UNIVERSITY *of York*

This is a repository copy of *Multivariate Patterns in the Human Object-Processing Pathway Reveal a Shift from Retinotopic to Shape Curvature Representations in Lateral Occipital Areas, LO-1 and LO-2*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/100747/

Version: Published Version

Article:

White Rose
university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Behavioral/Cognitive

Multivariate Patterns in the Human Object-Processing Pathway Reveal a Shift from Retinotopic to Shape Curvature Representations in Lateral Occipital Areas, LO-1 and LO-2

Richard J. W. Vernon,[1] **André D. Gouws,**[2] **Samuel J. D. Lawrence,**[1] **Alex R. Wade,**[1] **and** ⬤**Antony B. Morland**[1,2]

[1]Department of Psychology and [2]York Neuroimaging Centre, University of York, Heslington, York YO10 5DD, United Kingdom

Representations in early visual areas are organized on the basis of retinotopy, but this organizational principle appears to lose prominence in the extrastriate cortex. Nevertheless, an extrastriate region, such as the shape-selective lateral occipital cortex (LO), must still base its activation on the responses from earlier retinotopic visual areas, implying that a transition from retinotopic to "functional" organizations should exist. We hypothesized that such a transition may lie in LO-1 or LO-2, two visual areas lying between retinotopically defined V3d and functionally defined LO. Using a rapid event-related fMRI paradigm, we measured neural similarity in 12 human participants between pairs of stimuli differing along dimensions of shape exemplar and shape complexity within both retinotopically and functionally defined visual areas. These neural similarity measures were then compared with low-level and more abstract (curvature-based) measures of stimulus similarity. We found that low-level, but not abstract, stimulus measures predicted V1–V3 responses, whereas the converse was true for LO, a double dissociation. Critically, abstract stimulus measures were most predictive of responses within LO-2, akin to LO, whereas both low-level and abstract measures were predictive for responses within LO-1, perhaps indicating a transitional point between those two organizational principles. Similar transitions to abstract representations were not observed in the more ventral stream passing through V4 and VO-1/2. The transition we observed in LO-1 and LO-2 demonstrates that a more "abstracted" representation, typically considered the preserve of "category-selective" extrastriate cortex, can nevertheless emerge in retinotopic regions.

Key words: extrastriate; lateral occipital; object selective; representational similarity analysis; shape; V1

Significance Statement

Visual areas are typically identified either through retinotopy (e.g., V1–V3) or from functional selectivity [e.g., shape-selective lateral occipital complex (LOC)]. We combined these approaches to explore the nature of shape representations through the visual hierarchy. Two different representations emerged: the first reflected low-level shape properties (dependent on the spatial layout of the shape outline), whereas the second captured more abstract curvature-related shape features. Critically, early visual cortex represented low-level information but this diminished in the extrastriate cortex (LO-1/LO-2/LOC), in which the abstract representation emerged. Therefore, this work further elucidates the nature of shape representations in the LOC, provides insight into how those representations emerge from early retinotopic cortex, and crucially demonstrates that retinotopically tuned regions (LO-1/LO-2) are not necessarily constrained to retinotopic representations.

Introduction

Visual areas of the human brain can be identified by their retinotopic representations of the visual field (Engel et al., 1994; Wandell et al., 2007) or by their selectivity to stimulus categories (Malach et al., 1995; Kanwisher et al., 1997; Epstein and Kanwisher, 1998; Downing et al., 2001). These differing approaches persist primarily because category-selective areas appear to have far weaker retinotopy than early visual areas (Sayres and Grill-

Correspondence should be addressed to Prof. Antony B. Morland, Department of Psychology, University of York, Heslington, York YO10 5DD, UK. E-mail: antony.morland@york.ac.uk.
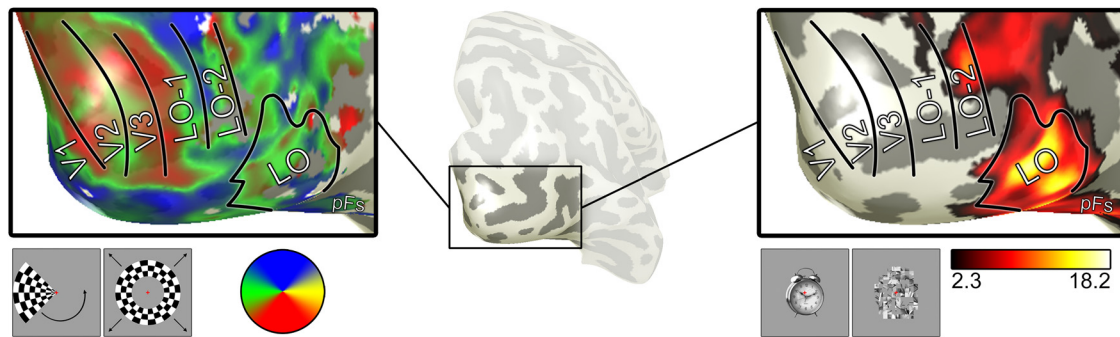
Figure 1. Identification of visual areas. Middle, The occipital area of interest (highlighted) from one representative participant, which is magnified in the panels to the left and right. In the left, the visual field representation of the cortex is shown (see color key below). Boundaries between visual areas are shown in black and correspond to the representations of the horizontal and vertical meridians. Data are given for responses to a rotating wedge (as inset below the data). Also shown is the outline of LO and general location of pFs, which is derived from the functional localizer of an object versus the functional localizer of scrambled objects. Data here are Z-score thresholded at Z = 2.3, but LO itself was defined using a sphere surrounding peak voxel (see Materials and Methods).

Spector, 2008), implying that such areas might map other (more "abstract") dimensions of stimulus content (Op de Beeck et al., 2008a).

There is considerable interest in the relationship between early, retinotopic representations and representations in later extrastriate cortex (Levy et al., 2001; Hasson et al., 2002; Malach et al., 2002; Larsson and Heeger, 2006; Hemond et al., 2007; Sayres and Grill-Spector, 2008; Schwarzlose et al., 2008; Arcaro et al., 2009; Grill-Spector and Weiner, 2014). We explored how topographic representations of the visual field shift to representations of more abstract stimulus properties and whether these organizational principles are mutually exclusive. Note that "abstract" is used throughout this paper to encompass representations that are abstracted from retinotopic representations predominating in earlier visual areas. We specifically investigated "curvature complexity" as a potential abstract feature that extrastriate areas may be tuned to, asking where in the visual hierarchy such representations might emerge.

It is well established that the lateral occipital complex (LOC) is selective for shapes over stimuli without coherent form (Malach et al., 1995; Grill-Spector et al., 2001). This region has been divided into the more posterior/dorsal lateral occipital cortex (LO) and the more anterior/ventral posterior fusiform sulcus (pFs), but their underlying organization remains a source of debate (Op de Beeck et al., 2008a). Multiple studies find that activity patterns within the LO correspond to similarity between shape features. For example, Op de Beeck et al. (2008b) found that perceptual shape similarity (cued predominantly by shape features, e.g., protrusion spikiness) corresponded with LO activity. Similarly, patterns of activity in the LO correlate with measures of aspect ratio and skew (Haushofer et al., 2008b) and the prominence of protrusions in radial frequency patterns (Drucker and Aguirre, 2009). Generally, these studies found that strictly physical similarity measures (e.g., pixelwise similarity) were poor predictors of LO activity, implying a move from retinotopic organization toward something more abstract.

This transition is unlikely to be abrupt. Retinotopy is the currency of early visual cortex, so retinotopy must ultimately be the foundation for higher-level representations (Peirce, 2015). This implies the existence of an intermediate stage in which retinotopic and more abstract representations coexist. We asked where this point might lie.

Two candidate regions for such a point are the retinotopically defined visual field maps LO-1 and LO-2 (Larsson and Heeger, 2006), lying nearby or overlapping with LO (Sayres and Grill-

Spector, 2008). LO-2 in particular responds well to objects (Fig. 1) and is causally involved in shape-processing tasks (Silson et al., 2013), implying some similarity with LO. Therefore, this lateral occipital aspect of the human brain offers an ideal site to explore how organizational principles transition from retinotopic to abstract.

We hypothesized that representations in V1–V3 would be driven strongly by low-level (retinotopic) similarity, whereas LO should be driven by more abstract measures of shape curvature. For LO-1 and LO-2 (lying between V3d and LO), their full hemifield representations imply retinotopic tunings, yet their overlap with LO implies more abstracted representations may also exist. Therefore, we hypothesized that a retinotopic-to-abstract transition may occur near or in LO-1 and LO-2.

To test our hypotheses, we created stimulus sets differing along dimensions of shape exemplar (e.g., bird/cat) and level of shape detail (i.e., how complex the shape outline is). BOLD responses to each of our stimuli were recorded under a rapid event-related fMRI design, and measures of neural similarity were extracted for the visual areas discussed above. These neural similarity measures could then be compared with various stimulus similarity measures to reveal which stimulus features are most salient for each visual area.

Materials and Methods

Participants
Twelve participants (mean ± SD age, 25.42 ± 4.78 years; eight males; all gave informed consent) were recruited from the University of York Department of Psychology. Each participant underwent one high-resolution structural scanning session, one retinotopic-mapping session, one localizer session, and two main functional sessions, totaling 5.25 h scanning per participant. In addition, all participants performed two 30 min behavioral sessions.

Preliminary data acquisition and analysis
All imaging data were acquired using a GE Healthcare 3 Tesla Sigma HD Excite scanner and a 16-channel head coil to improve signal-to-noise in the occipital lobe.

