Version Date: 11/12/21

The dorsal visual pathway represents object-centered spatial relations for object recognition

Vladislav Ayzenberg and Marlene Behrmann
Neuroscience Institute, Carnegie Mellon University

Corresponding Authors:
Vladislav Ayzenberg: vayzenbe@andrew.cmu.edu
Marlene Behrmann: behrmann@andrew.cmu.edu

Abstract

Although there is mounting evidence that input from the dorsal visual pathway is crucial for object processes in the ventral pathway, the specific functional contributions of dorsal cortex to these processes remains poorly understood. Here, we hypothesized that dorsal cortex computes the spatial relations among an object's parts – a processes crucial for forming global shape percepts – and transmits this information to the ventral pathway to support object categorization. Using multiple functional localizers, we discovered regions in the intraparietal sulcus (IPS) that were selectively involved in computing object-centered part relations. These regions exhibited task-dependent functional connectivity with ventral cortex, and were distinct from other dorsal regions, such as those representing allocentric relations, 3D shape, and tools. In a subsequent experiment, we found that the multivariate response of posterior IPS, defined on the basis of part-relations, could be used to decode object category at levels comparable to ventral object regions. Moreover, mediation and multivariate connectivity analyses further suggested that IPS may account for representations of part relations in the ventral pathway. Together, our results highlight specific contributions of the dorsal visual pathway to object recognition. We suggest that dorsal cortex is a crucial source of input to the ventral pathway and may support the ability to categorize objects on the basis of global shape.

Keywords: dorsal stream, ventral stream, two visual streams, object recognition, shape perception, visual cortex

Introduction

A central organizing principle of the brain is that the visual system is segregated into a ventral visual pathway for recognizing objects and a dorsal visual pathway for locating and interacting with objects (Mishkin et al., 1983; Ungerleider & Haxby, 1994). However, research increasingly shows that the dorsal pathway computes some of the same object properties as the ventral pathway (Farivar, 2009; Freud et al., 2020; Freud et al., 2016), and may even play a functional role in object recognition (Freud et al., 2020; Holler et al., 2019). Despite these findings, the dorsal pathway is rarely included in conceptual or computational models of visual recognition (Gauthier & Tarr, 2016; Zhuang et al., 2021). Indeed, artificial neural network models (ANNs) trained for object recognition are almost exclusively modelled on ventral cortex processes (Blauch et al., 2021; Kubilius et al., 2019). One potential reason for this exclusion, is that the specific functional contributions of the dorsal pathway to object recognition are poorly understood.

The primary function of the dorsal pathway has long been considered to be the computation of visuospatial information in the service of coordinating actions (Goodale & Milner, 1992; Mishkin et al., 1983). However, dorsal cortex, particularly the posterior parietal cortex (PPC), also computes object properties relevant for recognition. For instance, many studies find robust sensitivity to shape information in the PPC (Bracci & Op de Beeck, 2016; Freud et al., 2017; Georgieva et al., 2008), akin to ventral object regions such as the lateral occipital complex (LOC; Grill-Spector et al., 2001; Kourtzi & Kanwisher, 2001). As in LOC, dorsal shape representations are seemingly robust to changes in size and orientation, as well as format (i.e., 3D vs. 2D; Konen & Kastner, 2008; Vaziri-Pashkam & Xu, 2019). Object representations in the dorsal pathway also appear to code relatively abstract representations, such that the multivariate responses in PPC corresponds to perceived semantic similarity among objects, even when controlling for low-level visual properties (Bracci & Op de Beeck, 2016; Jeong & Xu, 2016).

Although these studies highlight the similarities between dorsal and ventral pathways, object representations in dorsal cortex are not simply redundant with those in the ventral cortex (Bracci & Op de Beeck, 2016; Freud et al., 2015; Vaziri-Pashkam & Xu, 2019). What, then, are the unique contributions of the dorsal stream to object recognition? One possibility, consistent with its role in visuospatial processing (Kravitz et al., 2011; Mishkin et al., 1983), is that dorsal cortex computes the spatial relations among an object's component parts – that is, the object's topological structure.

