 2014; Rice et al., 2014; Watson et al., 2014) has suggested that the organization of category representations in high-level visual cortex reflects brain selectivity for visual features, including relatively low-level dimensions, such as spatial frequency and local orientation content, which are typically associated with primary visual cortex. Similar suggestions have been proposed for representations in the dorsal visual pathway (Sakuraba et al., 2012). However, these studies did not directly compare object shape and object category information, thus making it difficult to separate contributions of these two factors.

The same problem arises for studies investigating category selectivity. Category distinctions are typically correlated with visual dimensions. Entities within the same object category share similar shape features. Faces are round and bodies are elongated. Most animals have four legs, a face, and round contours that largely differ from most man-made inanimate objects (e.g., a

Received June 16, 2015; revised Oct. 15, 2015; accepted Nov. 19, 2015.

Author contributions: S.B. and H.O.d.B. designed research; S.B. performed research; S.B. analyzed data; S.B. and H.O.d.B. wrote the paper.

This work was supported by European Research Council ERC-2011-Stg-284101, federal research action IUAP-P7/11, and Hercules Grant ZW11_10. We thank Nicky Daniels and Jessica Bulthé for technical assistance; and Niko Kriegeskorte for providing the multiple object arrangement code.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Stefania Bracci, Laboratory of Biological Psychology, KU Leuven, 3000, Belgium. E-mail: stefania.bracci@kuleuven.be.

DOI:10.1523/JNEUROSCI.2314-15.2016

Copyright © 2016 the authors 0270-6474/16/360432-13\$15.00/0

bookshelf). Given these constraints, it is not surprising that most studies investigating category selectivity do not control for shape differences among stimuli within and across classes. Typically, such studies resort to the *post hoc* application of relatively limited computational models to argue that category effects cannot be reduced to visual features. However, these visual features are never captured fully by the models (Kriegeskorte et al., 2008a; Op de Beeck et al., 2008b). Thus, it is hard to exclude the possibility that observed large-scale divisions (e.g., animate/inanimate) might be largely accounted for by object shape information.

To address these issues and compare the contribution of shape and category information within the two visual pathways, we implemented a two-factorial event-related fMRI design where shape and category membership are manipulated independently. This design allowed us to separate object shape and object category and investigate the contribution of the two factors.

Materials and Methods

Participants

The fMRI study included 15 right-handed adult volunteers (8 females; mean age, 24 years). One participant was excluded because of excessive head motion. All participants gave informed consent to take part in the fMRI experiment. The study was approved by the ethics committee of the KU Leuven.

Stimuli

The 54 stimuli are shown in Figure 1. Six categories of objects were included in an event-related design fMRI experiment: minerals, animals, fruit/vegetables, musical instruments, sport articles, and tools. Our stimulus set included categories that are superordinate to the basic level as originally defined by Rosch et al. (1976). We included superordinate categories such as animals and fruit instead of basic categories such as faces, bodies or cars to reduce low-level visual properties similarities across exemplars of the same category. In addition, functional categories such as musical instruments, sport articles, and tools were included to target dorsal stream representations. Each category consisted of 9 gray-scale images on a white background and had a size of $8^\circ \times 8^\circ$ (400×400 pixels). For each category, each image had unique shape properties, thus creating 9 subsets of images with similar shape properties. Thus, the category and shape dimension were orthogonal to each other; each shape type (e.g., round) contained 1 image from each of the 6 object categories and each object category (e.g., animals) contained 1 image from each shape type. Categories included in our stimulus set might have been biased toward selection of certain categories due to the necessity to dissociate the two dimensions. The nine summed images obtained by summing all images from each shape type are shown in the last row of Figure 1. The six summed images obtained by summing all images from each object category are shown in the last column of Figure 1. These six summed images suggest that object category could not be distinguished based on object shape properties. As a measure of image low-level shape properties (image silhouette), we computed pixelwise similarities among images (Op de Beeck et al., 2008b). For the silhouette model, the resulting dissimilarity matrix ($1 - \text{correlation}$) is reported in Figure 2A (leftmost column). The value in each cell (top triangle) of this dissimilarity matrix reflects pixel-based differences for each object pair (blue represents large similarity). For the silhouette model, the 2D arrangement derived from multidimensional scaling (MDS) is reported Figure 2B.

Spatial frequency

To ensure that category and shape information could not be inferred based on differences in the spatial frequency content of the stimuli, we tested similarities in spatial frequencies and local orientation content. Each image was Fourier transformed to quantify the spectral power as a function of spatial frequency (1) averaged across orientations and (2) for each orientation separately. Subsequently, for these values, we computed the absolute difference across pairs of stimuli, and the resulting dissimilarity matrices were correlated with the shape and category dissimilarity matrices derived from behavioral judgments. Confirming the indepen-

dence of these two factors, neither spatial frequency alone (shape: $r = 0.07$; category: $r = 0.01$) nor spatial frequency together with orientation information (shape: $r = 0.13$; category: $r = -0.03$) could predict behavioral judgments about object shape and category.

Behavioral similarity judgments

Similarity judgments for the category and shape dimensions were rated by an independent group of participants ($N = 16$) using the multiple object arrangement method (Kriegeskorte and Mur, 2012). Differently from pairwise similarity judgments, the multiarrangement method allows for measuring multiple similarity judgments in a single arrangement, thus allowing each item to be rated in the context of all the remaining items (Kriegeskorte and Mur, 2012). Each participant rated all 54 images used in the functional neuroimaging study. For shape similarity, participants were asked to arrange the images based on perceived object shape similarity. For semantic category similarity, participants were asked to arrange the images based on the semantic similarity among objects. Results were averaged across participants. The shape and semantic category models are summarized in Figure 2 by means of dissimilarity matrices (see Fig. 2A) and multidimensional scale arrangements (see Fig. 2B).

Scanning procedure

The study consisted of two separated sessions, each performed in separated days. Each session included experimental runs as well as localizer runs. The stimuli presentation was controlled by a PC running the Psychophysics Toolbox package (Brainard, 1997) in MATLAB (The MathWorks). Pictures were projected onto a screen and were viewed through a mirror mounted on the head coil.

Experimental runs. Each session included 8 experimental runs (16 in total), each lasting 7 min and 40 s (230 volumes per run). For each subject and for each run, a fully randomized sequence of 54 image trials (repeated 2 times) and 18 fixation trials (repeated 2 times) was presented. No special optimization sequence was adopted. Each trial was presented for 1500 ms, followed by a fixation screen for 1500 ms. Each run started and ended with 14 s of fixation. Within the whole experiment, each stimulus was repeated 32 times. Participants performed a 1-back real-world size judgment task by pressing a button with their right index or middle finger if the current image was smaller or larger relative to the image presented in the previous trial. The fingers associated with each response were counterbalanced across runs.

Localizer runs. Seven categories of objects were included in a block-design fMRI localizer: whole bodies, hands, faces, tools, chairs, places, and scrambled images. Each condition consisted of 18 grayscale images (400×400 pixels) on a white background. In total, 4 functional localizer runs (2 runs for each session) were included in the study, each lasting 5 min and 12 s. Within each run a fully randomized sequence of 7 category blocks (each repeated 4 times) interleaved with a fixation block lasting 16 s was presented. At the beginning and at the end of each run, an additional fixation block was presented for 14 s. Within each category block, images were presented at the center of the screen for 400 milliseconds with a blank interstimulus interval of 400 ms. Participants performed a 1-back repetition detection task by pressing a button with their right index finger any time the same picture was presented two times in succession. In each block, 1 or 2 repetitions were presented.

Imaging parameters

Data collection was performed on a 3T Philips scanner with a 32-channel coil at the Department of Radiology of the University Hospitals Leuven. MRI volumes were collected using echo planar (EPI) T2*-weighted scans. Acquisition parameters were as follows: repetition time (TR) of 2 s, echo time (TE) of 30 ms, flip angle (FA) of 90° , field of view (FoV) of 216 mm, and matrix size of 72×72 . Each volume comprised 37 axial slices (covering the whole brain) with 3 mm thickness and no gap. The T1-weighted anatomical images were acquired with an MP-RAGE sequence, with $1 \times 1 \times 1$ mm resolution.

Preprocessing and data analysis

Imaging data were preprocessed and analyzed using the Statistical Parametrical Mapping software package (SPM 8, Wellcome Department of

Table 1. ROIs localization^a

ROIs	Functional contrast	Mask	No. of subjects	MNI coordinates		Cluster size		<i>p</i> ^b
				Left hemisphere	Right hemisphere	Left hemisphere	Right hemisphere	
Object								
LOTc-object	Chairs > scrambled images		14	−28 −79 18	30 −74 27	153	173	0.16
VOTc-object	Chairs > scrambled images		13	−29 −70 17	32 −70 27	40	52	0.20
Body								
LOTc-body	Bodies > chairs		14	−30 −78 19	33 −70 28	73	200	<0.001
Face								
LOTc-face	[(faces > chairs) and (faces > bodies)]		13	−28 −78 18	29 −73 27	31	36	0.40
Body/face								
VOTc-body/face	[(bodies + faces > chairs)]		14	−29 −70 18	32 −68 26	19	70	<0.001
Hand								
LOTc-hand	[(hands > chairs) and (hands > bodies)]		12	−31 −77 19	33 −69 28	76	33	0.03
Scene								
TOS-scene	Scenes > chairs		14	−24 −65 42	28 −65 47	160	219	0.04
aPPA-scene	Scenes > chairs		13	−27 −65 41	28 −61 42	49	95	<0.001
pPPA-scene	Scenes > chairs		14	−27 −65 42	28 −64 42	33	98	0.001
Early visual cortex								
EVC-1	All categories > baseline	BA-17	14	−23 −71 33	28 −58 38	186	138	0.002
EVC-2	All categories > baseline	BA-18	14	−25 −80 13	27 −74 23	567	503	0.05
Parietal cortex								
IPS-hand	Hands > chairs		14	−32 −57 49	34 −55 77	113	189	0.02
SPL	All categories > baseline	BA-5/7	14	−27 −64 48	28 −61 52	124	162	0.10
IPL	All categories > baseline	BA-40	14	−31 −59 47	33 −56 50	474	547	0.20
Prefrontal cortex								
DPFC	All categories > baseline	BA-46	14	−29 −53 43	33 −39 45	97	221	<0.001

^aFor each ROI, functional contrast, anatomical mask (if used), number of subjects in which the ROI was localized, single subject average MNI coordinates, and cluster size are reported.