Structural scans. We acquired three 16-channel T1-weighted anatomical images (TR, 7.8 ms; TE, 3.0 ms; TI, 600 ms; voxel size, 1 × 1 × 1 mm³; flip angle, 12°; matrix size, 256 × 256 × 176; FOV, 25.6 cm), one eight-channel T1-weighted anatomical image to aid alignments (TR, 7.8 ms; TE, 2.9 ms; TI, 450 ms; voxel size, 1.13 × 1.13 × 1 mm³; flip angle, 20°; matrix size, 256 × 256 × 176; FOV, 29 cm), and one T2*-weighted fast gradient recalled echo scan (TR, 400 ms; TE, 4.3 ms; voxel size, 1 × 1 × 2 mm³; flip angle, 25°; matrix size, 128 × 128; FOV, 26 cm).

The three 16-channel T1 scans were aligned and averaged and then divided by the T2*-weighted data. This improved gray–white matter

contrast and partially corrected for the signal drop-off caused by use of a 16-channel coil. This average T1 was then automatically segmented into gray and white matter with manual refinements.

Retinotopic mapping scans. We collected six wedge scans plus two ring scans in each participant (TR, 3000 ms; TE, 30 ms; voxel size, $2 \times 2 \times 2$ mm^3; flip angle, 90°; matrix size, $96 \times 96 \times 39$; FOV, 19.2 cm). Wedges were 90° in size and rotated counterclockwise about a red fixation cross. Ring stimuli expanded about fixation. Both wedges and rings were high-contrast ($>98\%$, 400 cdm^{-2}) checkerboard stimuli that flickered at a rate of 6 Hz. Each scan contained eight cycles of wedges/rings, with 36 s per cycle, traversing a circular region of radius 14.34°. Participants maintained fixation throughout the scan.

Retinotopy data were analyzed using standard techniques (Wandell et al., 2007) as specified previously (Baseler et al., 2011). Regions of interest (ROIs) V1–V4, LO-1, and LO-2 (Larsson and Heeger, 2006) were identified in both hemispheres of all participants.

LOC localizer scans. Three 8-min localizer scans (using identical imaging parameters to those used in the retinotopic scans) followed an ABAB block design contrasting objects with scrambled objects. In total, 16 object blocks and 16 scrambled object blocks were used per scan (15 s blocks) with one image presented per second (0.8 s presentation, 0.2 s interstimulus interval). Participants maintained fixation on a central red cross while performing a one-back task in which there could be one, two, or no repeats within a given block (to ensure that attention was maintained). All stimuli were presented centrally on a full-screen mid-gray background (200 cdm^{-2}), and there were no baseline/rest periods between blocks.

For stimuli, we used 225 PNG images of easily recognizable objects manually extracted from their original backgrounds. These were converted to grayscale with a flattened (equalized) image histogram. Image size was set to subtend $4 \times 4°$ visual angle on average (exact size depended on image aspect ratio). To create scrambled stimuli, we split the objects and background into a grid with 20 rows and columns (square size $\sim 0.8 \times 0.8°$), and then all squares lying within the convex hull of the object were randomly permuted and rotated. This meant scrambled objects would contain all local details from the original objects, plus the same coarse outline, but would not be semantically recognizable. Because scrambling introduced sharp contrast edges between permuted squares, we applied a Gaussian filter (SD of 1 pixel) to both the objects and scrambled objects.

Localizer data were analyzed using FEAT (FMRI Expert Analysis Tool; Worsley, 2001). At the first (individual) level, we removed the first three volumes and used a high-pass filter cutoff point of 60 s to correct for low-frequency drift. Spatial smoothing was performed with a Gaussian kernel of 4 mm FWHM, and FILM prewhitening was used. To combine data within a participant, we ran fixed-effects analysis with cluster correction ($Z > 2.3$, $p < 0.05$), and then we defined LO and pFs using a method partially based on that proposed by Julian et al. (2012).

First, significant activation within each participant was binarized and linearly transformed into standard (MNI 152 T1 2 mm) space. To identify the "average" activation, the data were summed, spatially smoothed (Gaussian filter with 4 mm FWHM), and then divided by the number of participants (12). It was then thresholded at 0.6 to identify voxels in which 60% of participants show significant activation. The thresholded activation in each hemisphere was then manually bisected into LO and pFs masks, based primarily on anatomical location. These masks were then back-transformed into each participant's individual space. Finally, for each of the left and right hemisphere LO and pFs masks in each participant, we selected all active voxels lying within a sphere (10 mm radius) centered on that participant's peak voxel within the respective region. Using this method, we ensured that all participants had approximately the same number of voxels in their left and right LO and pFs ROIs (desirable for the multivoxel pattern analysis described below) and that their ROIs were all in the same approximate anatomical location. This approach also has the secondary advantage of reducing overlap between LO-2 and the posterior parts of LO (Table 1), allowing conclusions to be drawn for each region independently.

Table 1. Percentage overlap between LO-1, LO-2, and LO localizer data across all participants

Participant	Sphere LO ROI		Cluster corrected		Voxel corrected	
	LO-1	LO-2	LO-1	LO-2	LO-1	LO-2
P01	0	2.1	0.5	19.3	0	6.2
P02	0	2.7	56.3	99.5	35.4	98.7
P03	0	0	3.0	16.4	0.8	12.4
P04	0	0	9.0	61.0	3.5	30.2
P05	0	13.3	38.8	95.5	8.9	87.5
P06	0.4	5.6	44.6	89.6	31.3	74.8
P07	4.3	22.8	33.3	70.0	22.4	49.3
P08	0	0.4	8.5	62.9	1.9	40.0
P09	0	0	31.5	56.4	7.9	38.4
P10	0	0	15.5	86.6	5.0	75.5
P11	0	18.1	16.8	75.5	5.0	57.3
P12	0	0	0	2.1	0	0.1
Average	2.4	9.3	23.4	61.2	12.2	47.5
SD	2.8	8.8	18.5	32.6	12.7	32.3

We calculated the percentage of voxels in LO-1 and LO-2 that overlapped with LO or LOC localizer activity (objects > scrambled objects) under various conditions. The first "Sphere LO ROI" is the final LO ROI used for the purposes of this study, created using a sphere (10 mm radius) centered on a peak voxel (see Materials and Methods, LOC localizer scans). The second condition, "Cluster corrected," compares the percentage of voxels in LO-1 and LO-2 with the underlying activity on which the sphere ROI was created (i.e., cluster-corrected LOC localizer data, Z-score thresholded at $Z = 2.3$, $p < 0.05$). The final condition, "Voxel corrected," is an additional analysis for comparative purposes. We thresholded LOC localizer data on a voxel-by-voxel basis ($p < 10^{-5}$, uncorrected for multiple comparisons), a common approach in the literature (Fig. 2), and again compared percentage overlap with LO-1 and LO-2 voxels.

Main functional scans: stimuli

To investigate the cortical representations of shapes, we developed a (standard) stimulus set comprising three exemplars (animal outlines), taken from the set of Snodgrass and Vanderwart (1980) images that had been converted to silhouettes and rated for recognizability by De Winter and Wagemans (2004). The exemplars were filtered on the basis of their Fourier descriptor (FD) content (Zahn and Roskies, 1972) to give three levels of detail (low, mid, high), yielding nine stimuli (Fig. 2). More "detailed" shapes contain more frequency information, and so they have greater variation in curvature around their perimeters. As such, this can also be thought of as a manipulation in "curvature complexity." By altering the phase of one FD (see the methods below), we also created "scrambled" versions to remove any semantic associations, which has been raised as a potential confound in previous literature (Haushofer et al., 2008b). These scrambled stimuli were unrecognizable but matched to the "standard" stimuli in FD content (Fig. 2).

We calculated the set of FDs for a given image using the following procedure. The outermost boundary of the shape was extracted, and moving average smoothing was applied, correcting for pixelation. The resultant smoothed boundary was then interpolated using a periodic cubic spline curve to 4096 points. For every point around the contour, we then calculated both the distance and angle to the next point before normalizing distance to the range $0–2\pi$ and removing the linear trend in the angles. By performing Fourier analysis on the set of angles, we can create our set of FDs. Critically, we save out information to make this process completely reversible, so shapes can be manipulated in the FD and transformed back to view the results. We removed linear trends from the resultant x and y coordinates to prevent nonclosed boundaries.

To select three exemplars from the set of Snodgrass and Vanderwart (1980) images, we identified shapes that (1) exhibited a smooth, exponential decay of "power" as a function of the number of FDs (as estimated by the norm of the residuals from a fitted exponential decay curve), (2) had high recognizability (ratings from De Winter and Wagemans, 2004), and (3) had biologically plausible shapes with relatively distinct profiles (i.e., animals). Using these criteria, we selected three shapes, specifically a bird, a cat, and a rabbit, for each of these stimuli, and then we identified a "target FD," which was a lenient estimate of the number of FDs needed to accurately reproduce the shape.

We next aimed to render our stimuli at low, mid, and high levels of detail. The level of shape detail was manipulated by using a filter in the FD based on a Gaussian with a half-width at half-maximum (HWHM) that controlled the FD content of the shape. The filter was originally specified

to have a maximum height of 2 but was subsequently clipped such that its value was unity for all values that originally exceeded 1.

For the three detail levels, we chose not to match the stimuli in terms of the number of FDs, because the number of FDs needed to describe a shape can vary somewhat arbitrarily. Instead, we matched detail across stimuli in terms of relative amplitude. First, for our three shapes, we calculated the summed amplitude of the FDs (after normalizing it with the DC component) from one FD to the target FD defined above and then fitted a curve of form $y = a(1-e^{-bx})$ to the data (x indicates HWHM of Gaussian filter; y indicates the sum of amplitude spectrum). We took the summed amplitude using our target FD and, from our fitted curve, interpolated the HWHM needed to get ¼, ½, and ¾ of this value to create low, mid, and high detail boundaries (specifically, the HWHM values needed were 3.09/2.45/2.48, 7.16/5.60/5.85, and 13.14/10.02/11.14 for the low/mid/high detail bird/cat/rabbit, respectively).