Many studies have demonstrated that a description of part relations is crucial for forming invariant 'global shape' representations (Biederman, 1987; Hummel, 2000), which may be key for recognizing objects across variations in viewpoint or across category exemplars (Ayzenberg & Lourenco, 2019; Hummel & Stankiewicz, 1996). Indeed, an inability to represent the part relations results in marked deficits in object recognition (Behrmann et al., 2006). Such a representation may be particularly important for basic-level object categorization because members of a category typically have similar spatial structures, but vary in regards to their component parts (Ayzenberg & Lourenco, 2019; Barenholtz & Tarr, 2006; Rosch et al., 1976).

Surprisingly, few studies have investigated whether the dorsal pathway represents object-centered part relations, with most, historically, focusing on allocentric spatial coding (Haxby et al., 1991), and even fewer have examined the relation between such coding in the dorsal pathway and object recognition processes in the ventral pathway (c.f. Zachariou et al., 2017). Thus, in the current study, we tested whether the dorsal visual pathway represents the relations among component parts, and

the extent to which coding in this region is independent of allocentric relations and other object properties represented by the dorsal pathway, such as 3D shape and tools. Furthermore, we investigated the contributions of such a region to object recognition processes in the ventral pathway.

Materials and Methods

Participants

We recruited 12 participants (3 female; M_{age} = 27.50, SD = 3.61) for Experiment 1, in which functional regions of interest (ROIs) were identified, and 12 participants (6 female; M_{age} = 26.83, SD = 3.7) for Experiment 2, in which the ROIs' contributions to object recognition were explored. Where possible, the same participants completed both Experiment 1 and 2, so that their pre-defined ROIs could be used for analysis. In total, eight participants from Experiment 1 also participated in Experiment 2. The four new participants in Experiment 2 were scanned in a second session (following the procedure of Experiment 1) in order to define their individual functional ROIs. All participants were right-handed and had normal or corrected-to-normal visual acuity. Participants were recruited from the Carnegie Mellon University community, gave informed consent according to a protocol approved by the Institutional Review Board (IRB), and received payment for their participation.

MRI scan parameters

Scanning was done on a 3T Siemens Prisma scanner at the CMU-Pitt Brain Imaging Data Generation & Education (BRIDGE) Center. Whole-brain functional images were acquired using a 64-channel head matrix coil and a gradient echo single-shot echoplanar imaging sequence. The acquisition protocol for each functional run consisted of 48 slices, repetition time = 1 s; echo time = 30 ms; flip angle = 64°; voxel size = 3 × 3 × 3 mm. Whole-brain, high-resolution T1-weighted anatomical images (repetition time = 2300 ms; echo time = 2.03 ms; voxel size = 1 × 1 ×1 mm) were also acquired for each participant for registration of the functional images.

Data Analysis

All images were skull-stripped (Smith, 2002) and registered to the Montreal Neurological Institute (MNI) 2mm standard template. Prior to statistical analyses, images were motion corrected, de-trended, and intensity normalized. To facilitate functional-connectivity analyses, 18 additional motion regressors generated by FSL were also included. All data were fit with a general linear model consisting of covariates that were convolved with a double-gamma function to approximate the hemodynamic response function. Data used to define regions of interest (ROIs) was spatially smoothed using a 6 mm Gaussian kernel. All other data were unsmoothed. All data were analyzed using the peak 100 voxels within a region (as defined by the functional localizer) or using a 6mm sphere (~120 voxels) centered on the peak voxel. Qualitatively similar results were found for all analyses when ROI sizes were varied parametrically from 100 to 400 voxels (the size of the smallest ROI). Analyses were conducted using FSL (Smith et al., 2004), and the nilearn, nibabel, and Brainiak packages for in Python (Abraham et al., 2014; Kumar et al., 2020).

Experimental design

Experiment 1: Localization of object-centered part relations

Participants completed four localizer scans to measure voxels activated by object-centered part relations, allocentric relations, depth information, and tool images. We first used a ROI approach to define regions in parietal cortex that represent part relations. Then, we used independent data to test the selectivity of these ROIs to part relations or to other visual properties represented by the dorsal stream, namely allocentric relations (Haxby et al., 1991), depth information (Georgieva et al., 2008), and tools (Mahon et al., 2007). Furthermore, we conducted conjunction analyses to examine the degree of overlap between dorsal ROIs sensitive to part relations and the other dorsal properties. Finally, we conducted task-dependent functional connectivity analyses to examine the degree to which dorsal ROIs sensitive to part relations are correlated with ventral regions.