^b*p* values indicate pairwise *t* test contrasting ROI cluster size across hemispheres.

Cognitive Neurology, London) and MATLAB. Functional images underwent the following preprocessing steps: slice timing correction, spatial realignment (to the first image) to adjust for individual head motion, coregistration of functional and anatomical images, segmentation, and spatial normalization to an MNI template. Functional images were resampled to a voxel size of $3 \times 3 \times 3$ mm and spatially smoothed by convolution of a Gaussian kernel of 4 mm FWHM (Op de Beeck, 2010).

We modeled the preprocessed signal for each voxel, for each participant and for each of the 54 images using a GLM. The GLM included regressors for each condition of interest (54 conditions) and the 6 motion correction parameters (*x*, *y*, *z* for translation and for rotation). Each predictor's time course was modeled by a boxcar function convolved with the canonical hemodynamic response function in SPM.

ROIs

Fifteen ROIs, which covered the wider lateral and ventral surface of OTc (LOTc; VOTc), and part of parietal and frontal cortices were defined in each individual participant. ROIs were defined by means of an independent functional localizer and (when necessary) the anatomical WFU PickAtlas Toolbox (Wake Forrest University PickAtlas, <http://fmri.wfubmc.edu/cms/software>). Object-selective voxels (chairs > scrambled images) were localized in lateral and ventral occipital temporal cortex (Grill-Spector and Malach, 2004). Face-selective voxels [conjunction of (faces > chairs) and (faces > bodies)] and body-selective voxels (bodies > chairs) could be defined separately in LOTc (Puce et al., 1996; Downing et al., 2001) (LOTc-face, LOTc-body) but not in VOTc (Peelen and Downing, 2005; Schwarzlose et al., 2005), where face and body voxels (face + body > chairs) were combined in a single ROI (VOTc-face/body). Hand-selective voxels [conjunction of (hands > chairs) and (hands > bodies)] were defined in LOTc (LOTc-hand) (Bracci et al., 2010). Additional hand-selective voxels (hands > chairs) were defined in the intraparietal sulcus (IPS-hand). Scene-selective voxels (scenes > chairs) were defined in the transverse occipital sulcus (TOS-scene) (Nasr et al., 2011), posterior and anterior parahippocampal gyrus (pPPA-scene, aPPA-scene) (Arcaro et al., 2009). Early visual areas (EVC-1 and EVC-2; all categories > baseline) were restricted to anatomical masking by Brodmann areas BA-17 and BA-18, respectively. Superior parietal lobe (SPL; all categories > baseline) and inferior parietal lobe (IPL; all categories > baseline) were restricted to the anatomical mask BA-5/7 and BA-40, respec-

tively. Finally, dorsal prefrontal cortex (DPFC; all categories > baseline) was restricted to the anatomical mask BA-46. ROIs included all spatially contiguous voxels that exceeded the statistical uncorrected threshold $p < 0.001$. When <25 active voxels were found at this threshold, a more liberal threshold of $p < 0.01$ was applied. Only ROIs with at least 25 active voxels were included in an individual subject. To ensure that all ROIs were anatomically independent from each other, a hierarchical inclusion criterion was applied which reflected the functional criterion. For example, if a subset of object-selective voxels were also selective for bodies, object-selective voxels were defined after excluding body-selective voxels (those voxels where the response to bodies was significantly higher than chairs). Table 1 reports details on ROIs' localization (e.g., functional contrast, cluster size), and Figure 3 shows all ROIs in one representative participant. These ROIs provide a continuous and comprehensive window on the large cortical area activated when perceiving objects, the latter being displayed in gray on the small brain maps in Figure 3 (all categories > baseline, $p = 0.00001$, uncorrected). With the exception of the region around the central sulcus, which is probably activated due to the execution of the motor response, a large part of the regions activated by viewing the object images is covered by the aforementioned ROIs. The choice of combining functional and anatomical criteria to define our ROIs, over possible alternatives such as anatomical parcellation, was preferred. This combined method (in several variations), very common in research on the visual system, allows determining functionally specific ROIs by including a large portion of visually active voxels and at the same time excluding voxels that do not show reliable visually object-related information.

The ROIs differed in size. Whereas differences in ROIs size can lead to differential results in classification-based analyses, correlation-based analyses are not affected by different ROIs' size. To confirm this, we repeated our analyses using the same number of voxels for each ROI. As expected, exactly the same results were obtained in the two analyses.

Multivoxel pattern analysis

We used correlation-based multivoxel pattern analysis to analyze how the spatial response pattern in individual ROIs differs between experimental conditions (Haxby et al., 2001). Parameter estimates ("responses") for each condition (relative to baseline) were extracted for each voxel in an ROI, for each participant and each run, and normalized per run by subtracting the mean response across all conditions for each voxel separately. The full dataset was divided into two independent subsets of

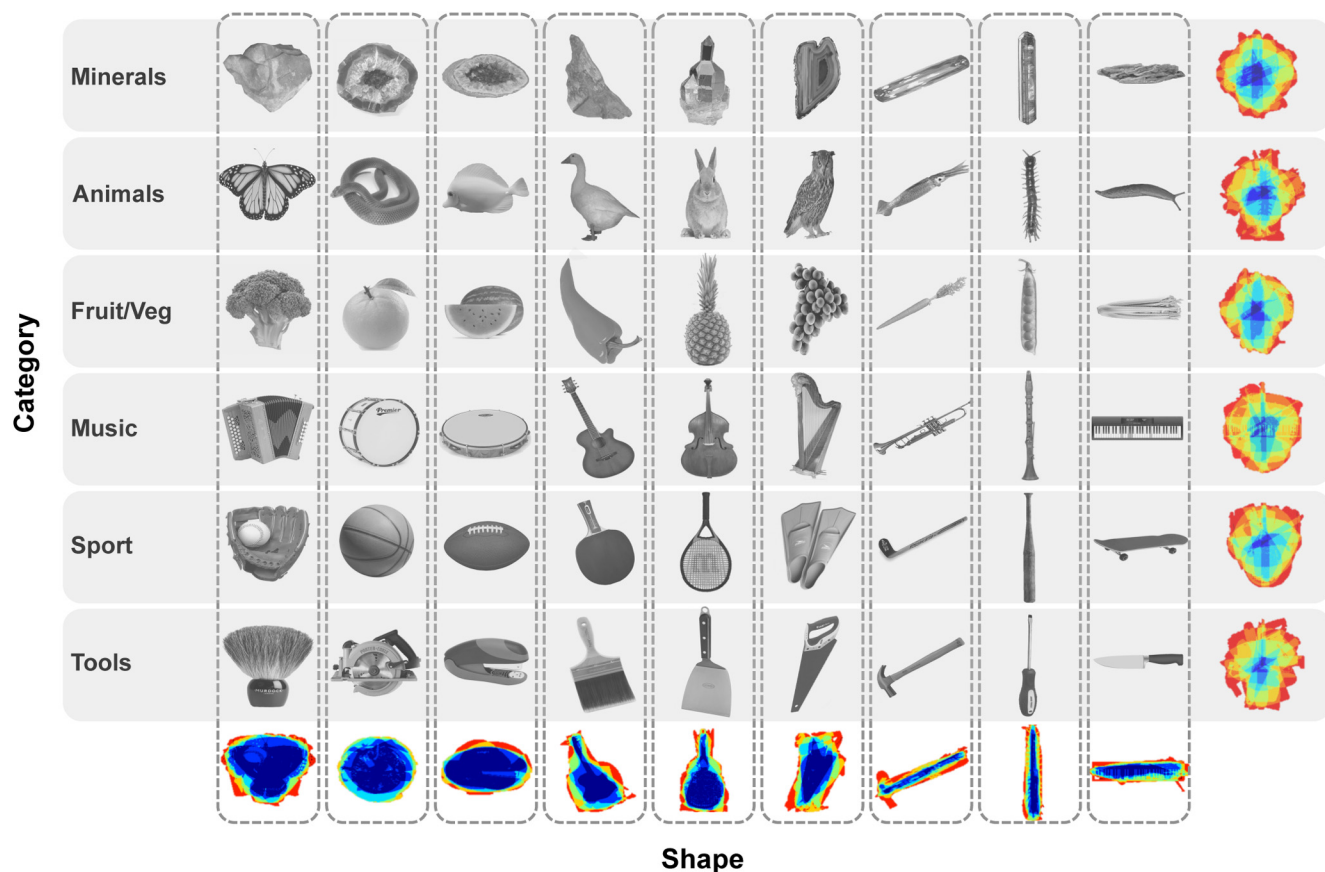


Figure 1. Experimental stimuli. The stimulus set consisted of 54 unique images comprising 6 object categories (rows) and 9 shape types (columns). Each object category (e.g., minerals) included 9 images (one from each shape type), with unique shape features. Each shape type included 6 images (one from each object category) with similar shape features. The pixelwise overlap obtained by summing all images from each shape type and each object category are shown in the last row and last column, respectively. Analyses of the 54×54 dissimilarity matrix (Fig. 2A, B, leftmost column), obtained from the pixel-based overlap between pairs of images, reveal how strongly this physical measure of low-level dissimilarity is dominated by shape: large differences between stimuli from the same object category and small differences between stimuli from the same shape type (Fig. 1, last row and last column). Thus, object category could not be distinguished based on shape information.

runs (Set 1 and Set 2). The multivoxel patterns of activity associated with each condition (e.g., fish) in Set 1 were correlated with the activity patterns in Set 2. This procedure of splitting the data in two was repeated 100 times. Correlations were averaged across the 100 iterations, thus resulting in an asymmetric 54×54 correlation matrix for each participant and ROI. Subsequently, the two halves (above and below the diagonal) of the correlation matrix were averaged, and only the upper triangle of the resulting symmetric matrix was used in the following analyses. To test whether the response pattern in an ROI conveyed information about stimulus identity, we compared the average of within-condition correlations (diagonal cells) for each ROI with the average of between-condition correlations (off diagonal cells). Pairwise t tests across participants revealed significant reliability of response patterns for each ROI ($p < 0.01$; for all tests). Thus, the multivoxel patterns convey information about the presented conditions in all the ROIs. Subsequently, correlation matrices were converted into dissimilarity matrices ($1 - \text{correlation}$) and used as neural input for the representational similarity analysis (RSA) (Kriegeskorte et al., 2008a). As before (Op de Beeck et al., 2008b), we correlated the behavioral dissimilarity matrices for shape and semantic category with the neural dissimilarity matrix of each ROI. Resulting correlations were Fisher transformed $\{0.5 \times \log[(1 + r)/(1 - r)]\}$.