The above process was used to create our standard stimuli. For the scrambled stimuli, we performed the same procedure, except that the phase of the FD with most power in the low complexity shapes was rotated counterclockwise through 90° (Fig. 2). For the cat and rabbit, this was FD 2, and for the bird, it was FD 3. This manipulation meant that, at low complexity, the shapes mostly changed through a rotation, whereas at high complexity, interactions with higher frequency FDs caused our shapes to be completely unrecognizable. Critically, these scrambled stimuli share the same FDs (albeit with one phase change) as our standard stimuli.

The area of each stimulus was matched to the area of a square with a length of 6° of visual angle. The profile of each shape outline was then rendered as the fourth derivative of a Gaussian (Wilkinson et al., 1998) at 50% contrast, yielding a peak spatial frequency of 1.68 cycles/°.

Main functional scans: data acquisition and analysis
Both the standard and scrambled functional scan sessions comprised five 8.5 min stimulus presentation runs (TR, 2000 ms; TE, 30 ms; voxel size, $2 \times 2 \times 2.5$ mm^3; flip angle, 90°; matrix size, $96 \times 96 \times 26$; FOV, 19.2 cm). We used a rapid event-related design, in which the stimulus presentation order had been counterbalanced and optimized (jittered) using Optseq2 (https://surfer.nmr.mgh.harvard.edu/optseq). Five stimulus presentation sequences were generated, one per run, and were used in order of most-to-least efficient (to ensure that most efficient runs were used when participants were most alert). Stimulus presentation order was identical between the standard and scrambled scans, and 108 stimuli were presented per run (12 of each stimulus). Stimuli were presented centrally (centered on shape centroid) for 0.6 s with a median interstimulus interval of 3.4 s (interquartile range, 2.4–6.4 s; range, 1.4–18.4 s).

Participants maintained fixation on a red cross (0.60°) and performed a one-back task (there were on average 10.6 sequential repeats per run). We added 10 s to the start of each scan to allow the magnetization to reach a steady state, and for all but the first two of the participant's standard scan sessions, we appended 20 s to the end of the scan. This ensured that we were capturing the complete hemodynamic response for the final stimuli presented. During each scan, we recorded video of the participant's right eye and later extracted eyeblinks using custom-written software.

Data were analyzed using FEAT (Worsley, 2001). The first five volumes of each run were discarded (ensuring the scanner had reached stable magnetization), the high-pass filter cutoff point was set to 100 s (correcting for low-frequency drift), FILM prewhitening was used, and motion was corrected. Motion parameters were also entered as confound covariates. As in the study by Op de Beeck et al. (2008b), we applied spatial smoothing (Gaussian kernel with FWHM of 4 mm; twice voxel
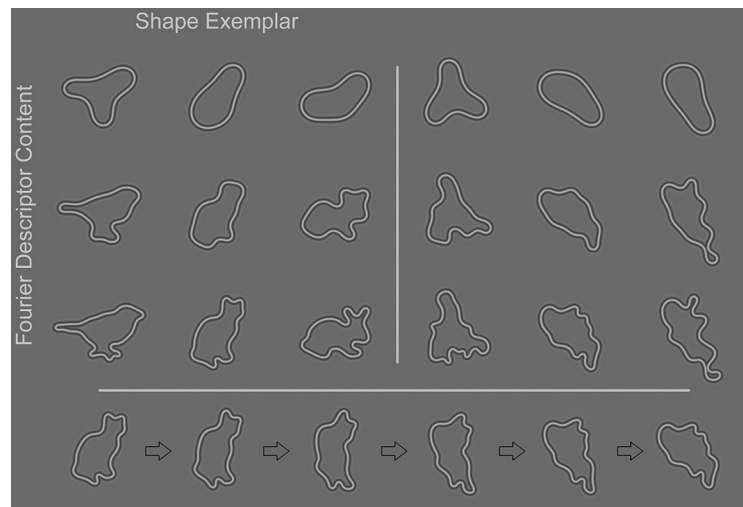

Figure 2. Stimuli used in the study. The standard and scrambled stimuli are shown on the top left and top right, respectively. Each column represents one exemplar, and each row represents one level of FD-Content (low, mid, and high detail; see Materials and Methods). The scrambled stimuli were created by rotating the phase of the FD with the most power through 90° counterclockwise, as shown for one stimulus (in incremental phase rotations of 18°) along the bottom row.

size). All nine stimuli were entered as separate explanatory variables (EVs). Blinks were also added as an EV (modeled as a 200 ms event from start of blink) because they can be a potential source of noise (Hupé et al., 2012; Gouws et al., 2014). Contrasts were set up to compare each individual stimulus to baseline plus one contrast comparing the activity of all stimuli over baseline. To combine runs within a participant, we ran fixed-effects analysis using cluster correction ($Z > 2.3$, $p < 0.05$), and all data were retained in the high-resolution structural space.

Percentage signal change was calculated on a voxel-by-voxel basis based on the methods suggested by Mumford (2007).

ROI restriction
All ROIs were restricted based on the contrast of all stimuli over baseline from the functional scans of the opposite study (i.e., analysis using the standard stimulus set will use ROIs constrained using activity from the scrambled stimulus set). This was primarily performed because our stimuli only occupied a relatively small part of the visual field. As such, we would expect a smaller proportion of voxels in earlier visual areas such as V1 (with small receptive field sizes) to respond to our stimuli, in contrast to later areas such as LO-2 with receptive field sizes that respond to much larger regions of the visual field (Dumoulin and Wandell, 2008; Amano et al., 2009); therefore, constraining ROIs helps to control for such differences. We used the activation from the opposite study to constrain ROIs because the stimuli in both studies occupy approximately the same spatial extent. This method also avoids concerns of circularity (Kriegeskorte et al., 2009) and does not cause the loss of power inherent in alternative approaches (e.g., split-half analysis).

Correlation analysis and shape similarity metrics
Our main (correlation) analysis explored the patterns of activity elicited by our various shape stimuli and attempted to predict those patterns. We first created neural similarity matrices that captured for a given ROI the similarity (in terms of the pattern of neural activity) between all pairwise combinations of our shape stimuli. To achieve this, we took the whole-brain activation elicited by each of our nine stimuli (in units of percentage signal change), and then for every ROI collapsed (concatenated) across hemisphere (V1-V4, LO-1, LO-2, LO, pFs), we extracted the pattern of activity specific to that area. We then iterated over all (36) pairwise combinations of our shapes and correlated the patterns of activity in each ROI independently (i.e., for each pairwise combination, we correlated the activity elicited by the two shapes in that pair). In this way, higher correlations indicate which sets of stimuli elicited more similar patterns of activity in each of our ROIs.

To explore why certain stimuli elicit more similar patterns of activity in each region, the neural similarity matrices were then further correlated

with a variety of different stimulus similarity measures, split into three broad categories. First, a perceptual measure was captured in a behavioral session completed after the scan (described below). Second, we had three low-level measures intended to capture "retinotopic" shape similarity; these metrics depended highly on the exact location of the contours of our stimuli. The low-level predictors were as follows.

(1) Pixelwise distance between the gray values in pairs of our images. Although often a relatively crude metric, given that our shapes have been matched for size and translation, this should nevertheless capture some aspect of image overlap.

(2) GIST descriptors. Each image is convolved with set of Gabor filters at eight different orientations and four different spatial frequencies (Oliva and Torralba, 2001). These are split and averaged with a 4×4 grid. This results in 512 ($8 \times 4 \times 16$) unique values per image, which should capture the lower-level visual properties of the shape.

(3) A contour discrepancy measure, essentially Procustes analysis without matching for rotation. The "raw" contours for a given pair of shapes were matched (in a least-squares sense) for scale and translation (because this is how our final stimuli were presented to participants), we then took the average distance between all corresponding coordinates around those contours.

Finally, we had four more "abstract" measures that were intended to capture the curvature in our shapes (e.g., number and magnitude of protrusions) regardless of their spatial location. These included the following.

(1) Minima/Maxima. The number of minima and maxima (or concavities and convexities) around the contours of the shape.

(2) The number of FDs (or specifically, the HWHM of the Gaussian filter) needed to create our shapes. This is our most direct proxy for level of shape detail.

(3) Shape compactness. This was represented with the area of the shape over the area of a circle with the same perimeter as that shape. Under this definition, a circle is (intuitively) the most compact shape, and any deviations from circularity should decrease compactness.

(4) A "convex perimeter" measure. This was the perimeter of the shape over the perimeter of the convex hull of the shape. A convex hull is the smallest convex boundary that completely encapsulates the shape to which it is being fitted. As such, concavities in the profile of the shape should increase the shape perimeter but not necessarily the convex hull perimeter.

For all predictors, the final similarity metric for a given pair (Shape1–Shape2) was the Euclidean distance between the value for Shape1 and the value for Shape2. For measures returning multiple values (e.g., pixelwise distance), the metric was the average Euclidean distance between the values for Shape1/Shape2. This metric was then inverted (subtracted from zero) so that larger numbers would represent greater similarity (i.e., zero represents perfect similarity).

Perceptual similarity measure

To acquire perceptual similarity ratings for our shapes, all participants performed two behavioral sessions (one per stimulus set) at least 1 week after the corresponding functional scan. These were primarily included because Op de Beeck et al. (2008b) found that perceptual similarity predicted both LO and pFs activity, whereas Haushofer et al. (2008b) found that perceptual similarity only predicted activity in pFs. Therefore, our behavioral components aimed to address the ambiguity regarding the role of perceptual shape similarity for neural representations in LO and pFs.