For each localizer, we defined posterior and anterior parietal ROIs by overlaying posterior intraparietal sulcus (pIPS) and anterior IPS (aIPS) binary masks and selecting voxels within those masks that survived a whole-brain cluster-corrected threshold ($p < .001$). Broad pIPS and aIPS masks were created by combining IPS0 with IPS1 and IPS1with IPS2 probabilistic masks, respectively, from the Wang et al. (2014) atlas. For comparison of the activation profiles from dorsal regions, an object-selective ROI in the ventral stream was defined similarly within the lateral occipital complex (LOC) probabilistic parcel (Julian et al., 2012).

Object-centered part relations localizer. Participants completed six runs (320 s each) of an object-centered part relations localizer consisting of blocks of object images where either the spatial arrangement of component parts varied from image to image (part-relations condition), while the parts themselves stayed the same; or the local features of the component parts varied from image to image (feature condition), while the spatial arrangement of the parts stayed the same (Figure 1A). Objects could have one of 10 possible spatial arrangements, and one of 10 possible part features. Spatial arrangements were selected to be qualitatively different from one another as outlined by the recognition-by-components (RBC) model (e.g., end-to-end; end-to-middle; Biederman, 1987). The component parts were comprised of qualitatively different features as outlined by the RBC model (e.g., sphere, cube). All stimuli subtended ~6° visual angle on screen.

Each block of the part relations localizer contained 20 images, displaying each spatial arrangement or part feature twice per block depending on the condition. Each image was presented for 800 ms with a 200 ms interstimulus interval (ISI) for a total of 20 s per block. To minimize visual adaptation, the location of object images on the screen varied by ~2° every trial. The image order within the block was randomized. Participants also viewed blocks of a fixation cross (20 s). Participants viewed 5 repetitions of each block per run, with blocks presented in a pseudorandom order under the constraint that all three block types (relations, feature, fixation) were presented once before repetition. To maintain attention, participants performed an orthogonal one-back task, in which they responded via key press when detecting the repetition of an image on consecutive presentations.

Object-centered part relations ROIs in pIPS and aIPS were defined in each individual using 4 out of the 6 runs as those voxels that responded more to the part-relations than the feature condition. Selectivity was measured for each voxel in an ROI by extracting standardized parameter estimates for each condition (relative to fixation) in left out runs (2 out of 6).

Allocentric relations localizer. Participants completed two runs (368 s each) of an allocentric relations localizer wherein some blocks they judged whether displayed objects had the same allocentric relations, in this case the same distances between objects (distance condition), or had

the same brightness (brightness condition; Zachariou et al., 2017). A nearly identical display was shown in both conditions, consisting of two diagonally arranged displays, each containing a line and circle (Figure 1B). In the distance condition, the allocentric relations (i.e., distances) between the line and circle, either matched across the two displays or differed. In the brightness condition, the brightness of the circles across the two displays either matched or differed. On each trial, participants were required to indicate whether the two displays were the same or different (according to distance or brightness). Each display subtended ~4° visual angle on screen. Prior to the start of the scan, participants' individual sensitivity to distance and brightness (blocked) was measured using an adaptive task where the distances and brightness of the stimuli was titrated until accuracy on each of the tasks was approximately 75%.

Each block contained 10 distance or brightness trials, in which five trials had matching displays and five trials had different displays. Each trial was presented for 1700 ms with a 300 ms interstimulus interval (ISI) for a total of 20 s per block. The trial order within the block was randomized. Participants also viewed blocks of fixation (20s). Participants viewed 6 repetitions of each block per run, with blocks presented in a pseudorandom order under the constraint that all three block types (distance, brightness, fixation) were presented once before repetition.

Allocentric relation ROIs were defined in each individual as those voxels that responded more to the distance than the brightness condition. Selectivity was measured for each voxel in an ROI by extracting standardized parameter estimates for each condition (relative to fixation).