To take into account the noise in the data, for each ROI we computed an estimate of the reliability of the data, which provides an indication of the maximum correlations we can expect given the signal-to-noise ratio of the data (Op de Beeck et al., 2008b). For each subject and each ROI, the 54×54 correlation matrix was correlated with the averaged correlation matrix of the remaining participants. Values were averaged across participants. The resulting correlation values capture noise inherent to a

single subject as well as noise caused by intersubject variability. This measure of reliability gives an estimate of the highest correlation we can expect in each ROI when correlating behavioral dissimilarity (e.g., shape model) and neural dissimilarity (e.g., activation pattern in each ROI). We provide this measure as a reference in all the relevant data figures (gray-shaded background bars). It would also be possible to normalize for reliability by dividing all correlations through the reliability, and such approach results in very similar statistics and conclusions as our main, not normalized analyses.

MDS and hierarchical cluster analysis

MDS and hierarchical cluster analysis were used to visualize and compare neural similarity structures in all ROIs and similarity structures related to pixel-based overlap (silhouette similarity) and the behavioral models (shape similarity and category similarity). Metric MDS was performed using MATLAB function “mdscale” normalized with the sum of squares of the dissimilarities. The hierarchical cluster analysis was performed using the MATLAB function “linkage” using the nearest distance default method.

Results

Shape and category information in the ventral and dorsal pathway

We collected behavioral and neural data on a set of 54 images (Fig. 1). An independent group of participants ($N = 16$) performed similarity judgments on all images for the shape and the semantic category dimension (see Materials and Methods). As

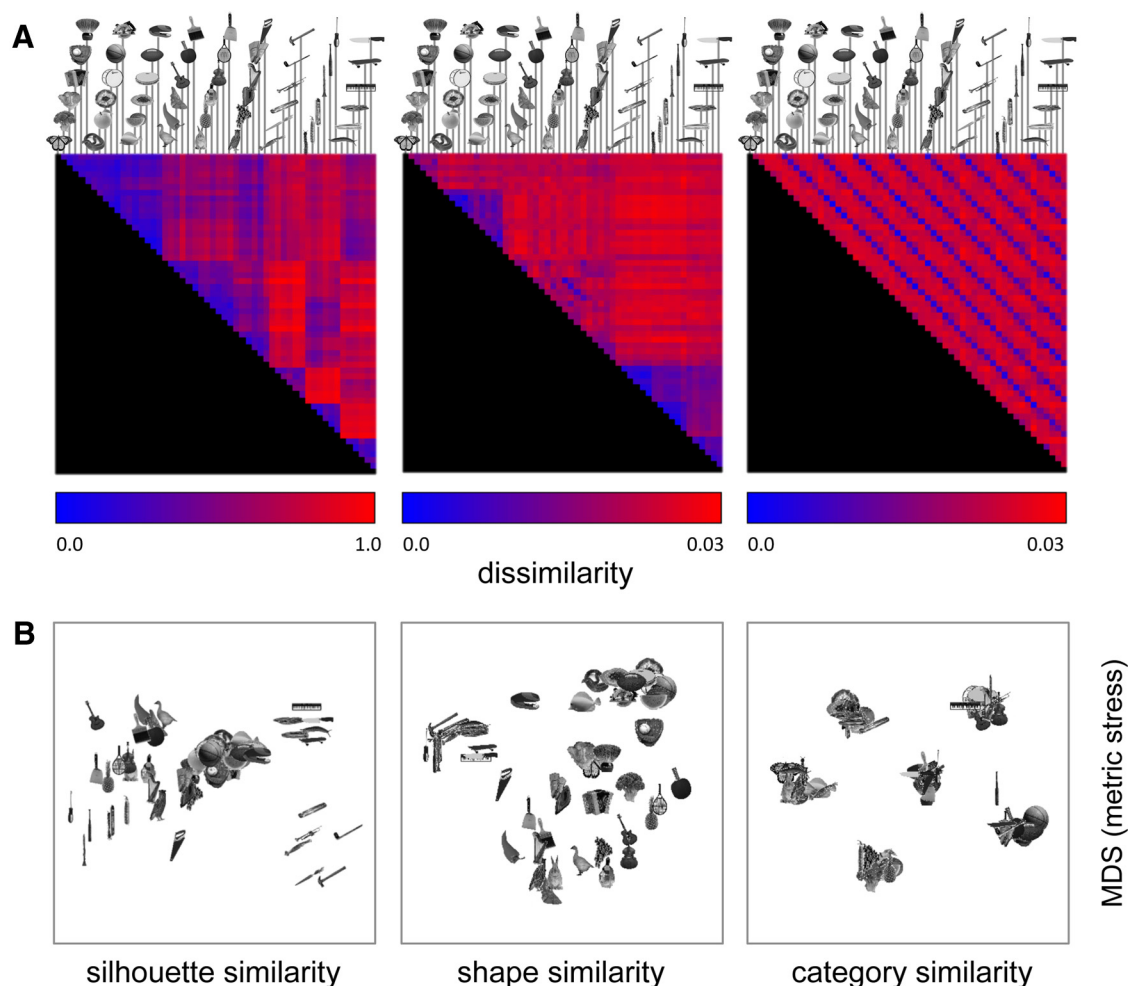


Figure 2. Models. **A**, Mean representational dissimilarity matrices (red represents large dissimilarities) and **B** 2D arrangements derived from MDS for the silhouette model (leftmost panel), the perceived shape model (middle panel), and the category model (rightmost panel). Correlations between dissimilarity matrices are as follows: silhouette similarity and shape similarity ($r = 0.25$); silhouette similarity and category similarity ($r = -0.05$); and shape similarity and category similarity ($r = -0.01$).

intended, the two dimensions were independent ($r = -0.01$) and revealed a very different representational space (Fig. 2B).

To investigate how information about shape and category is distributed throughout the ventral and dorsal pathway, dissimilarity matrices derived from behavioral judgments (Fig. 2) were compared with neural dissimilarity matrices derived from ROI activity patterns (see Materials and Methods) by means of RSA (Kriegeskorte et al., 2008a). Defined ROIs (see Materials and Methods) covered a large cortical area of visually active voxels within both visual pathways (Fig. 3). As shown in Figure 4A, the neural similarity in most ROIs (BA-17, BA-18, TOS-scene, pPPA-scene, aPPA-scene, LOTC-object, LOTC-face, LOTC-body, LOTC-hand, VOTC-object, VOTC-body/face, and SPL) showed significance above baseline correspondence with shape similarity as rated behaviorally ($p < 0.004$ for all tests; Fig. 4A, green asterisks). We refer to these regions as shape-sensitive ROIs. Shape information was not present in IPS-hand, IPL, and DPFC. Different results were observed for the category dimension. Whereas category information was not present in early visual areas (EVC-1, EVC-2: $t < 1$ for both tests) and scene-selective areas (TOS, pPPA, aPPA: $t < 1$ for all tests), significance above baseline category information was observed in object/face/body/hand selective areas in lateral and ventral OTC (LOTC-object, LOTC-face, LOTC-body, LOTC-hand, VOTC-object, VOTC-

body/face), SPL, IPL, IPS-hand, and DPFC ($p < 0.01$ for all tests; Fig. 4A, orange asterisks). We refer to these regions as category-sensitive ROIs. Thus, category-related information is present, even when shape similarity is orthogonal to category membership and goes against it so that stimuli with high shape similarity belong to different categories.