The stimuli used were identical to those described above (e.g., in size/position), except that they were rendered on a 400×400 pixel background of noise (i.e., every pixel was randomly set to a value between 0 and 255). We also generated 100 noise masks. The stimuli and masks were presented in a circular aperture that smoothed out to mid-gray at the edges (a Gaussian filter with FWHM set to 90% of the diameter of the circle).

The participant's task was to rate pairs of stimuli on a 1–6 Likert scale (extremely, very, slightly dissimilar–slightly, very, extremely similar). Stimulus pairs were presented with the following timings: noise mask (200 ms), Shape1 (50 ms), noise mask (500 ms), Shape2 (50 ms), noise

mask (200 ms). The experiment was split into four main blocks. The first block was a practice trial in which all (36) pairwise combinations of shapes were presented once, to familiarize participants with the task and comparisons. The second block contained four sets, and the third and fourth blocks contained three sets of all pairwise stimuli. The experiment paused between blocks to provide a rest interval (participants pressed "space" when they were ready to continue). Within each set of pairwise stimuli, the ordering was random; the pair ordering alternated such that, if Shape1–Shape2 was a comparison in set one, set two would compare Shape2–Shape1. No comparisons were made between identical stimuli. Each behavioral session lasted ~30 min.

Multidimensional scaling

In addition to the correlation approach, we also used multidimensional scaling (MDS) on the neural similarity matrices from all ROIs across both stimulus sets. This allows us to visualize what the "shape space" of each ROI may look like. This MDS approach used PROXSCAL (Busing et al., 1997) in SPSS (IBM SPSS Statistics 20), with a weighted Euclidean or INDSCAL (Individual Differences Scaling; Carroll and Chang, 1970) model. This assumes a common dimensionality across all participants but allows for individual variability through the use of weightings (i.e., one participant may preferentially weight dimension 1 over dimension 2, whereas another participant may do the converse, but ultimately both participants use the same dimensions).

Although the MDS was primarily used for visualization purposes, we nevertheless aimed to "quantify" the resultant solutions; two different approaches were used to achieve this. First, we used a brute-force (stimulus blind) clustering approach. Specifically, for every ROI, we iterated over all possible permutations of the nine stimuli, chunking each permutation into three sets of three items. The average inter-item distance was calculated within and then across each set, and the permutation that minimized this value was taken as the clustering solution. Essentially, this method simply identifies the "best-fitting" set of three clusters (each containing three items) for all MDS solutions.

Our second approach specifically aimed to assess whether our stimulus dimensions (shape exemplar, FD-Content) were present in the extracted solutions. This analysis is analogous to the first, except that we just took the two stimulus permutations that clustered stimuli either by exemplar (bird/cat/rabbit) or level of detail (low/mid/high). For each ROI, the average inter-item distance for both permutations was calculated (as above) and then subtracted from the average distance across all pairwise combinations of shapes. If the result of this calculation was positive (i.e., greater than zero), then it would indicate some degree of clustering.

Statistical analysis

All correlations were transformed to Fisher's Z scores for averaging and statistical testing. For ANOVAs, the Greenhouse–Geisser correction was used when the assumption of sphericity had been violated (as indicated by Mauchly's test), and corrected degrees of freedom are reported. All *post hoc* tests are Bonferroni's corrected.

Results

As an initial approach, we assessed percentage signal change, as shown in Figure 3, across the low, mid, and high detailed shapes, using ROI $\times$ FD-Content repeated-measures ANOVAs.

Both standard and scrambled stimulus sets showed significant main effects of ROI (standard, $F_{(7,77)} = 4.74$, $p = 1.8 \times 10^{-4}$; scrambled, $F_{(7,77)} = 8.11$, $p = 2.3 \times 10^{-7}$) and FD-Content (standard, $F_{(2,22)} = 9.79$, $p = 0.001$; scrambled, $F_{(2,22)} = 12.54$, $p = 0.2.3 \times 10^{-4}$). There was no significant interaction in the standard stimuli ($F_{(14,154)} = 1.15$, $p = 0.318$), but a significant interaction emerged in the scrambled stimulus set ($F_{(14,154)} = 2.00$, $p = 0.021$).

To explore the significant main effects, we ran Bonferroni's-corrected *post hoc* tests across all ROIs and FD-Content levels. This revealed significantly diminished activity in pFs when compared with V3, V4, LO-1, and LO (standard, $p = 0.005$, $p = 0.022$,
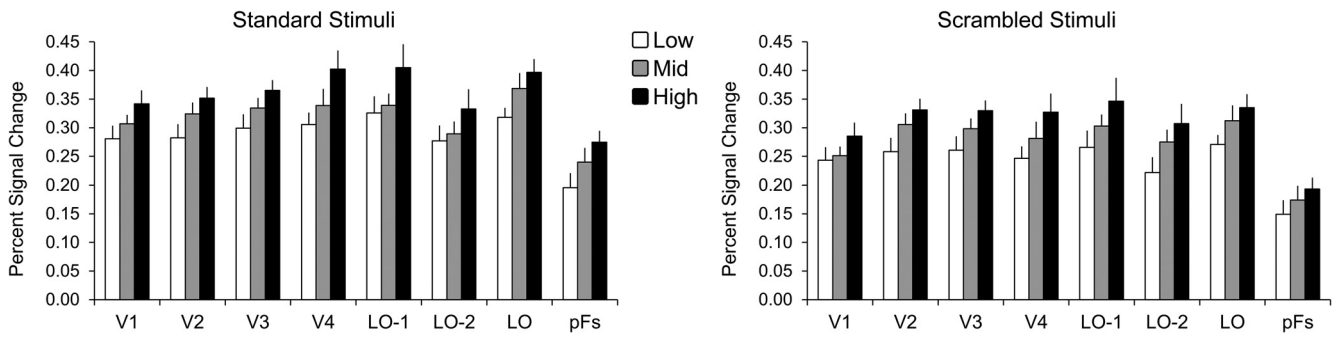
Figure 3. Percentage signal change for different brain regions. Results are shown for the standard (left) and scrambled (right) stimuli, when comparing FD-Content collapsed across exemplar. Error bars indicate the SEM.
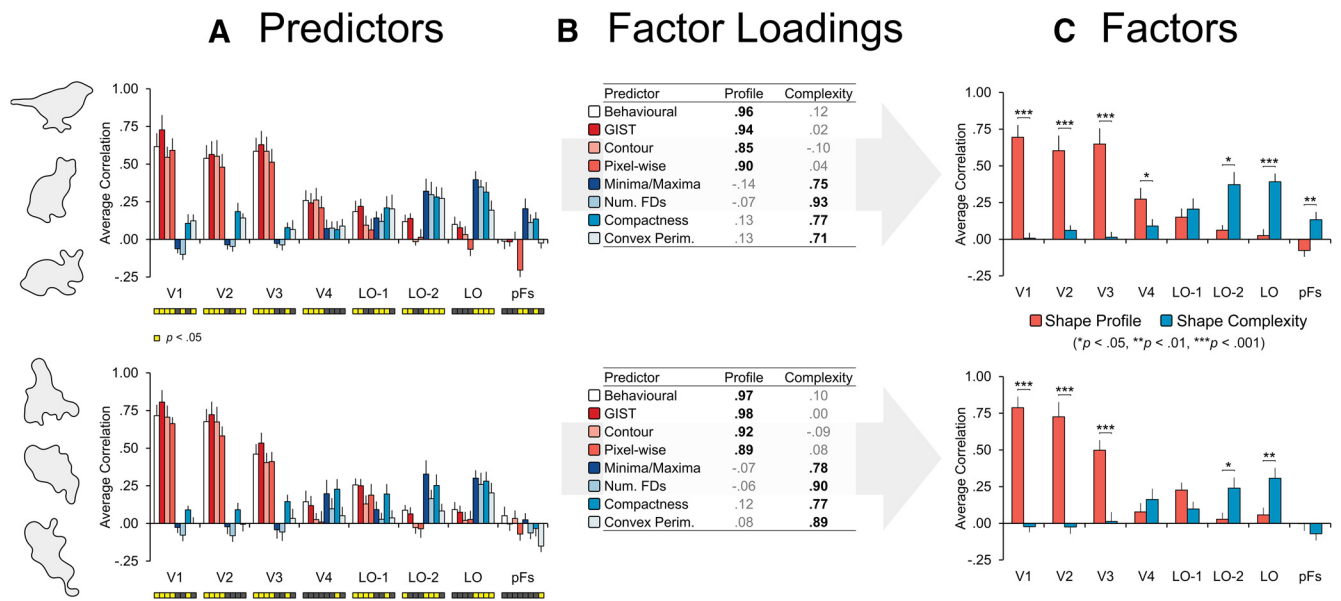

Figure 4. Correlations between predictors, factors, and neural similarity. *A*, The correlations between our predictors and neural similarity. For illustrative purposes, bars that differ significantly from zero (two-tailed 1-sample *t* tests, uncorrected for multiple comparisons) are highlighted below each graph. *B*, The loadings of these predictors in the two factors (Shape-profile, Shape-complexity) derived from PCA. *C*, The correlations between those factors and the neural similarity matrices. The top and bottom rows of panels correspond to data obtained with standard and scrambled stimuli, respectively. The middle tables also provide a color key to the predictors plotted in the left panels. In the rightmost panel, significance is derived from (uncorrected) two-tailed paired samples *t* tests, *p < 0.05, **p < 0.01, ***p < 0.001. Error bars indicate the (conventional) SEM.