Depth localizer. Participants completed two runs (308 s each) of a depth localizer wherein they viewed blocks of object images that contained 3D shapes as defined from depth shading cues (3D condition), or 2D shapes with comparable low-level properties (2D condition; Figure 1C). Each condition was comprised of ten 3D or 2D object images from (Georgieva et al., 2008). All stimuli were ~6° visual angle on screen. Each block contained 20 images, displaying each possible 3D or 2D image twice per block. Each image was presented for 700 ms with a 100 ms interstimulus interval (ISI) for a total of 16 s per block. The image order within the block was randomized. Participants also viewed blocks of fixation (16 s). Participants viewed 6 repetitions of each block per run, with blocks presented in a pseudorandom order under the constraint that all three block types (3D, 2D, fixation) were presented once before repetition. To maintain attention, participants performed an orthogonal one-back task, responding to the repetition of an image on consecutive presentations.

Depth ROIs were defined in each individual as those voxels that responded more to the 3D than the 2D condition. Selectivity was measured for each voxel in an ROI by extracting standardized parameter estimates for each condition (relative to fixation) in left out runs.

Tool and object localizer. Participants completed two runs (340 s) of a tool localizer wherein they viewed blocks of object images that contained tools (tool condition), manipulable non-tool objects (non-tool condition), or box-scrambled object images (scrambled conditions; Figure 1D). Each condition was comprised of ten tools, non-tools, or scrambled object images from (Chen et al., 2018; Chen et al., 2016). Each block contained 20 images, displaying each possible tool, non-tool, or scrambled image twice per block. All stimuli subtended ~6° visual angle on screen. Each image was presented for 700 ms with a 100 ms interstimulus interval (ISI) for a total of 16 s per block. The image order within the block was randomized. Participants also viewed blocks of fixation (16 s). Participants viewed 5 repetitions of each block per run, with blocks presented in a pseudorandom order under the constraint that all four block types (tool, non-tool, scrambled, fixation) were

presented once before repetition. To maintain attention, participants performed an orthogonal one-back task, responding to the repetition of an image on consecutive presentations.

Tool ROIs were defined in each individual as those voxels that responded more to the tool than the non-tool condition. Object ROIs in LOC were defined as those voxels that responded more to objects (tool + non-tool) than scrambled. Selectivity was measured for each voxel in an ROI by extracting standardized parameter estimates for each condition (relative to fixation).