Next, we assessed how the representational content is changing across ROIs. To compare representational content across ROIs, we performed second-order correlations across ROI's correlation matrices averaged across subjects. The resulting dissimilarity matrix ($1 - \text{correlation}$; Fig. 4B) captures similarities in representational content among ROIs: similarities in two ROIs' representational content (e.g., BA17 and BA18) suggest that these ROIs represent the stimuli in a similar manner. The application of MDS to this ROI similarity matrix revealed a 2D arrangement (Fig. 4C) in which the first (horizontal) dimension seems related to the anatomical posterior-to-anterior axis, and the second dimension to the ventral-to-dorsal axis. In Figure 4C, the color-coding of the ROIs in terms of their selectivity for shape (green color-coded), category (orange color-coded), or both types of information (yellow color-coded) suggests a transition in representational content from shape-sensitive ROIs to category-sensitive ROIs along the anatomical posterior-to-anterior axis. To quantify the relationship between the type of selectivity and

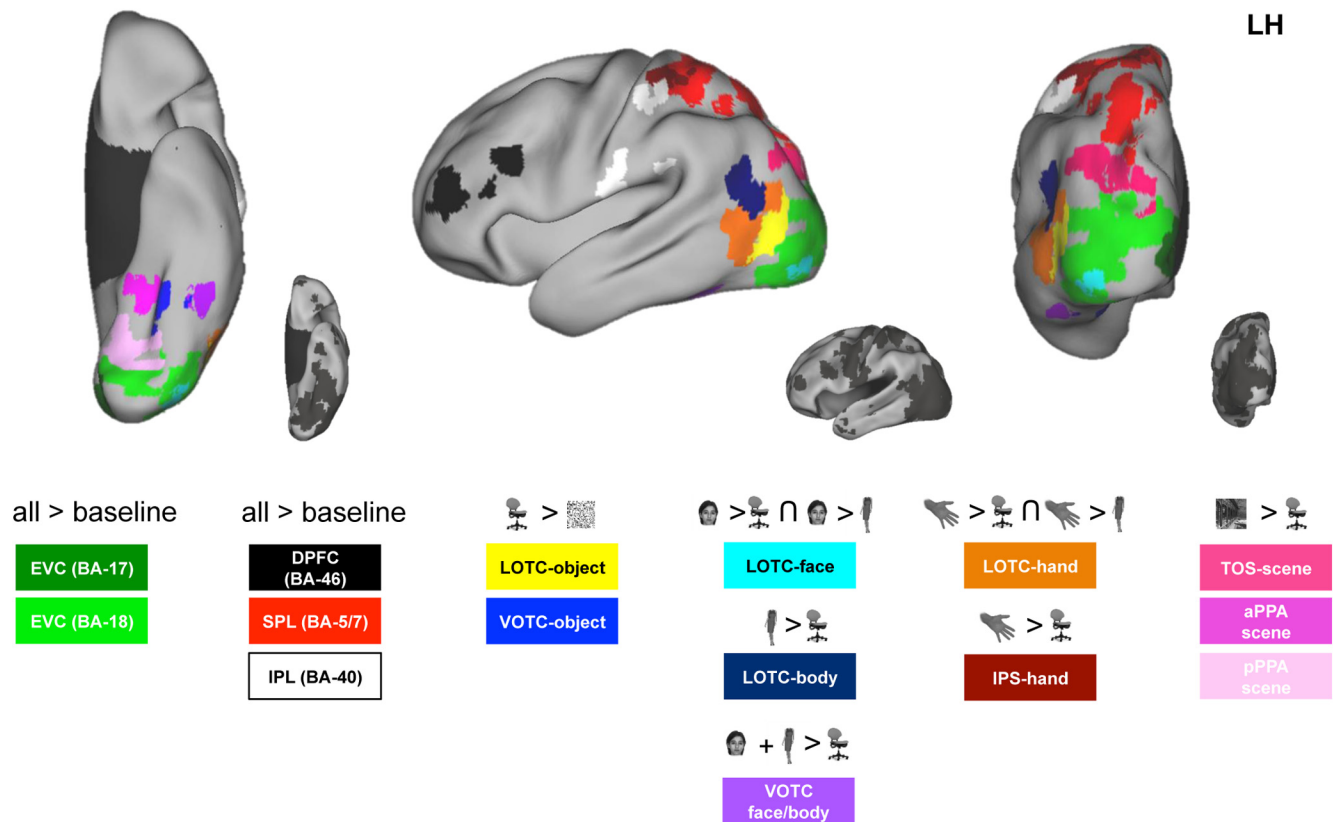


Figure 3. ROIs. Individual-participant ROIs are shown for the left hemisphere of one representative participant on a ventral, lateral, and posterior view of the inflated PALS human brain template (Van Essen, 2005). Small brain maps display the significantly activated voxels when contrasting all categories included in the localizer with the fixation baseline ($p = 0.00001$, uncorrected). LH, Left hemisphere.

the anatomical position of an ROI, we correlated the ROI's mean anatomical location on the y -axis with the amount of category/shape information present in each ROI ($\text{cs-index} = (\text{category, shape})/(\text{category} + \text{shape})$). Further confirming results shown in Figure 4C, we found a highly positive correlation between ROI's mean anatomical location on the y -axis and the cs-index ($r = 0.76$); from posterior to anterior, shape information decreases and category information emerges. Together, these results show a transition from object shape to object category along the posterior-to-anterior anatomical axis, with many regions in high-level visual cortex encoding both types of information.

In sum, we observed that category selectivity could not be reduced to object visual properties, such as perceived shape. Nevertheless, many of the high-level visual regions encoded both dimensions of object images: the shape and the category to which they belong. In the next sections, we will further characterize information content for shape-sensitive and category-sensitive ROIs separately.

Characterizing information content in category-sensitive ROIs

What type of “category” information is represented in category-sensitive ROIs? The hierarchical cluster analysis (see Material and Methods) performed on category-sensitive ROIs (Fig. 4A, orange asterisks) revealed two main clusters reflecting differences in ROIs' representational content (Fig. 5A): one cluster for more ventral ROIs in occipitotemporal cortex (light blue color-coded: LOTc-object, LOTc-body, VOTc-object, VOTc-face/body, LOTc-face, LOTc-hand) and one cluster for more dorsal ROIs in parietal and prefrontal areas (dark blue color-coded: IPS-

hand, SPL, DPFC, IPL). This clustering suggests differential category-related information content for areas within the dorsal and ventral visual pathway. Figure 5D shows the category clusters for one representative subject on a brain template.

The candidate hypotheses for the nature of these differences can be found in the literature. Although our results add to the increasing evidence that object representations are encoded in both visual pathways (Konen and Kastner, 2008; Grill-Spector and Weiner, 2014), the two streams are supposed to support different computations; whereas the ventral stream processing primarily supports object perception, such as the animate/inanimate division (Warrington and Shallice, 1984), the dorsal stream processing sustains action-related computations (Buxbaum et al., 2014). Does information content in the ventral and dorsal cluster reflect this distinction? Up to now, no single study directly compared these two hypotheses in ventral as well as dorsal visual cortex.

To address this question, we used RSA to compare neural similarity matrices derived from ROIs' activity patterns (averaged across ROIs within each cluster) to two category models: (1) The animate/inanimate model (Fig. 5B, light blue border) captures the animate/inanimate division previously reported in visual cortex (Kriegeskorte et al., 2008b; Konkle and Caramazza, 2013). This model assumes high correspondence between neural patterns for two animate objects (two animals) and for two inanimate objects (e.g., a mineral and a musical instrument), but low correspondence between neural patterns for one animate and one inanimate object (e.g., an animal and a mineral). (2) The action/nonaction model (Fig. 5B, dark blue border) captures sen-

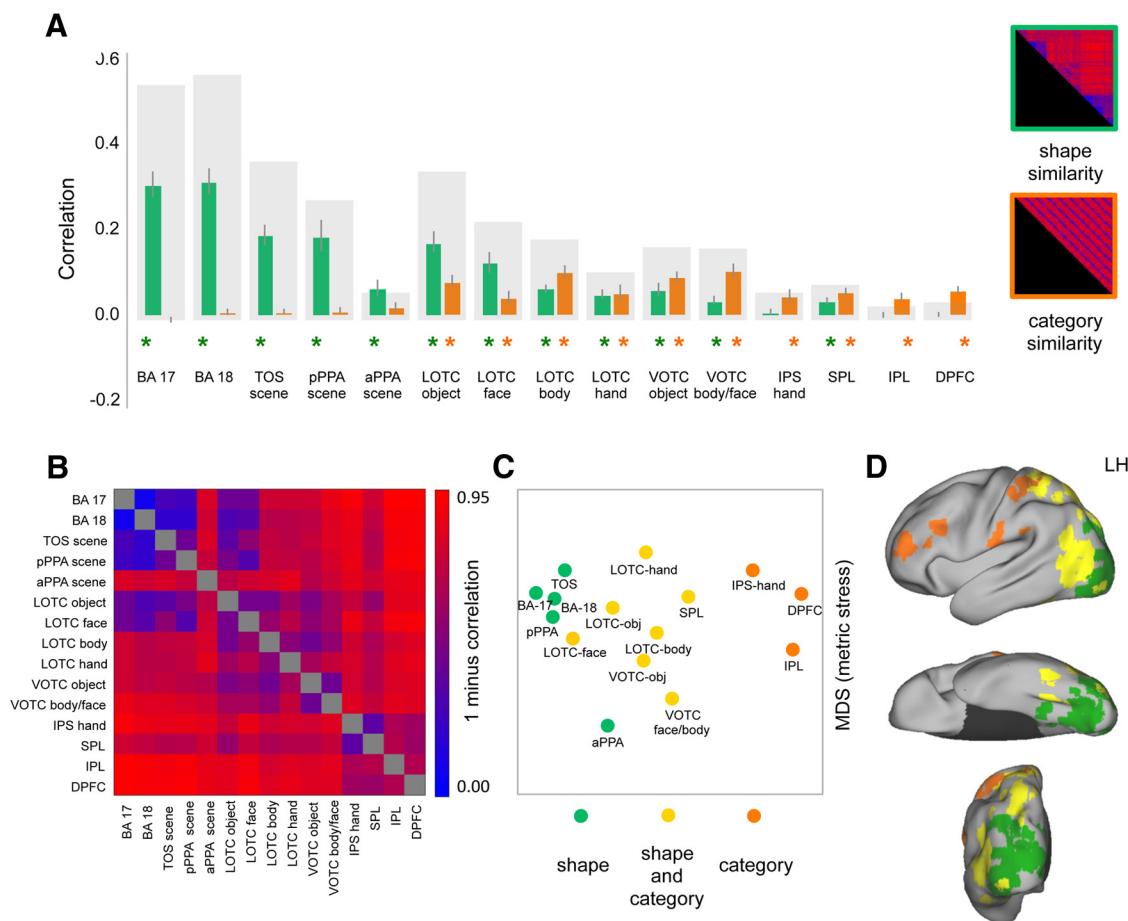


Figure 4. Representational similarity analysis for shape and category. **A**, Results of ROI RSA for the shape similarity (green color-coded) and the category similarity (orange color-coded). Green and orange asterisks indicate ROIs with significant shape and category information, respectively. For each ROI, the gray-shaded background bar represents the reliability of the correlational patterns in each ROI, which provides an approximate upper bound of the observable correlations between behavioral and neural data. Error bars indicate SEM. **B**, The dissimilarity matrix, derived from second-order correlations across ROI's correlation matrices (averaged across subjects), shows similarities across ROI's representational content. Blue represents similar ROI's structural content. **C**, MDS, performed on the ROI dissimilarity matrix from **B**, shows ROI pairwise distances in a 2D arrangement. Pairwise distances reflect response-pattern similarity: ROIs positioned next to each other have a similar information content, whereas ROIs positioned far from each other show dissimilar information content. **D**, Shape-sensitive (green color-coded), category-sensitive (orange color-coded), and both shape/category-sensitive (yellow color-coded) ROIs are shown for one representative subject on the inflated PALS human brain template (Van Essen, 2005). LH, Left hemisphere.

sitivity to action-related properties of objects, which might be more emphasized in dorsal stream areas. Our stimulus set includes three action-related object categories (sport articles, musical instruments, and tools). The action/nonaction model predicts high correspondence between neural patterns for two action-related objects (e.g., a musical instrument and a tool) and two nonaction objects (e.g., an animal and a mineral), but low correspondence between neural patterns for one action-related and one nonaction object (e.g., a musical instrument and a mineral).