$p = 0.039$, and $p = 0.001$ respectively; scrambled, $p = 0.003$, $p = 0.006$, $p = 0.003$, and $p = 4.6 \times 10^{-4}$, respectively), plus V1 and V2 in the scrambled stimulus set ($p = 0.049$, $p = 9.8 \times 10^{-5}$, respectively). No other differences were significant (all p values > 0.089). For FD-Content, high detailed shapes elicited significantly greater activity than low detailed shapes (standard, $p = 0.002$; scrambled, $p = 0.002$), but there were no significant differences between low and mid (standard, $p = 221$; scrambled, $p = 0.080$) or mid and high (standard, $p = 136$; scrambled, $p = 0.056$) detailed shapes.

Given the significant interaction in the scrambled stimulus set's analysis, we ran one-way repeated measures ANOVAs comparing FD-Content within each ROI. All ROIs showed significant main effects of FD-Content (all $F_{(2,22)}$ values > 5.27, all p values < 0.014), so Bonferroni's-corrected *post hoc* tests were used to compare across FD-Content levels. High detailed shapes elicited significantly greater activity than low detailed shapes in all ROIs (all p values < 0.042), plus significantly greater activity than mid detailed shapes in V4 ($p = 0.012$) and LO-1 ($p = 0.018$). Mid detailed shapes elicited significantly greater activity than low detailed shapes in LO-2 ($p = 0.025$) and broadly V2 ($p = 0.051$)

plus V3 ($p = 0.052$). No other comparisons emerged as significant (all p values > 0.068).

Correlation analysis

We next turn to our main analysis, taking the patterns of neural activity in each ROI and correlating them with our perceptual, low-level and more abstract similarity predictors. We reasoned that, if a predictor correlated well with neural similarity within a given ROI, then that predictor should inform us about the nature of the shape representation within that region.

The correlations between the similarity in activations and predictors for our ROIs are shown in Figure 4A. In general, we found that results for the perceptual measure were very similar to results for the low-level measures, and these in turn were strong predictors of activity patterns within V1–V3 and to a lesser extent V4. Conversely, our abstract measures were strong predictors of LO-2 and LO activity. In LO-1, neither low-level nor abstract measures appeared to dominate.

Prompted by the pattern of results, we asked whether our predictors could be reduced to a smaller number of dimensions. Therefore, we used principal component analysis (PCA) on the

predictors with orthogonal (varimax) rotation and found for both stimulus sets two well defined (independent) dimensions emerged (Fig. 4B, Factor Loadings). A first "Shape-profile" factor (standard, 42.70%; scrambled, 44.94% variance explained) clearly captured the perceptual and low-level measures. This appeared to be characterizing the general spatial overlap between our shape outlines. A second "Shape-complexity" factor (standard, 31.77%; scrambled, 35.25% variance explained) captured our more abstract image metrics linked to the curvature and FD-Content present in the shape outlines. We chose not to include a third factor because the first two factors already accounted for 74.48% (standard stimuli) and 80.19% (scrambled stimuli) of the variance; a third factor would have accounted for little additional variance (standard, 11.68%; scrambled, 9.92% variance explained). Furthermore, only the first two factors in both stimulus sets had eigenvalues >1 in accordance with Kaiser's (1960) criterion, and Horn's (1965) parallel analysis also suggested the retention of just two factors for both stimulus sets.

We then correlated these Shape-profile and Shape-complexity factors with neural similarity for all ROIs (Fig. 4C). This produced a clear pattern of results, with the Shape-profile factor dominating in early visual cortex but losing prominence in later extrastriate regions, in which the Shape-complexity factor takes over. This pattern both highlights and clarifies the general trends noted previously in the individual predictor correlations depicted in Figure 4A. The resultant correlations were evaluated with ROI × factor ANOVAs for each stimulus set. For both stimulus sets, significant main effects of ROI (standard, $F_{(7,77)} = 13.20$, $p = 4.7 \times 10^{-11}$; scrambled, $F_{(7,77)} = 20.93$, $p = 1.7 \times 10^{-15}$) and factor (standard, $F_{(1,11)} = 15.43$, $p = 0.002$; scrambled, $F_{(1,11)} = 40.21$, $p = 5.5 \times 10^{-5}$) emerged. However, these are qualified by highly significant interactions (standard, $F_{(2.80,30.80)} = 35.24$, $p = 7.7 \times 10^{-10}$; scrambled, $F_{(7,77)} = 36.83$, $p = 4.4 \times 10^{-22}$), making interpretation of main effects difficult. As such, we ran paired-samples t tests within each ROI to compare factors.

For both standard and scrambled stimuli, the Shape-profile factor correlations were significantly greater than those for the Shape-complexity factor in V1, V2, and V3 (standard, $p = 1.0 \times 10^{-6}$, $p = 1.5 \times 10^{-4}$ and $p = 1.1 \times 10^{-5}$, respectively; scrambled, $p = 1.0 \times 10^{-7}$, 3.0×10^{-5}, and 2.0×10^{-4}, respectively). V4 showed a Shape-profile preference for the standard ($p = 0.048$) but not scrambled ($p = 0.361$) stimuli.

Conversely, we found significant Shape-complexity preferences in LO-2 and LO (standard, $p = 0.011$, $p = 6.2 \times 10^{-5}$, respectively; scrambled, $p = 0.031$, $p = 0.004$, respectively), plus pFs for the standard ($p = 0.006$) but not scrambled ($p = 0.293$) stimuli. In LO-1, we found no significant differences between the Shape-profile and Shape-complexity factors (standard, $p = 0.591$; scrambled, $p = 0.073$); this could either suggest that we failed to capture variance in LO-1 or that both factors play some role in describing its activity. To explore whether the correlations were greater than zero, we ran one-sample t tests that found significant results for both the Shape-profile (standard, $p = 0.014$; scrambled, $p = 5.1 \times 10^{-4}$) and Shape-complexity (standard, $p = 0.009$; scrambled, $p = 0.041$) factors. This implies that the hypothesized retinotopic-to-functional transition could be occurring here.

Result reliability

To verify the robustness of our findings, we assessed alternative explanations and potential issues that might confound interpretation.

First, we aimed to address whether the shifting neural representation from V1-V3 to LO-1, LO-2, and LO is specific to these ROIs or is it a more general property of neural tunings when moving anteriorly through visual cortex (perhaps because of increasing receptive field sizes for example)? To test this, we analyzed two additional retinotopically defined areas, V3A/B (anterior and dorsal to V3d) and combined regions VO-1 and VO-2 (VO-1/2; Brewer et al., 2005) that extend anteriorly from V4 (the VO-1/VO-2 boundary was not always clear across participants so they were collapsed to a single ROI). These ROIs were again constrained with the "all stimuli over baseline" contrasts (as per the main study). For the standard stimuli, the average correlation between the Shape-profile and Shape-complexity factors with neural similarity in V3A/B were $r = 0.30$ and $r = 0.07$, respectively, and they differed significantly ($t_{(11)} = 3.65$, $p = 0.004$). A similar profile of results was found for VO-1/2 (Shape-profile, $r = 0.20$; Shape-complexity, $r = 0.00$; $t_{(11)} = 2.86$, $p = 0.015$). For the scrambled stimuli, the corresponding V3A/B correlations were $r = 0.16$ and $r = 0.14$, which in this case were not significantly different ($t_{(11)} = 0.19$, $p = 0.854$). In VO-1/2, there was no evidence that either factor was represented (Shape-profile, $r = 0.05$; Shape-complexity, $r = 0.02$; $t_{(11)} = 0.52$, $p = 0.612$). Whereas the correlations with scrambled stimuli were generally weaker, the fact that with standard stimuli the Shape-profile correlation was significantly greater than the Shape-complexity correlation in both V3A/B and VO-1/2 demonstrates that more anterior ROIs do not necessarily transition to the more abstract/curvature-tuned representation that emerged around LO-1 and LO-2.

Second, a potential issue is the use of spatial smoothing. It is possible that smoothing blurred the boundaries between ROIs, and perhaps results such as the Shape-complexity preference of LO-2 could be explained by spread from neighboring LO. To test this, we re-ran all analysis without spatial smoothing and correlated the Shape-profile and Shape-complexity factors with the resultant neural similarity measures across all ROIs. In general, very small differences were observed, with average losses of 0.77% (standard stimuli) and 1.93% (scrambled stimuli) variance explained when spatial smoothing was not used. As a lenient test for differences, we then ran two-tailed paired-samples t tests to compare correlations in the original analysis with those produced without spatial smoothing. The results reported in Table 2 suggest that generally spatial smoothing was beneficial, because correlations in larger ROIs that you would expect to be less susceptible to spatial smoothing (e.g., V1 or LO) were significantly reduced without it. In contrast, for a smaller ROI (LO-1), we observed a significant increase in the Shape-complexity correlation, implying that a lack of spatial smoothing was actually beneficial. Importantly, the interpretation of the data analysis remains the same whether or not spatial smoothing was used.