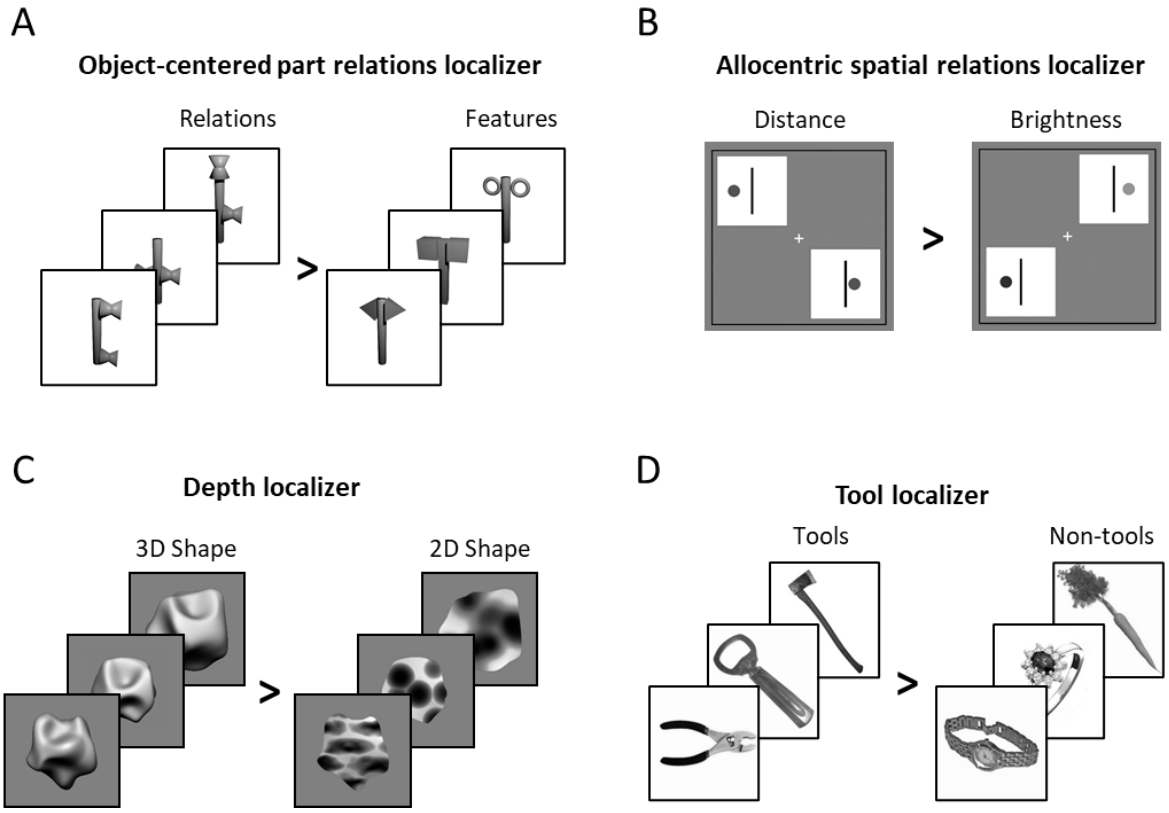


Figure 1. Example stimuli from the (A) object-centered part relations, (B) allocentric relations (C) depth, (D) and tool localizers used in Experiment 1.

Task-dependent functional connectivity. We conducted psychophysiological interaction (PPI; Friston et al., 1997) analyses to examine whether there is task-dependent functional connectivity between dorsal regions involved in computing part relations, and ventral regions involved in object recognition (Friston et al., 1997). A contrastive psychological task covariate was created from the part relations localizer by assigning timepoints corresponding to part-relations blocks a value of 1 and assigning timepoints corresponding to feature blocks a value of -1, then convolving the covariate with a standard HRF. Physiological covariates were generated from each participant's cleaned residual timeseries by extracting the timeseries from a 6 mm sphere centered on the peak voxel in dorsal ROIs that respond more to the relations than feature condition in the part relations localizer. Finally, a psychophysiological interaction covariate was created for each participant by multiplying the psychological and physiological covariates.

For each participant, 4 runs (randomly selected) of the object-centered spatial localizer were used to identify the peak voxel that responded more to the spatial than feature condition in pIPS and

aIPS parcels. The cleaned residual timeseries from the left-out two runs were then normalized, concatenated, and then further regressed on the psychological and physiological covariates generated for those runs. A whole-brain functional connectivity map was generated by correlating the residual timeseries of every voxel with the interaction covariate, and applying a fisher transform on the resulting map.

Data were analyzed in a cross-validated manner, such that every possible permutation of localizer (4 runs) and left-out runs (2 runs) was used to define the seed region separately, and then analyze connectivity. An average map was created by computing the mean across all permutations and a final group map was created by computing the mean across subjects. Significant voxels were determined by standardizing the group map and applying FDR-correction ($p < 0.05$). Together, this procedure ensures that any correlation between regions is driven by the task-dependent neural interaction, and not by the baseline correlation between regions or shared task activation.

Experiment 2: Basic-level object categorization in parietal ROIs

We tested whether the multivariate pattern in parietal ROIs that represent object-centered part relations can support basic-level object categorization. We further used representational similarity analyses (RSA), to examine the visual contributions of these ROIs to object recognition. Finally, we used multivariate functional connectivity analyses to examine the degree to which the part relation ROIs in the dorsal pathway interact with the ventral pathway.

To this end, participants completed 8 runs (330 s each) where they viewed images of common objects. The object set was comprised of five categories (boat, camera, car, guitar, lamp) each with five exemplars. Objects were selected from the ShapeNet 3D model dataset (Chang et al., 2015) and rendered to have the same orientation, texture, and color. The original texture and color information was removed to ensure that similarity among objects was on the basis of shape similarity, rather than other features. All stimuli subtended ~6° visual angle on screen (see Figure 2).

Objects were presented in an event-related design with the trial order and ISI optimized to maximize efficiency using Optseq2 (https://surfer.nmr.mgh.harvard.edu/optseq/). Each stimulus was presented for 1 s, with a jittered ISI between 1 and 8 seconds. Participants viewed 4 repetitions of each object per run. For each participant, parameter estimates for each object (relative to fixation) were extracted for each voxel. Responses to the stimuli in each voxel were then normalized by subtracting the mean response across all stimuli.