Results from RSA were tested in a 2×2 ANOVA with Cluster (ventral, dorsal) and Model (animate/inanimate, action/nonaction) as within-subject factors. Results revealed a significant Cluster \times Model interaction ($F_{(1,13)} = 31.15$, $p = 0.00009$; Fig. 5B), as such indicating that the differences in the relation between models and representational content in the ventral and dorsal cluster can be captured by these two models. *Post hoc* pairwise t tests further confirmed this dissociation: whereas in the ventral cluster (Cluster 1) the animate/inanimate model could significantly better explain the neural pattern relative to the action/nonaction model ($t_{(13)} = 4.70$, $p = 0.0004$; Fig. 5B), in the dorsal cluster (Cluster 2) the action/nonaction model was significantly

more related to the neural data relative to the animate/inanimate model ($t_{(13)} = 2.9$, $p = 0.01$; Fig. 5B).

The correlations between the best model and the neural similarity matrix are not as high as they could be given the reliability of the data; thus, representations in the ventral and the dorsal stream are not captured fully by any of these models. Nevertheless, the models capture important aspects of those representations. This is visible in the 2D space formed by the dimensions that capture most variation in the neural similarity matrices according to MDS. These spatial configurations illustrate the animate/inanimate (Fig. 5C, left) and the action/nonaction (Fig. 5C, right) division in the ventral and the dorsal cluster, respectively. Thus, despite the fact that both ventral and dorsal regions show category information, this first direct comparison of the two pathways through RSA confirms that the informational content in ventral and dorsal regions differs and reflects some of the hypothesized different computations happening in the ventral and dorsal pathway.

Characterizing information content in shape-sensitive ROIs

What type of “shape” information is represented in shape-sensitive ROIs? The hierarchical cluster analysis (see Materials

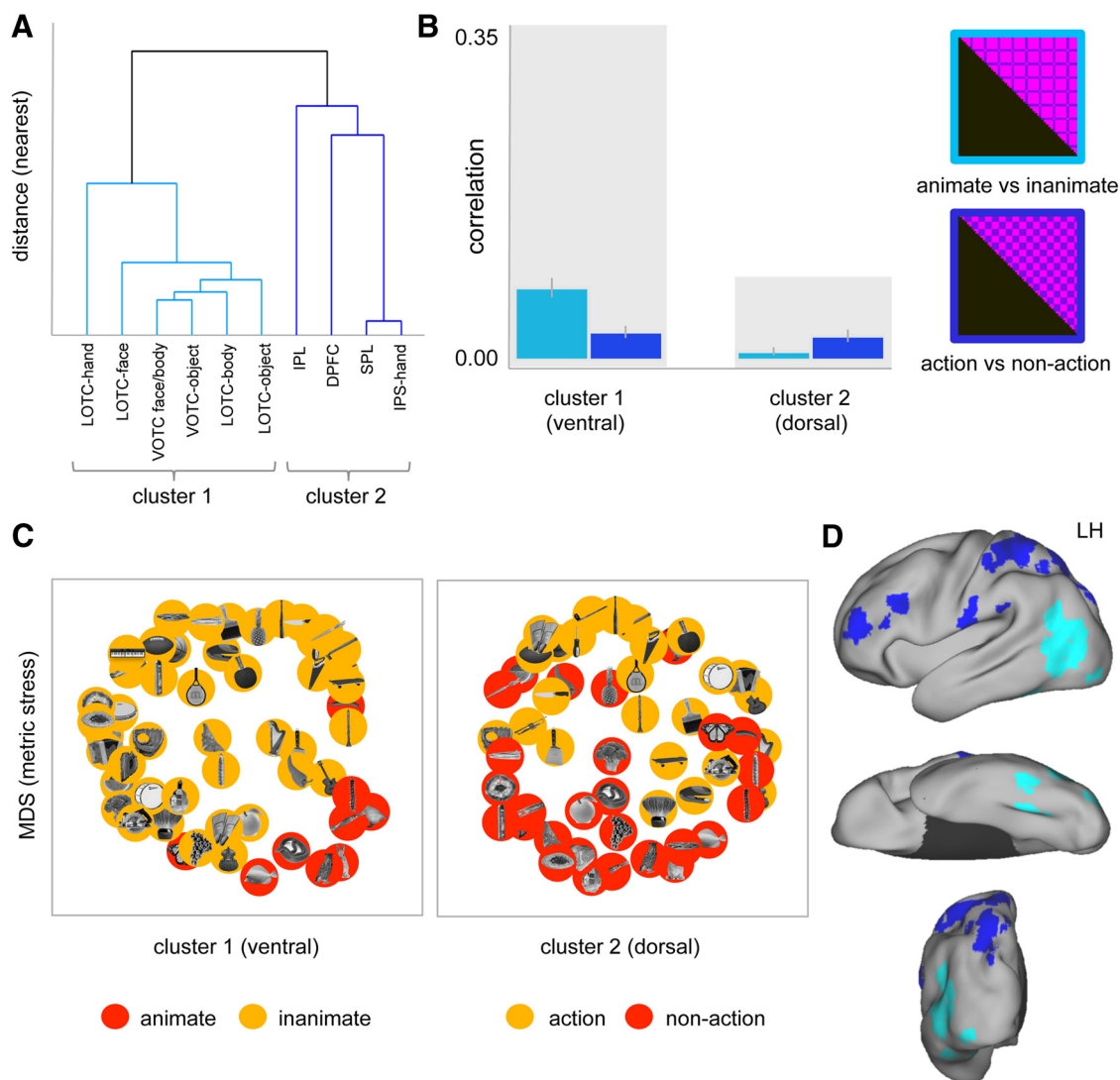


Figure 5. Characterizing information content in category-sensitive ROIs. **A**, The hierarchical plot derived from a hierarchical cluster analysis shows ROI's activity-pattern similarity structure for category-sensitive ROIs. Results revealed two separated clusters: one (Cluster 1) for ventral stream areas and one (Cluster 2) for dorsal stream areas. **B**, Results of RSA at the cluster level for the animate/inanimate model (light blue color-coded) and the action/nonaction model (dark blue color-coded). For each cluster, the gray-shaded background bar represents the reliability of the correlational patterns in each cluster. Error bars indicate SEM. **C**, MDS, performed on neural dissimilarity matrices ($1 - \text{correlation}$), averaged across ROIs within each cluster, shows object pairwise distances in a 2D space for the ventral cluster (left) and the dorsal cluster (right). Pairwise distances reflect response-pattern similarity: the animate/inanimate division and the action/nonaction division are clearly visible in the ventral cluster (left) and the dorsal cluster (right), respectively. **D**, ROIs from Cluster 1 (light blue color-coded) and Cluster 2 (dark blue color-coded) are shown for one representative subject on the inflated PALS human brain template (Van Essen, 2005). LH, Left hemisphere.

and Methods) performed on shape-sensitive ROIs (Fig. 4A, green asterisks) revealed two main clusters (Fig. 6A). Cluster 1 included BA-17, BA18, TOS-scene, pPPA-scene, LOTC-face, and LOTC-object. Cluster 2 included VOTC-face/body, VOTC-object, LOTC-body, and LOTC-hand. Two additional ROIs (SPL, aPPA-scene) did not group with either cluster and were excluded from subsequent analyses. The ROIs in the two shape clusters differed in their anatomical location; along the posterior-to-anterior anatomical y -axis, single-subject MNI coordinates, averaged across ROIs within each cluster, were significantly more posterior in Cluster 1 than in Cluster 2 ($t_{(11)} = 22.2$, $p < 0.0001$; only subjects where all ROIs could be defined were included in this analysis). Figure 6D shows the shape clusters for one representative subject on a brain template. Together, these results suggest differential shape-related information content in posterior (Cluster 1) and anterior (Cluster 2) shape-sensitive ROIs.