Finally, the main correlation analysis examined how our factors predict neural similarity on an individual basis, but this contains individual variability as a source of noise. To provide a cleaner picture of our results, we therefore collapsed neural similarity across participants and compared the resultant correlation matrices with our factors. No differences emerged in the pattern of correlations, although as expected on the basis of reducing noise, our factors now captured more variance (Fig. 5A). Collapsing neural similarity across participants also allows us to visualize our results. Figure 5B depicts for both stimulus sets the averaged neural similarity matrix between all pairwise combinations of the shape stimuli in four key ROIs, as well as the pairwise stimulus similarity described by the two factors (Shape-profile, Shape-

Table 2. Comparing shape factor/neural similarity correlations with and without spatial smoothing

	V1	V2	V3	V4	LO-1	LO-2	LO	pFs
Standard stimuli								
Shape-profile	0.69/0.71	0.60/0.63	0.65/0.60	0.27/0.24	0.15/0.10	**0.06/0.02***	**0.02/−0.01***	−0.08/−0.05
Shape-complexity	0.01/−0.01	0.06/0.06	0.01/0.03	0.09/0.06	0.20/0.23	0.37/0.32	0.39/0.33	**0.13/0.05***
Scrambled stimuli								
Shape-profile	**0.79/0.76***	**0.73/0.68***	**0.50/0.37****	0.08/0.06	0.23/0.19	0.03/0.05	**0.06/−0.01****	0.00/−0.03
Shape-complexity	**−0.02/−0.05***	−0.02/−0.02	0.01/0.04	0.16/0.13	**0.10/0.13****	0.24/0.20	**0.31/0.21***	−0.07/−0.06

Exploring the influence of spatial smoothing on correlational data. Spatial smoothing (Gaussian kernel with 4 mm FWHM) was used when analyzing our main functional scans, potentially blurring representations across smaller ROIs. To test whether this was an issue, we compared our main results (see Fig. 4C) with and without spatial smoothing applied (significance from uncorrected 2-tailed paired-samples t test). Generally, a lack of spatial smoothing appears to reduce correlations slightly, but the drop in variance explained is minimal because such spatial smoothing does not appear to be a confound in our analysis. Format: with spatial smoothing/without spatial smoothing. *p < 0.050, **p < 0.010, ***p < 0.001.
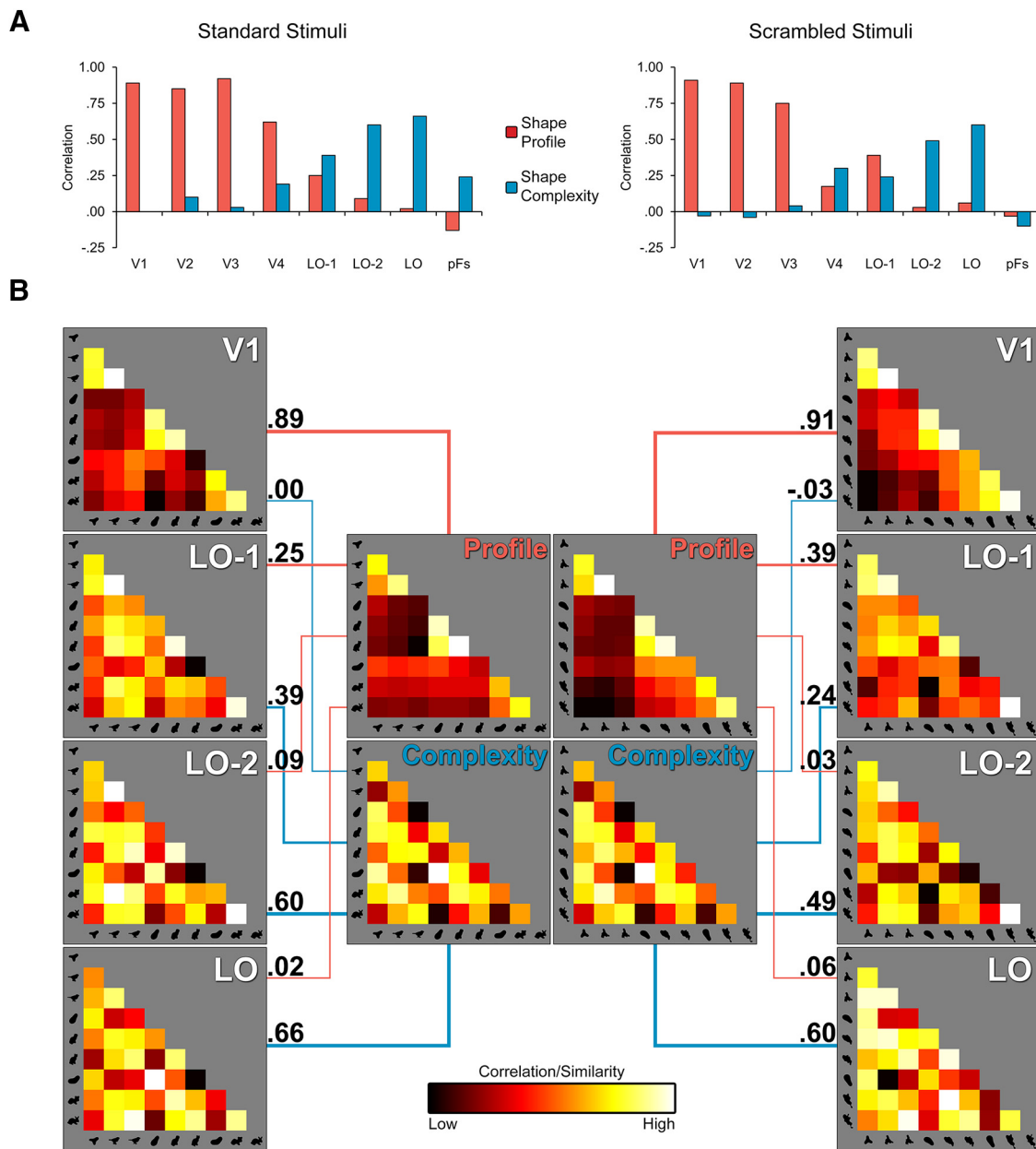

Figure 5. Group mean neural similarity matrices and their correlation with Shape-profile and Shape-complexity factors. **A**, Correlations between standard (left) and scrambled (right) group mean similarity matrices and the Shape-profile and Shape-complexity factors. **B**, Neural similarity matrices for the standard (left) and scrambled (right) stimuli in V1, LO-1, LO-2, and LO (collapsed across participant), plus factor similarity matrices centrally. Here "brighter" colors represent greater similarity; note that, to enhance visibility, the color map has been scaled independently and linearly in each image based on the minimum and maximum correlations, maximizing the range of colors used in each image. As with our main correlation analysis, we see that V1 corresponds closely to the Shape-profile factor, whereas LO-2 and LO show greater similarity with the Shape-complexity factor. LO-1 is linked to both factors, with the exact balance changing slightly between stimulus sets.
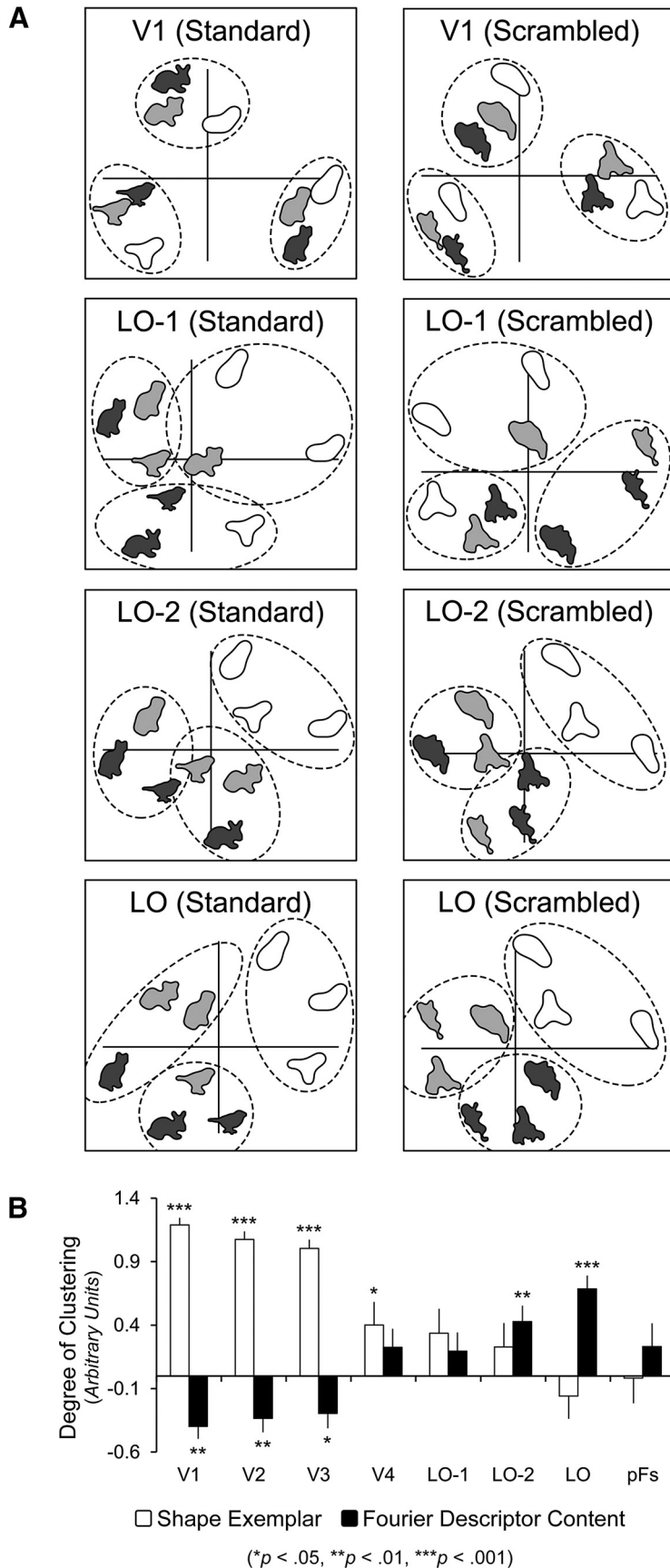
complexity). This approach provides some insight into the organizational principles underlying our Shape-profile and Shape-complexity factors. For example, the similarity between the "mid detail bird" and "mid detail rabbit" (column 2, row 2 from bottom left in each similarity matrix) is very low in the Shape-profile factor because of the minimal spatial overlap, but they are highly similar in the Shape-complexity factor. This implies that our Shape-complexity factor captures some index of shape curvature or complexity not grounded in retinotopic coordinates.