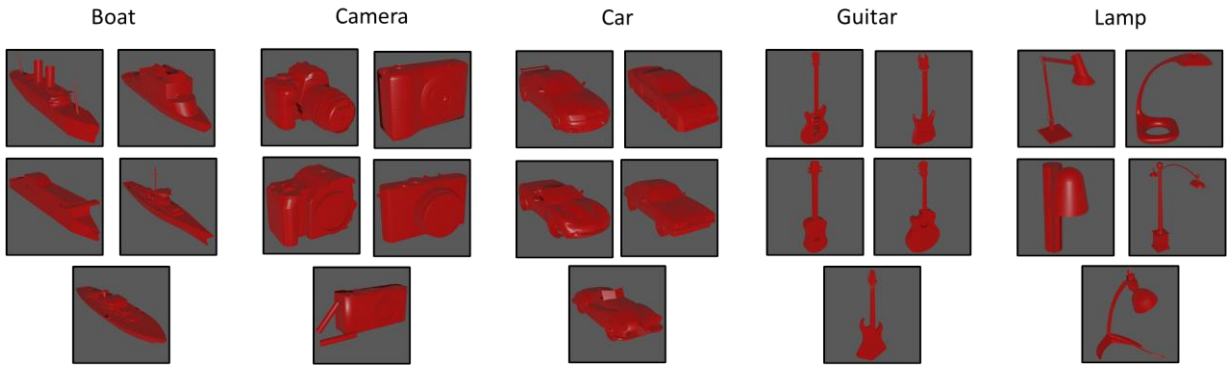


Figure 2. Object stimuli presented in Experiment 2. Participants viewed five exemplars from five categories in an event-related design.

Representational similarity analyses. A 25 × 25 symmetric neural representational dissimilarity matrix (RDM) was created for each ROI and participant by correlating (1-Pearson correlation) the voxel-wise responses for each stimulus with every other stimulus in a pairwise fashion. Neural RDMs were then Fisher transformed and averaged across participants separately for each ROI. Only the upper triangle of the resulting matrix (excluding the diagonal) was used in subsequent analyses.

Neural RDMs were compared to RDMs created from a model that approximates the spatial relations among component parts, namely a model based on the medial axis shape skeleton. Shape skeletons provide a quantitative description of the spatial arrangement of component parts via internal symmetry axes (Blum, 1973), and are tolerant to variations in the parts themselves (Ayzenberg et al., 2019; Feldman & Singh, 2006). For our skeletal model, we used a flux-based medial axis algorithm (Dimitrov et al., 2003; Rezanejad & Siddiqi, 2013) which computes a 'pruned' skeletal structure tolerant to local variations (Feldman & Singh, 2006). Skeletal similarity between objects was computed as the mean Euclidean distance between each point on one object's skeleton structure with the closest point on a second object's skeleton structure.

We also compared neural RDMs for models of low- and high-level vision, namely the Gabor-jet model, a model of image-similarity that approximates the response profile of early visual regions (Margalit et al., 2016), and the penultimate layer of CorNet-S, a recurrent artificial neural network designed to approximate the response profile of the ventral visual pathway in monkeys (Kubilius et al., 2019). Object similarity for both Gabor-jet and CorNet-S were computed as the mean Euclidean distance between feature vectors for each object image (see Figure 8A).

Multivariate connectivity analyses. We conducted multivariate pattern dependence (MVPD) analyses (Anzellotti et al., 2017) to examine whether dorsal ROIs involved in computing part relations interact with ventral object regions during object viewing. MVPD tests the degree to which the multivariate activation timeseries of a seed region accounts for the variance of the multivariate activation timeseries of a target region.

For each participant, data were split into a training (6 runs) and test (2 runs) set. A multivariate timeseries was generated from each participant's cleaned residual timeseries training data by extracting the timeseries of each voxel from a 6 mm sphere centered on the peak voxel in dorsal ROIs that responds more to the part-relations than feature blocks in the object-centered relations

localizer. The dimensionality of the voxel timeseries was then reduced by applying principal components analysis (PCA) and selecting the components that explain 90% of the variance. The same procedure was then repeated for a target region using a searchlight with 6 mm sphere. Next, using the training data, a linear regression was fit separately on each component of the target region using the components from the seed region as predictors. This procedure results in a series of beta weights describing the linear mapping between the principal components of the seed region to each individual principal component of the target region. For computational efficiency, the searchlight was conducted within an extended visual cortex mask created using an atlas from Wang et al. (2014) comprised of occipital, dorsal, and ventral visual cortices.