In the visual system, information about shape is processed throughout a series of hierarchical stages, so that early visual areas process image low-level visual properties, such as position and orientation (Hubel, 1963), and extrastriate visual areas represent perceived object shape (Haushofer et al., 2008; Op de Beeck et al., 2008b) in a way that is tolerant to changes in object position, size, and orientation (Grill-Spector et al., 1999; James et al., 2002). To test whether representational content in the posterior and anterior “shape” cluster reflects this known hierarchical shape processing, we used RSA to compare neural similarity matrices derived from ROIs' activity patterns (averaged across ROIs within each cluster) to two shape models: (1) the “low-level” shape model (silhouette model), based on image pixel-wise similarities; and (2) the “high-level” shape model (shape similarity), derived from shape similarity judgments, which is the shape model used up to now in Results. Figures 2B and 6B illustrate the two models and their differences. The two shape models cor-

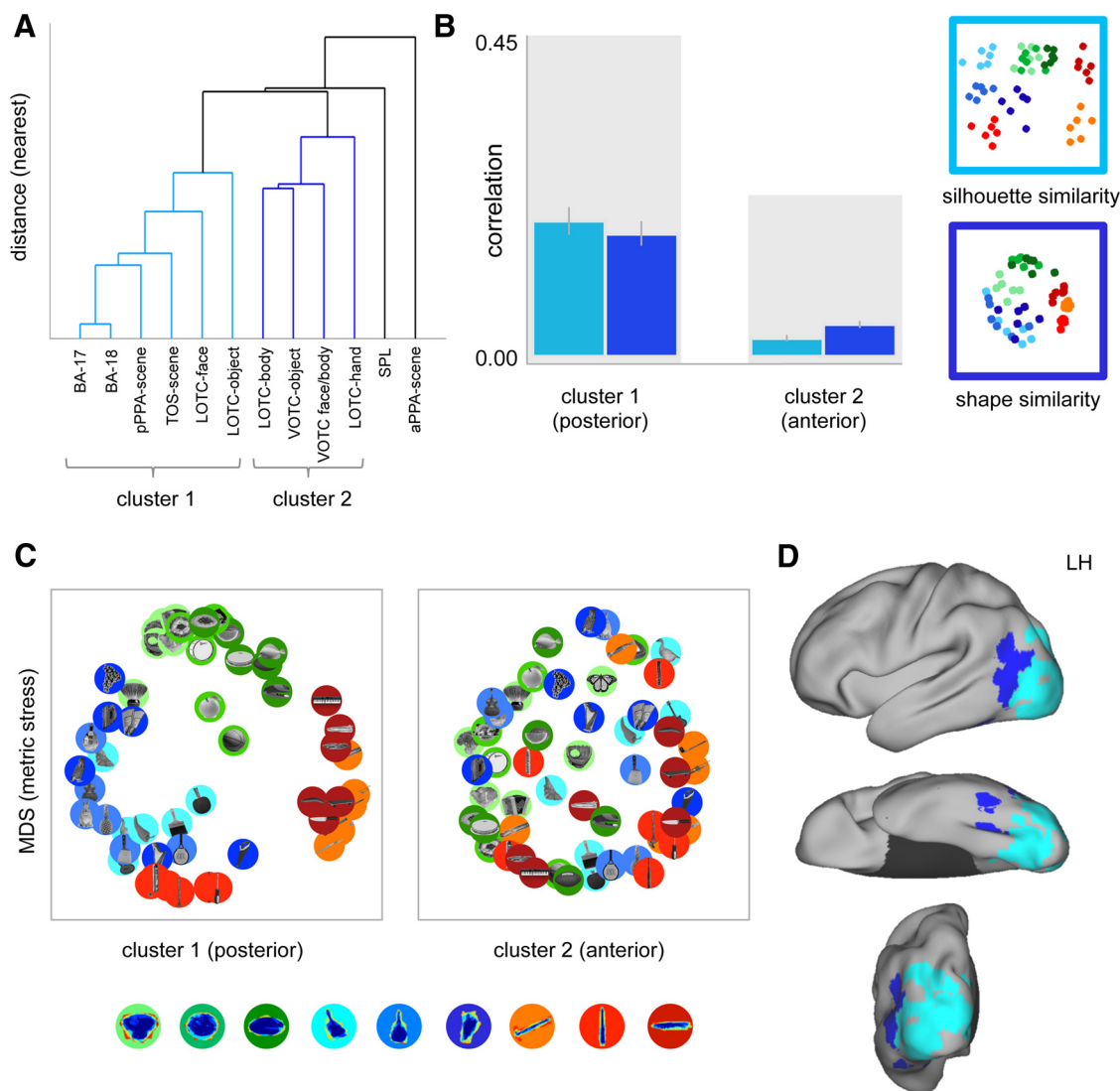


Figure 6. Characterizing information content in shape-sensitive ROIs. **A**, The hierarchical plot derived from the hierarchical cluster analysis shows ROI's activity-pattern similarity structure for shape-sensitive ROIs. Results revealed two separated clusters: Cluster 1 (BA-17, BA-18, pPPA-scene, TOS-scene, LOTC-face, LOTC-object) and Cluster 2 (LOTC-body, VOTC-object, VOTC-face/body, LOTC-hand). **B**, Results of RSA at the cluster level for the silhouette model (light blue color-coded) and the perceived shape model (dark blue color-coded). For each cluster, the gray-shaded background bar represents the reliability of the correlational patterns in each cluster. Error bars indicate SEM. **C**, MDS, performed on neural dissimilarity ($1 - \text{correlation}$) matrices averaged across ROIs within each cluster, shows object pairwise distances in a 2D space for the posterior cluster (left) and the anterior cluster (right). Pairwise distances reflect response-pattern similarity. **D**, ROIs from Cluster 1 (light blue color-coded) and Cluster 2 (dark blue color-coded) are shown for one representative subject on the inflated PALS human brain template (Van Essen, 2005). LH, Left hemisphere.

relate only partially ($r = 0.25$). Clear differences between the two models are obvious from visual inspection of the MDS solutions (Fig. 2B): in the perceived shape model, shape types cluster in three main subdivisions: elongated shapes (different shades of red), round shapes (different shades of green), and triangular shapes (different shades of blue). These divisions are not present in the silhouette model where the three elongated shape types are largely segregated. Stated otherwise, the judged shape similarity shows more tolerance for the image orientation of elongated stimuli.

Results from the RSA were tested in a 2×2 ANOVA with Cluster (posterior, anterior) and Model (silhouette, shape) as within-subject factors. This analysis revealed a significant Cluster \times Model interaction ($F_{(1,13)} = 11.6$, $p = 0.005$; see Fig. 6B), as such confirming differences in the relation between models and representational content in the posterior and anterior shape-sensitive ROIs. *Post hoc* pairwise t tests further confirmed this difference: whereas in the posterior cluster (Cluster 1) the silhouette model could significantly better explain the neural pattern

relative to the shape model ($t_{(13)} = 2.20$, $p = 0.05$; Fig. 6B), in the anterior cluster (Cluster 2) the shape model was significantly more related to the neural data relative to the silhouette model ($t_{(13)} = 4.71$, $p = 0.0004$; Fig. 6B). There was also a main effect of Cluster ($F_{(1,13)} = 90.5$, $p < 0.0001$), with much higher correlations overall in Cluster 1 than in Cluster 2. This main effect is at least in part a trivial consequence of the differences in reliability of the multivoxel patterns in the two clusters (Fig. 6B, gray bars). We should also note that the fit with the best shape model is far from perfect in each cluster: the highly significant correlations between the best model and the neural similarity data are smaller than what could be expected given the reliability of the data.

In Figure 6C, the 2D arrangements derived from MDS illustrate the representational structure in the posterior (left panel) and anterior (right panel) "shape" cluster. Consistent with the models, in the posterior cluster, vertical elongated objects (red color-coded stimuli) and horizontal elongated objects (dark red color-coded stimuli) were largely dissociated. In the anterior cluster,

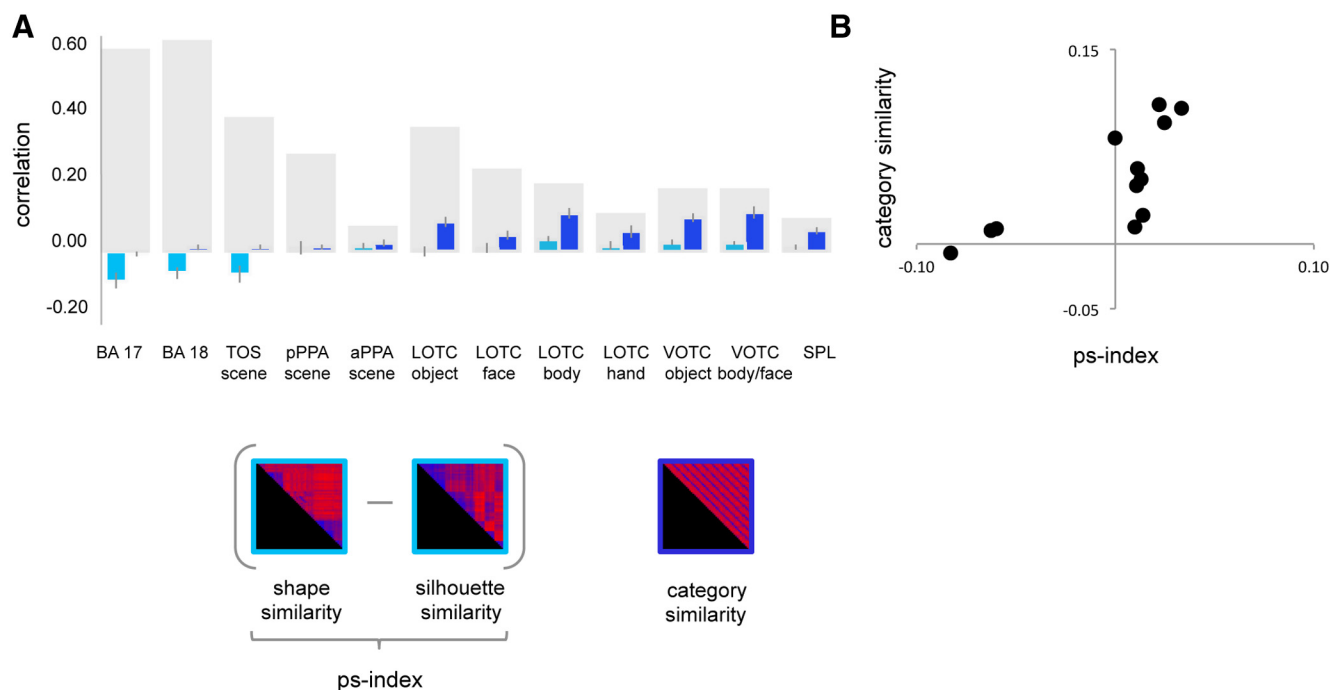


Figure 7. Interaction between shape and category. **A**, Results of ROI RSA for the perceived shape (ps-) index (light blue color-coded) and the category similarity (dark blue color-coded) in shape-sensitive ROIs. The ps-index reflects the relative amount of high-level compared with low-level shape information and was computed by subtracting “low-level” shape silhouette information (Fig. 6B, light blue) from “high-level” perceived shape information (Fig. 6B, dark blue) separately for each shape-sensitive ROI. For each ROI, the gray-shaded background bar represents the reliability of the correlational patterns in each ROI. Error bars indicate SEM. **B**, Scatterplot of the relationship between the ps-index and category information (averaged across subjects) across the 12 shape-responding ROIs.

clustering of individual stimuli was not very obvious, consistent with a lower reliability of the data and the possible contribution of other factors, such as category information (see Relation between the representation of shape and category), but overall elongated vertical objects were often in close proximity to elongated horizontal objects.