MDS

Our final exploratory approach, MDS, used the neural similarity matrices from a given ROI to test for underlying structure in terms of how that ROI responds to shape stimuli (see MDS solutions for representative ROIs in Fig. 6A). Only two-dimensional solutions were extracted because this allows for easy visualization, allowing us to perceptually assess the resultant solutions and providing some idea of how each ROI may be representing the shape stimuli. Crucially, although our correlation analysis could only explore neural similarity with respect to our factors (based on stimulus properties), MDS takes no account of the relations between stimuli because it is based on neural data alone.

We first ran the "stimulus blind" clustering analysis on our representative ROIs to provide some insight into the underlying structure. Specially, for each ROI, we identified the optimal 3 × 3 grouping (i.e., three groups of three items) that minimized the average pairwise distance between the items within each group (see

←

each ROI (see Materials and Methods, Multidimensional scaling). Here we see strong shape exemplar clustering in V1, a general lack of clustering in LO-1, and evidence for FD-Content clustering in LO-2 and (to a greater extent) LO. Also of note, for LO-2, there is some evidence of an orthogonal separation between FD-Content and shape exemplar. The dispersion accounted for each ROI MDS solution is (standard stimuli/scrambled stimuli) as follows: V1, 0.93/0.95; LO-1, 0.82/0.86; LO-2, 0.86/0.86; LO, 0.89/0.86. ***B***, Graphical depiction of clustering based on the stimulus dimensions (shape exemplar; FD-Content; Fig. 2); here we plot the mean distance between all stimuli minus the distance between all pairs within a given stimulus dimension. A value of zero implies that clustering within a dimension is no greater than the average clustering across all pairwise combinations of items, whereas positive values imply some degree of clustering. Significance values are taken from one-sample *t* tests to assess whether the clustering differs significantly from zero, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. Error bars indicate SEM.

Figure 6. Solutions from MDS. ***A***, Extracted two-dimensional solutions for four representative brain regions in the standard (left) and scrambled (right) stimulus sets. For FD-Content, low, mid, and high detailed shapes have been shaded in white, mid-gray, and dark-gray, respectively. The dashed rings highlight the tightest set of three clusters (each containing 3 items) for

Materials and Methods). The clusters (Fig. 6A, circled items) indicated that, in V1, items with a similar shape profile were grouped together, whereas the clusters in later extrastriate areas (namely LO-2 and LO) were primarily grouped on the basis of shape/curvature complexity. No consistent clustering emerged in LO-1.

Next, we performed analysis on the MDS solutions across all ROIs to specifically determine whether one or both of our original stimulus dimensions (from Fig. 2) were present in the resultant solutions [i.e., do shapes cluster on stimulus exemplar (bird, cat, or rabbit) or FD content (low, mid, and high detail)]. Note that we are not exploring how our factors (Shape-profile, Shape-complexity) relate to the MDS solutions because MDS essentially just re-describes the neural similarity data (albeit with reduced dimensionality); the main correlation analysis described above has already explored the factor/neural similarity relationship in the "purest" sense. No significant differences in clustering patterns emerged between the standard and scrambled stimulus sets (paired-samples t tests, all p values >0.173) and so for simplicity we collapsed across stimuli. The clustering analysis in Figure 6B again reveals a shift in the nature of shape representations between early and late visual cortices. Shape exemplar is most influential in early areas such as V1, whereas FD-Content (or level of shape detail) predominates in later extrastriate regions LO-2 and LO. Quantifying this, one-sample t tests indicated that the shape exemplar dimension showed significant clustering in V1 ($p = 9.1 \times 10^{-14}$), V2 ($p = 4.5 \times 10^{-12}$), V3 ($p = 5.2 \times 10^{-11}$), and V4 ($p = 0.041$). In fact, for V1, V2, and V3 the FD-Content dimension showed significant negative clustering (i.e., dispersion; V1, $p = 0.001$; V2, $p = 0.008$; V3, $p = 0.022$), likely attributable to the tight exemplar clusters driving down the average pairwise distance. Only LO-2 and LO showed significant clustering on the FD-Content dimension (LO-2, $p = 0.003$; LO, $p = 5.7 \times 10^{-6}$). No dimensions in LO-1 or pFs differed significantly from zero (all p values >0.099), implying a lack of clustering.

Generally, these results support our main findings. Early retinotopic regions are dominated by the spatial layout of the outline of a shape, whereas lateral occipital ROIs (namely LO-2/LO) are dominated by more abstract measures (likely the size and number of protrusions around each shape). In LO-1, neither dimension appears to dominate, so it is a viable candidate for the transition point between these two organizational principles. Interestingly, LO-2 appeared to broadly split the shape exemplar and FD-Content dimensions orthogonally (Fig. 6B), implying sensitivity to shape profile that was not captured in our factor analysis (although hints emerged in the raw predictors, see behavioral and GIST predictors; Fig. 4A). This would fit with the retinotopic organization of the area, but our clustering analysis did show that the FD-Content dimension nevertheless dominates (in line with the main analysis).

Discussion

A general pattern of results emerged across our studies using standard and scrambled stimuli, with activity in lower-level retinotopic ROIs (V1–V3) being better predicted by the spatial layout of a shape, whereas extrastriate lateral occipital regions LO-2 and LO were better predicted by more abstract shape features. A middle ground emerged in LO-1, which could represent a transitional point between these two organizational principles. The more abstract representation appears to be specific to the lateral occipital cortex, because more abstract tunings were not found in ventral ROIs V4 and VO-1/2 or the more dorsal ROI V3A/B.

First, we note that very similar results were identified across both stimulus sets, despite their considerable differences, demonstrating the robustness of the findings. It also reinforces the lack of semantic influences in both early visual areas and LO, in line with previous work (Kim et al., 2009). This places LO as perhaps an earlier shape-processing region, one that encodes the form of the shape regardless of its novelty, recognizability or familiarity.

The results from a standard GLM analysis identified a generally flat profile of responses to our shape stimuli, with the exception of pFs, which showed diminished activity. This highlights the importance of exploring representations using multivariate approaches, given that total signal change revealed very few differences along the visual hierarchy. We also noted consistently greater activation for high complexity shapes (with slight differences between areas), likely attributable to the increased amount of detail in their contours.

Turning to our main findings, both our correlation and MDS results implied that there were (at least) two general organizational principles in the visual cortex. First, we have low-level organization (Shape-profile), in which the precise layout of the shape drives activation (i.e., translation, scaling, or rotation would alter similarity measures under this definition). Unsurprisingly, this measure was a strong predictor for early retinotopic cortex. Second, we have a more abstract organization (Shape-complexity) that appears to be capturing curvature in the profile of the shape (translation, scaling, rotation would not alter abstract similarity measures). This measure dominated in LO.

Of key interest is how these two organizational principles relate to LO-1 and LO-2. LO-1 appeared to contain influences of both low-level and abstract representations, whereas LO-2 showed remarkable similarity to the results profile of LO. In effect, we seem to have a retinotopic ROI responding nonretinotopically. However, receptive field sizes increase progressively through the visual hierarchy (Dumoulin and Wandell, 2008; Amano et al., 2009), implying that LO-1 and (to a greater extent) LO-2 will be pooling information over larger areas compared with earlier regions. Larger receptive field sizes could allow for a more lenient response profile to visual stimuli, meaning that regions such as LO-2 can respond to more abstract features (such as curvature-complexity) instead of the strict Shape-profile-based activation of areas with smaller receptive fields. However, it is worth noting that our stimuli were presented in a relatively limited region of the visual field, perhaps biasing tunings in LO-2 away from a retinotopic response profile. It is likely that our low-level measures of shape overlap would play a greater role in predicting neural similarity if stimulus spatial position was an experimental manipulation, particularly given evidence of sensitivity to spatial position in LO (Grill-Spector et al., 1999; Sayres and Grill-Spector, 2008; Schwarzlose et al., 2008). Nevertheless, spatial jitter may have masked the more subtle shape representations that emerged in this study; therefore, controlling for stimulus position was likely beneficial.

Given that LO-2 in particular shows some overlap with LO (Sayres and Grill-Spector, 2008) is causally involved in shape-processing tasks (Silson et al., 2013) and both LO-2 and LO appear to share the same abstracted representations, there is strong, converging evidence that this region is a preliminary object-processing stage as Larsson and Heeger (2006) posited. However, although it is possible that this more abstract representation emerges solely within LO-2, its complete lack of Shape-profile tuning implies that there may be some previous transitional point at which more abstract tunings start to emerge. A good candidate

for the transition is LO-1, in which patterns of activity could be explained (albeit weakly) by both low-level and more abstract stimulus properties. As such, although the functions of LO-1 and LO-2 are clearly dissociable to some extent (Silson et al., 2013), a stream of activation passing from LO-1 to LO-2 and then LO (and beyond) seems plausible. This has a second implication, that LO could contain more explicit retinotopy than the general lower visual-field biases that have been identified previously (Sayres and Grill-Spector, 2008); indeed, retinotopic maps have been reported in this area (Kolster et al., 2010). Although we have clearly found that a more abstract dimension predominates in this area, the same is true for LO-2, which does contain retinotopic maps. As such, an abstract representation does not preclude underlying retinotopy.

The nature of the abstract representation emerging in LO-2 and LO is perhaps less intuitive than the concept of low-level physical similarity. Our abstract similarity measures were clearly proxies for some underlying construct, given that they reduced so cleanly to a single dimension, and this dimension appears to be capturing the contours, or more specifically the curvature of the shape. This finding is in accordance with macaque literature demonstrating the importance of curvature information for the ventral visual stream (Kayaert et al., 2011; Yue et al., 2014). It also corroborates Drucker and Aguirre (2009), who found that amplitude manipulations in composite radial frequency patterns (making protrusions more/less salient) were linked to LO neural similarity. However, we also found that the number of curves in the profile of a shape strongly predicted similarity in LO-2 and LO (see minima/maxima; Fig. 4A), implying that both amplitude and frequency are important. Future work could evaluate how the two variables underpin shape representations in LO. It would also be valuable to probe the respective roles of convexity and concavity. It is well established that convexities are generally highly salient visual features (Bertamini and Wagemans, 2013), and convexities elicit greater activation than concavities (Haushofer et al., 2008a). Moreover, LO can represent shapes when parts of the contours are deleted or obscured (Kourtzi and Kanwisher, 2001; Lerner et al., 2002; Stanley and Rubin, 2003; Hayworth and Biederman, 2006). Convexities and concavities must occur together, and so fully representing both could lead to redundancy. The broader evidence therefore points to an underlying representation based on shape convexity.