The beta weights from the training data are then used to generate a predicted multivariate timeseries for left-out runs of the target region, which is then correlated (Pearson) with the actual observed timeseries of the target region. A final fit value is computed as the weighted mean of correlations across target region principal components, with the weighting of each correlation determined by the proportion of variance explained by each target component. A single map for each participant is created by averaging the weighted correlations following 5-fold cross-validation, and then Fisher transforming the correlations. A final group map is created by computing mean across participants. Significant voxels were determined by standardizing the group map and applying FDR-correction ($p < 0.05$).

Results

Experiment 1: Selectivity for object-centered relations in the dorsal pathway

ROI definition. See Table 1 for a summary of significant group-level clusters from every localizer. The part relations localizer (4 runs) identified significant clusters in pIPS and aIPS in the right hemisphere (rpIPS, raIPS) of every participant and in 10 out of 12 participants in the left hemisphere (lpIPS, laIPS; see Figure 3A). Likewise, a group averaged map created using 2 runs (left out to measure selectivity) from every participant also revealed significant clusters in pIPS and aIPS, though these were found exclusively in the right hemisphere (see Figure 3B).

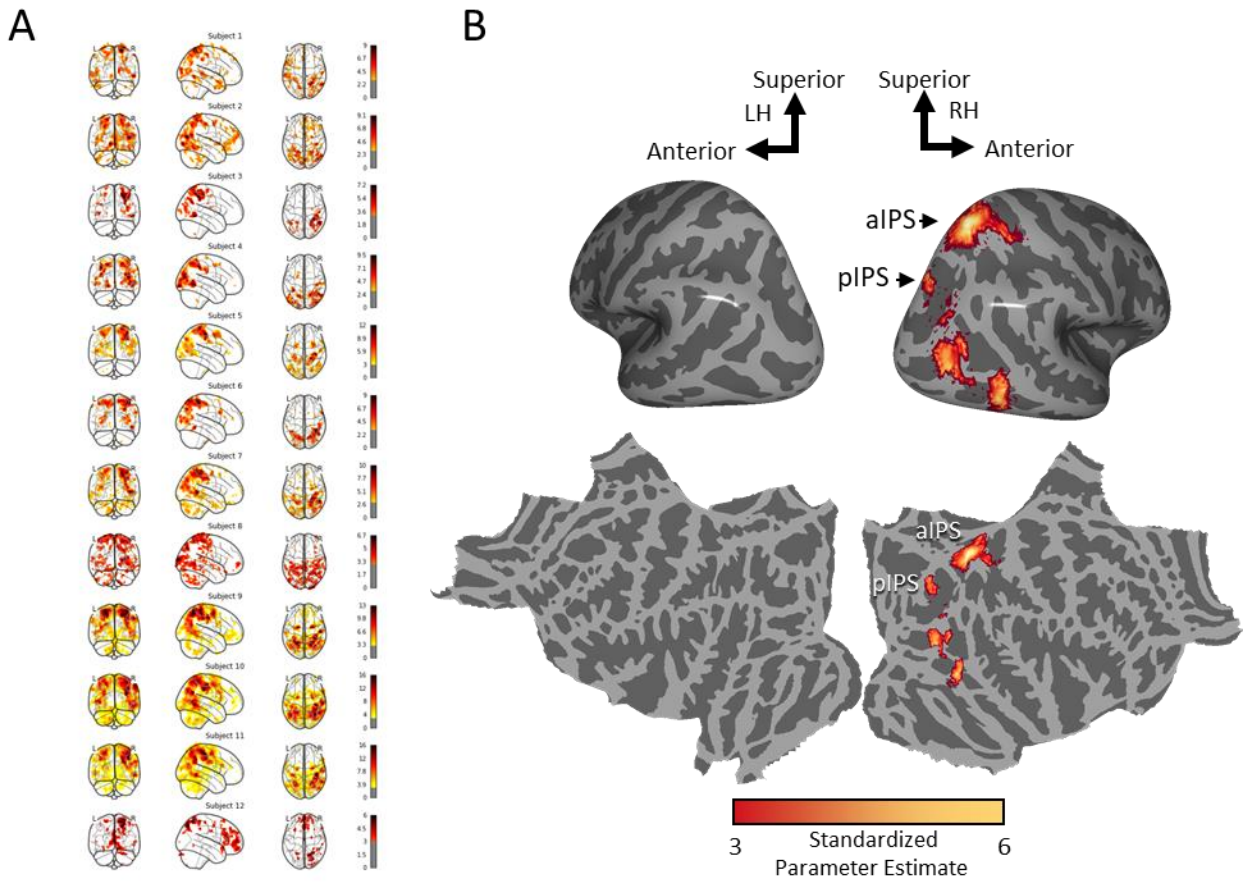


Figure 3. Significant activation to part relations (versus features) condition from the object-centered part relations localizer displayed (A) for each individual participant and in (B) a group average map inflated (above) and flattened (below). Values reflect the standardized parameter estimate.

Table 1. Significant group level clusters for the object-centered part relations, allocentric spatial relations, and tool localizer. MNI Coordinates correspond to the peak voxel within each cluster. The depth localizer is not listed because there were no significant clusters at the group level.

Localizer		Region	MNI Coordinate		
			x	y	z
Object part relations					
	1	R Posterior Intraparietal Sulcus (IPS0)	26	-76	44
	2	R Ventral Intraparietal Complex (VIP)	22	-58	64
	3	R Middle Temporal Area (MT)	42	-78	12
	4	R Temporal Parietal Junction (TPJ)	52	-60	-2
Allocentric spatial relations					
	1	L Intraparietal sulcus (IPS1)	-26	-72	24
	2	L Ventral Intraparietal Complex (VIP)	-16	-68	58
	3	R Ventral Intraparietal Complex (VIP)	16	-62	56