Relation between the representation of shape and category

The orthogonal manipulation of shape and category has allowed us to assess the separate contribution of each factor. In addition, we can investigate potential relations between the two factors. Despite the relatively artificial dissociation in our stimulus set, in a more general context, shape can be a reliable cue to recognize, identify, and categorize an object. Not just any shape feature is useful though, and in the literature it has been suggested that a shape representation useful for object recognition and basic-level categorization should be sensitive to features that allow transformation-invariant object recognition (Biederman, 1987; Kayaert et al., 2003). The above results showed that we have the sensitivity in our dataset to differentiate between low-level pixel-based shape features and more high-level shape representations: along the visual pathway, from posterior to anterior, there is a progression from “low-level” to “high-level” shape representation. If this subjective shape perception (e.g., recognizing elongate shapes regardless of orientation) has any role in the ability to categorize objects at a superordinate level (e.g., animate vs inanimate), then we expect a close relationship between subjective perception of shape and semantic category sensitivity across the ROIs sensitive to shape (Fig. 4A, all ROIs marked in green).

To investigate this question, we analyzed all ROIs with significant shape sensitivity (BA-17, BA-18, TOS-scene, pPPA-scene, aPPA-scene, LOTc-object, LOTc-face, LOTc-body, LOTc-hand, VOTc-object, VOTc-body/face, and the SPL). We calcu-

lated the relative amount of high-level compared with low-level shape information, referred to as the perceived shape index (ps-index) by subtracting the correlation with the “low-level” shape silhouette model (Fig. 6B, light blue) from the correlation with the “high-level” perceived shape model (Fig. 6B, dark blue) separately for each shape-sensitive ROI. Results for the ps-index and for category similarity (as reported in Fig. 4A, orange color-coded) are shown for all shape-sensitive ROIs in Figure 7A. The correlation between the ps-index and category information (averaged across subjects) was highly significant across the 12 ROIs ($r_{(10)} = 0.74$, $p = 0.006$), thus suggesting a close relationship between hierarchical shape processing and category selectivity in visual areas (Fig. 7B). This statistical analysis only takes into account the variation across ROIs to determine significance. The same analysis performed on single subjects confirmed that the relation between the ps-index and category information across shape-sensitive ROIs is significantly positive when considering the variation across subjects (one-sample t test computed across individual subject’s correlation values: $t_{(11)} = 17.24$, $p = 0.0001$; only subjects where all ROIs could be defined were included in this analysis). We note, however, that these results, including only 12 ROIs, are consistent with several variants of a positive relationship, including a more dichotomous distinction between low-ps regions and high-ps regions.

In sum, even though we have shown that semantic category selectivity cannot be reduced to shape selectivity, nor vice versa, we observed a close association between the two dimensions so that shape representations include more high-level shape properties in more category-sensitive regions.

Discussion

In addition to shared functional/semantic properties, objects within the same category typically share similar visual features.

Here we created a stimulus set where the category and the shape dimension were orthogonal to each other, thus allowing disentangling their unique contribution to object representations. We found evidence for the following: (1) category representations with different properties in the ventral and the dorsal stream; (2) shape representations of varying complexity; and (3) association between category and shape representations.

First, the pattern of activity in many regions in lateral and ventral occipitotemporal cortex was related to the category membership of stimuli. As such, our results put into perspective recent findings suggesting that category representations in the ventral visual pathway can be, in part, reduced to relatively simple visual properties (Rice et al., 2014; Watson et al., 2014). Nevertheless, the ventral category-sensitive regions, often regarded as the highest stage in visual information processing, also show shape sensitivity in addition to category sensitivity. As such, the present study allows us to conclude that lateral and ventral occipitotemporal representations contain information about category as well as shape. Thus, neither dimension alone can fully explain representational content in high-level visual areas. Given that the exact weight of visual and more semantic factors might depend upon various factors, further follow-up studies are needed to determine the relative weight given to these dimensions using different semantic categories (e.g., faces) and/or manipulating different shape features (e.g., mid-level visual features).

There are, of course, other visual properties beyond shape. Nevertheless, shape is a very prominent alternative explanation for apparent category selectivity (Baldassi et al., 2013; Rice et al., 2014; Watson et al., 2014). Importantly, our manipulation of overall shape also introduces large variations in other properties, making it unlikely that any other visual property could explain the consistent category selectivity. Thus, category selectivity is remarkably tolerant to large variations in shape and other visual properties. We note, however, that our conclusions refer to the regional representational content, not the organizing principles driving brain representations. This distinction is very important. Indeed, our findings do not exclude the possibility that conceptual representations derive from object visual properties, semantic/functional knowledge, or a combination of both factors. On this latter point, interesting developmental studies have shown how during development, throughout experience, infants switch from visual to semantic properties to represent conceptual categories (Keil and Batterman, 1984; Keil, 1994).

These results revealed that both visual streams encode information about object category. However, dorsal and ventral representations differed significantly in terms of their category information content. In our stimulus set, regions in the ventral visual pathway mostly represent the animate/inanimate division. This result confirms the conclusion of previous studies (Kriegeskorte et al., 2008b), now with a stimulus set that dissociates this category distinction from other visual properties, and even without including exemplars from two very prominent classes of animate stimuli: human faces and bodies. Conversely, dorsal stream areas represent whether an object, regardless of its shape properties, is functionally associated with an action (musical instruments, sport articles, tools) or not (minerals, animals, fruit/vegetables). This is consistent with the proposed role of object representations in the dorsal stream. Neuropsychological studies have shown that, whereas lesions in the ventral stream drastically affect object recognition (Warrington and Shallice, 1984), parietal lobe lesions impair hand-object interactions, such as the ability to manipulate objects according to their function (Buxbaum

et al., 2014). This evidence, together with functional neuroimaging studies showing, in parietal areas, selectivity for tools but not to other graspable objects (Chao and Martin, 2000; Valyear et al., 2007), suggest that object representations within the dorsal pathway might encode object functional and motor representations necessary to perform skillful actions.

As a second point, our results illustrate the distribution and transformations in how shape is represented. Shape representations were remarkably widespread throughout the ventral visual pathway, including more anterior occipitotemporal regions (note that anterior temporal areas were not included in our ROIs), where shape and category information coexisted. Within the shape domain, we observed a progression of shape representations from early visual areas to high-level visual areas, whereas representational content in more posterior areas (e.g., BA 17/18) was best predicted by the silhouette model; higher up in high-level visual areas, representations reflected perceived shape similarities (Fig. 6). The properties of high-level shape representations were in line with earlier reports (Eger et al., 2008) showing an increased degree of orientation invariance (all elongated objects cluster together despite differences in orientation). These results confirm evidence from multivariate analyses obtained earlier with artificial objects (Haushofer et al., 2008; Op de Beeck et al., 2008b) and show that these earlier conclusions hold for images of real, everyday objects. Perceived shape similarity of familiar categories is represented in many category-selective regions.

As a third point, our results suggest a relationship between shape and category representations. The degree to which information content in shape-sensitive ROIs reflected perceived rather than low-level shape properties was closely related to the degree to which that same ROI encoded category information (Fig. 7). These results suggest that object shape representations in high-level visual cortex might be influenced by the interaction with object semantic knowledge, or vice versa. This perspective differs from the point of view taken in most existing literature. Many studies have argued in favor of one dimension, category selectivity (Kanwisher, 2010), or particular visual properties (Ishai et al., 1999). Our findings indicate that either dimension cannot be explained by the other. Furthermore, these dimensions show interesting associations that might inform on why many properties coexist. Op de Beeck et al. (2008a) already suggested that category selectivity might be based upon the coincidence of multiple features, each of which might be correlated to some degree with category distinctions. Many of such feature maps have been demonstrated, including eccentricity biases (Levy et al., 2001; Hasson et al., 2002), curvature/shape (Brincat and Connor, 2004; Yamane et al., 2008), and spatial frequency (Rajimehr et al., 2011). Nevertheless, category selectivity cannot be reduced to a simple linear combination of these features. As the current study shows, category selectivity is robust even when the other features are not consistently associated with category membership or even manipulated independently from category membership. Instead, as suggested by the found association between perceived shape sensitivity and category sensitivity, it might be worthwhile to reverse our viewpoint and question to what extent can we understand the existing feature maps by assessing their usefulness for categorizing objects.

Indeed, the observed feature maps and their relationships are difficult to understand without taking into account the relationship to category representations. As far as our results are concerned, the goal of recognizing and categorizing objects might be an important factor to understand the transition from low-level

to high-level shape representations. As another example, the bias of face-selective regions to prefer curved objects, lower spatial frequencies, and foveally presented stimuli is hard to explain without resorting to the concept of a face (Op de Beeck et al., 2008a). In early visual cortex, neurons that prefer foveal stimuli typically process higher spatial frequencies, not lower spatial frequencies. Thus, why would one and the same region prefer lower spatial frequencies and foveally presented stimuli if not for the fact that face recognition typically involves the processing of lower spatial frequencies of foveated faces? Thus, to understand how and why the functional organization for such visual properties is correlated with category selectivity, we might have to consider the association of these properties with category information throughout everyday visual experience.