Finally, it is intriguing that the abstract shape features failed to reliably predict neural activity within pFs. This region was less active compared with other areas, so perhaps responses were not robust enough to establish clear neural similarity measures. Alternatively, individual differences in perceived shape similarity could underlie this discrepancy. Our behavioral measure was originally intended to capture such between-subject differences, but it appears as though participants were predominantly using spatial overlap to determine perceptual shape similarity, given that the behavioral measure collapsed into the same dimension as the low-level predictors. This meant that our metric for perceptual similarity did not correlate with LO or pFs activity in contrast to previous work (Haushofer et al., 2008b; Op de Beeck et al., 2008b). Haushofer et al. in particular noted that responses within pFs (but not LO) reflected a more "implicit" metric of shape similarity (how likely two different shapes were to be confused as identical), which perhaps is better suited to measuring subtle perceptual differences than explicitly asking for similarity ratings. Such perceptual differences may underlie the unexplained variance in pFs activity in our study.

To conclude, we identified two orthogonal organizational principles in the visual cortex: Shape-profile and Shape-complexity. The Shape-profile factor reflected the spatial layout of our shape stimuli, whereas the Shape-complexity factor instead reflected changes in curvature around the shapes perimeter. Accordingly, our Shape-profile factor was a strong predictor for early retinotopic cortex, whereas the Shape-complexity factor correlated well with shape-selective LO. Critically, LO-1 and LO-2 (two retinotopic regions lying between early visual cortex and later extrastriate regions) show intermediate representations. LO-1 has influences from both factors, whereas in LO-2, the Shape-complexity factor dominates. We argue that this represents a transitional point between two previously discrete approaches: retinotopy and functional selectivity.

References

Amano K, Wandell BA, Dumoulin SO (2009) Visual field maps, population receptive field sizes, and visual field coverage in the human MT+ complex. J Neurophysiol 102:2704–2718. CrossRef Medline

Arcaro MJ, McMains SA, Singer BD, Kastner S (2009) Retinotopic organization of human ventral visual cortex. J Neurosci 29:10638–10652. CrossRef Medline

Baseler HA, Gouws A, Haak KV, Racey C, Crossland MD, Tufail A, Rubin GS, Cornelissen FW, Morland AB (2011) Large-scale remapping of visual cortex is absent in adult humans with macular degeneration. Nat Neurosci 14:649–655. CrossRef Medline

Bertamini M, Wagemans J (2013) Processing convexity and concavity along a 2-D contour: figure-ground, structural shape, and attention. Psychon Bull Rev 20:191–207. CrossRef Medline

Brewer AA, Liu J, Wade AR, Wandell BA (2005) Visual field maps and stimulus selectivity in human ventral occipital cortex. Nat Neurosci 8:1102–1109. CrossRef Medline

Busing F, Commandeur JJ, Heiser WJ, Bandilla W, Faulbaum F (1997) PROXSCAL: a multidimensional scaling program for individual differences scaling with constraints. Softstat 97:67–74.

Carroll JD, Chang JJ (1970) Analysis of individual differences in multidimensional scaling via an n-way generalization of "Eckart-Young" decomposition. Psychometrika 35:283–319. CrossRef

De Winter J, Wagemans J (2004) Contour-based object identification and segmentation: stimuli, norms and data, and software tools. Behav Res Methods Instrum Comput 36:604–624. 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

Drucker DM, Aguirre GK (2009) Different spatial scales of shape similarity representation in lateral and ventral LOC. Cereb Cortex 19:2269–2280. CrossRef Medline

Dumoulin SO, Wandell BA (2008) Population receptive field estimates in human visual cortex. Neuroimage 39:647–660. CrossRef Medline

Engel SA, Rumelhart DE, Wandell BA, Lee AT, Glover GH, Chichilnisky EJ, Shadlen MN (1994) fMRI of human visual cortex. Nature 369:525. CrossRef Medline

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

Gouws AD, Alvarez I, Watson DM, Uesaki M, Rodgers J, Morland AB (2014) On the role of suppression in spatial attention: evidence from negative BOLD in human subcortical and cortical structures. J Neurosci 34:10347–10360. 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, Edelman S, Avidan G, Itzchak Y, Malach R (1999) Differential processing of objects under various viewing conditions in the human lateral occipital complex. Neuron 24:187–203. CrossRef Medline

Grill-Spector K, Kourtzi Z, Kanwisher N (2001) The lateral occipital complex and its role in object recognition. Vision Res 41:1409–1422. 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, Baker CI, Livingstone MS, Kanwisher N (2008a) Privileged coding of convex shapes in human object-selective cortex. J Neurophysiol 100:753–762. CrossRef Medline

Haushofer J, Livingstone MS, Kanwisher N (2008b) Multivariate patterns in object-selective cortex dissociate perceptual and physical shape similarity. PLoS Biol 6:e187. CrossRef Medline

Hayworth KJ, Biederman I (2006) Neural evidence for intermediate representations in object recognition. Vision Res 46:4024–4031. CrossRef Medline

Hemond CC, Kanwisher NG, Op de Beeck HP (2007) A preference for contralateral stimuli in human object- and face-selective cortex. Ferrari P, ed. PLoS One 2:e574. CrossRef Medline

Horn JL (1965) A rationale and test for the number of factors in factor analysis. Psychometrika 30:179–185. CrossRef Medline

Hupé JM, Bordier C, Dojat M (2012) A BOLD signature of eyeblinks in the visual cortex. Neuroimage 61:149–161. CrossRef Medline

Julian JB, Fedorenko E, Webster J, Kanwisher N (2012) An algorithmic method for functionally defining regions of interest in the ventral visual pathway. Neuroimage 60:2357–2364. CrossRef Medline

Kaiser HF (1960) The application of electronic computers to factor analysis. Educ Psychol Meas 20:141–151. CrossRef

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

Kayaert G, Wagemans J, Vogels R (2011) Encoding of complexity, shape, and curvature by macaque infero-temporal neurons. Front Syst Neurosci 5:51. CrossRef Medline

Kim JG, Biederman I, Lescroart MD, Hayworth KJ (2009) Adaptation to objects in the lateral occipital complex (LOC): shape or semantics? Vision Res 49:2297–2305. CrossRef Medline

Kolster H, Peeters R, Orban GA (2010) The retinotopic organization of the human middle temporal area MT/V5 and its cortical neighbors. J Neurosci 30:9801–9820. CrossRef

Kourtzi Z, Kanwisher N (2001) Representation of perceived object shape by the human lateral occipital complex. Science 293:1506–1509. CrossRef Medline

Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI (2009) Circular analysis in systems neuroscience: the dangers of double dipping. Nat Neurosci 12:535–540. CrossRef Medline

Larsson J, Heeger DJ (2006) Two retinotopic visual areas in human lateral occipital cortex. J Neurosci 26:13128–13142. CrossRef Medline

Lerner Y, Hendler T, Malach R (2002) Object-completion effects in the human lateral occipital complex. Cereb Cortex 12:163–177. 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. CrossRef Medline

Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy WA, Ledden PJ, Brady TJ, Rosen BR, Tootell RB (1995) Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc Natl Acad Sci U S A 92:8135–8139. CrossRef Medline

Malach R, Levy I, Hasson U (2002) The topography of high-order human object areas. Trends Cogn Sci 6:176–184. CrossRef Medline

Mumford J (2007) A guide to calculating percent change with featquery. Unpublished technical report. Available at http://mumford.bol.ucla.edu/perchange_guide.pdf.

Oliva A, Torralba A (2001) Modeling the shape of the scene: a holistic representation of the spatial envelope. Int J Comput Vis 42:145–175. CrossRef

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

Peirce JW (2015) Understanding mid-level representations in visual processing. J Vis 15:5. CrossRef Medline

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

Schwarzlose RF, Swisher JD, Dang S, Kanwisher N (2008) The distribution of category and location information across object-selective regions in human visual cortex. Proc Natl Acad Sci U S A 105:4447–4452. CrossRef Medline

Silson EH, McKeefry DJ, Rodgers J, Gouws AD, Hymers M, Morland AB (2013) Specialized and independent processing of orientation and shape in visual field maps LO1 and LO2. Nat Neurosci 16:267–269. CrossRef Medline

Snodgrass JG, Vanderwart M (1980) A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. J Exp Psychol Hum Learn 6:174–215. CrossRef Medline

Stanley DA, Rubin N (2003) fMRI activation in response to illusory contours and salient regions in the human lateral occipital complex. Neuron 37:323–331. CrossRef Medline

Wandell BA, Dumoulin SO, Brewer AA (2007) Visual field maps in human cortex. Neuron 56:366–383. CrossRef Medline

Wilkinson F, Wilson HR, Habak C (1998) Detection and recognition of radial frequency patterns. Vision Res 38:3555–3568. CrossRef Medline

Worsley KJ (2001) Statistical analysis of activation images. In: Functional MRI: an introduction to methods (Jezzard P, Matthews PM, Smith SM, eds), pp 251–270. Oxford: Oxford University Press.

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

Zahn CT, Roskies RZ (1972) Fourier descriptors for plane closed curves. IEEE Trans Comput 21:269–281.