In conclusion, our results provide a significant advance in the debate of to what extent object shape and object category underlie the functional organization of object representations in visual cortex. We created a stimulus set that allowed disentangling the shape and the category dimension. Notably, our results show that object category representations in both visual pathways cannot be reduced to object shape properties. At the same time, shape and category information interacts throughout the visual hierarchy shaping object perception in a fundamental way, ultimately leading to successful object recognition.

References

- Arcaro MJ, McMains SA, Singer BD, Kastner S (2009) Retinotopic organization of human ventral visual cortex. *J Neurosci* 29:10638–10652. [CrossRef Medline](#)
- Baldassi C, Alemi-Neissi A, Pagan M, Dicarlo JJ, Zecchina R, Zoccolan D (2013) Shape similarity, better than semantic membership, accounts for the structure of visual object representations in a population of monkey inferotemporal neurons. *PLoS Comput Biol* 9:e1003167. [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. [CrossRef Medline](#)
- Biederman I (1987) Recognition-by-components: a theory of human image understanding. *Psychol Rev* 94:115–147. [CrossRef Medline](#)
- Bracci S, Ietswaart M, Peelen MV, Cavina-Pratesi C (2010) Dissociable neural responses to hands and non-hand body parts in human left extrastriate visual cortex. *J Neurophysiol* 103:3389–3397. [CrossRef Medline](#)
- Brainard DH (1997) The Psychophysics Toolbox. *Spat Vis* 10:433–436. [CrossRef Medline](#)
- Brincat SL, Connor CE (2004) Underlying principles of visual shape selectivity in posterior inferotemporal cortex. *Nat Neurosci* 7:880–886. [CrossRef Medline](#)
- Buxbaum LJ, Shapiro AD, Coslett HB (2014) Critical brain regions for tool-related and imitative actions: a componential analysis. *Brain* 137:1971–1985. [CrossRef Medline](#)
- Chao LL, Martin A (2000) Representation of manipulable man-made objects in the dorsal stream. *Neuroimage* 12:478–484. [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](#)
- Eger E, Ashburner J, Haynes JD, Dolan RJ, Rees G (2008) fMRI activity patterns in human LOC carry information about object exemplars within category. *J Cogn Neurosci* 20:356–370. [CrossRef Medline](#)
- Goodale MA, Milner AD (1992) Separate visual pathways for perception and action. *Trends Neurosci* 15:20–25. [CrossRef Medline](#)
- Grill-Spector K, Malach R (2004) The human visual cortex. *Annu Rev Neurosci* 27:649–677. [CrossRef Medline](#)
- Grill-Spector K, Weiner KS (2014) The functional architecture of the ventral temporal cortex and its role in categorization. *Nat Rev Neurosci* 15:536–548. [CrossRef Medline](#)
- Grill-Spector K, Kushnir T, Hendler T, Edelman S, Itzhak Y, Malach R (1998) A sequence of object-processing stages revealed by fMRI in the human occipital lobe. *Hum Brain Mapp* 6:316–328. [CrossRef Medline](#)
- Grill-Spector K, Kushnir T, Edelman S, Avidan G, Itzhak Y, Malach R (1999) Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* 24:187–203. [CrossRef Medline](#)
- Hasson U, Levy I, Behrmann M, Hendler T, Malach R (2002) Eccentricity bias as an organizing principle for human high-order object areas. *Neuron* 34:479–490. [CrossRef Medline](#)
- Haushofer J, Livingstone MS, Kanwisher N (2008) Multivariate patterns in object-selective cortex dissociate perceptual and physical shape similarity. *PLoS Biol* 6:e187. [CrossRef Medline](#)
- Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293:2425–2430. [CrossRef Medline](#)
- Hubel DH (1963) The visual cortex of the brain. *Sci Am* 209:54–62. [CrossRef Medline](#)
- Ishai A, Ungerleider LG, Martin A, Schouten JL, Haxby JV (1999) Distributed representation of objects in the human ventral visual pathway. *Proc Natl Acad Sci U S A* 96:9379–9384. [CrossRef Medline](#)
- James TW, Humphrey GK, Gati JS, Menon RS, Goodale MA (2002) Differential effects of viewpoint on object-driven activation in dorsal and ventral streams. *Neuron* 35:793–801. [CrossRef Medline](#)
- Kanwisher N (2010) Functional specificity in the human brain: a window into the functional architecture of the mind. *Proc Natl Acad Sci U S A* 107:11163–11170. [CrossRef Medline](#)
- Kayaert G, Biederman I, Vogels R (2003) Shape tuning in macaque inferior temporal cortex. *J Neurosci* 23:3016–3027. [Medline](#)
- Keil FC (1994) Explanation based constraints on the acquisition of word meaning. In: *The acquisition of the Lexicon* (Gleitman L, Landau B, eds), pp 169–196. Cambridge, MA: Massachusetts Institute of Technology.
- Keil FC, Batterman N (1984) A characteristic-to-defining shift in the acquisition of word meaning. *J Verbal Learn Verbal Behav* 23:221–236. [CrossRef](#)
- 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](#)
- Kobatake E, Tanaka K (1994) Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *J Neurophysiol* 71:856–867. [Medline](#)
- Konen CS, Kastner S (2008) Two hierarchically organized neural systems for object information in human visual cortex. *Nat Neurosci* 11:224–231. [CrossRef Medline](#)
- Konkle T, Caramazza A (2013) Tripartite organization of the ventral stream by animacy and object size. *J Neurosci* 33:10235–10242. [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 (2012) Inverse MDS: inferring dissimilarity structure from multiple item arrangements. *Front Psychol* 3:245. [CrossRef Medline](#)
- Kriegeskorte N, Mur M, Bandettini P (2008a) Representational similarity analysis: connecting the branches of systems neuroscience. *Front Syst Neurosci* 2:4. [CrossRef Medline](#)
- Kriegeskorte N, Mur M, Ruff DA, Kiani R, Bodurka J, Esteky H, Tanaka K, Bandettini PA (2008b) Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron* 60:1126–1141. [CrossRef Medline](#)
- Levy I, Hasson U, Avidan G, Hendler T, Malach R (2001) Center-periphery organization of human object areas. *Nat Neurosci* 4:533–539. [Medline](#)
- Nasr S, Liu N, Devaney KJ, Yue X, Rajimehr R, Ungerleider LG, Tootell RB (2011) Scene-selective cortical regions in human and nonhuman primates. *J Neurosci* 31:13771–13785. [CrossRef Medline](#)
- Nasr S, Echavarria CE, Tootell RB (2014) Thinking outside the box: rectilinear shapes selectively activate scene-selective cortex. *J Neurosci* 34:6721–6735. [CrossRef Medline](#)
- Op de Beeck HP (2010) Against hyperacuity in brain reading: spatial smoothing does not hurt multivariate fMRI analyses? *Neuroimage* 49:1943–1948. [CrossRef Medline](#)
- Op de Beeck HP, Haushofer J, Kanwisher NG (2008a) Interpreting fMRI data: maps, modules and dimensions. *Nat Rev Neurosci* 9:123–135. [CrossRef Medline](#)
- Op de Beeck HP, Torfs K, Wagemans J (2008b) Perceived shape similarity

- among unfamiliar objects and the organization of the human object vision pathway. *J Neurosci* 28:10111–10123. [CrossRef Medline](#)
- Peelen MV, Downing PE (2005) Selectivity for the human body in the fusiform gyrus. *J Neurophysiol* 93:603–608. [CrossRef Medline](#)
- Puce A, Allison T, Asgari M, Gore JC, McCarthy G (1996) Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. *J Neurosci* 16:5205–5215. [Medline](#)
- Rajimehr R, Devaney KJ, Bilenko NY, Young JC, Tootell RB (2011) The “parahippocampal place area” responds preferentially to high spatial frequencies in humans and monkeys. *PLoS Biol* 9:e1000608. [CrossRef Medline](#)
- Rice GE, Watson DM, Hartley T, Andrews TJ (2014) Low-level image properties of visual objects predict patterns of neural response across category-selective regions of the ventral visual pathway. *J Neurosci* 34:8837–8844. [CrossRef Medline](#)
- Rosch E, Mervis CB, Gray WD, Johnson DM, Boyes-Braem P (1976) Basic objects in natural categories. *Cogn Psychol* 8:382–439.
- Sakuraba S, Sakai S, Yamanaka M, Yokosawa K, Hirayama K (2012) Does the human dorsal stream really process a category for tools? *J Neurosci* 32:3949–3953. [CrossRef Medline](#)
- Schwarzlose RF, Baker CI, Kanwisher N (2005) Separate face and body selectivity on the fusiform gyrus. *J Neurosci* 25:11055–11059. [CrossRef Medline](#)
- 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.
- Van Essen DC (2005) A population-average, landmark- and surface-based (PALS) atlas of human cerebral cortex. *Neuroimage* 28:635–662. [CrossRef Medline](#)
- Warrington EK, Shallice T (1984) Category specific semantic impairments. *Brain* 107:829–854. [CrossRef Medline](#)
- Watson DM, Hartley T, Andrews TJ (2014) Patterns of response to visual scenes are linked to the low-level properties of the image. *Neuroimage* 99:402–410. [CrossRef Medline](#)
- Yamane Y, Carlson ET, Bowman KC, Wang Z, Connor CE (2008) A neural code for three-dimensional object shape in macaque inferotemporal cortex. *Nat Neurosci* 11:1352–1360. [CrossRef Medline](#)
- Yue X, Pourladian IS, Tootell RB, Ungerleider LG (2014) Curvature-processing network in macaque visual cortex. *Proc Natl Acad Sci U S A* 111:E3467–E3475. [CrossRef Medline](#)