	4	L Secondary Somatosensory Cortex (S2)	-38	-38	48
	5	R Secondary Somatosensory Cortex (S2)	45	-40	63
	6	R V3A/V3B	34	-78	16
	7	L Middle Temporal Area (MT)	-48	-72	2
	8	L Fundal Superior Temporal (FST)	-48	-66	-6
Tools					
	1	L Lateral Interparietal Area (LIP)	24	-58	64
	2	R Ventral Intraparietal Complex (VIP)	-22	-54	58
	3	L Middle Temporal Area (MT)	-46	-76	6
	4	L Temporal Parietal Junction (TPJ)	-58	-72	0
	5	R Temporal Parietal Junction (TPJ)	56	-68	4

Selectivity for part relations. To test whether these ROIs are *selective* for object-centered part relations, we examined the response in this region (relative to fixation; see Material and Methods) to (1) activation in the relations blocks of the part relations localizer (independent runs), as well as the other dorsal conditions, namely, (2) distance as determined from the allocentric relations localizer, (3) 3D shape from the depth localizer, and (4) tools from the tool localizer.

A repeated-measures ANOVA with ROI (pIPS, aIPS), hemisphere (left, right), and condition (part relations, distance, 3D shape, tools) as within-subjects factors revealed that there was a significant main-effect of condition, $F(3, 24) = 8.26$, $p < .001$, $\eta_p^2 = 0.53$. There were no other main-effects or interactions (*ps* > .102). Post-hoc comparisons (Holm-Bonferroni corrected) revealed that activation to the part-relations condition was higher than distance ($t[11] = 4.64$, $p < .001$, $d = 1.55$), 3D shape ($t[11] = 4.16$, $p = .002$, $d = 1.39$), and tool ($t[11] = 4.48$, $p = .008$, $d = 1.16$) conditions. Thus, these analyses suggest that the dorsal pathway represents object-centered part relations, and that this representation is independent of allocentric spatial relations and other object properties represented by the dorsal pathway.

Although these analyses did not reveal a significant difference between left and right hemisphere ROIs, examination of the group map suggests that the part relations may be more strongly represented in the right hemisphere. To explore these possible differences, we also analyzed each

ROI separately. Note, due to the exploratory nature of this analysis, these results should be interpreted with caution.

Separate repeated measures ANOVAs were conducted for participants' left and right pIPS and aIPS which revealed main-effects of condition in all four regions (lpIPS: $F[3, 33] = 3.92$, $p = .021$, $\eta_p^2 = 0.33$; rpIPS: $F[3, 33] = 8.70$, $p < .001$, $\eta_p^2 = 0.44$; laIPS: $F[3, 33] = 4.69$, $p = .009$, $\eta_p^2 = 0.34$; raIPS: $F[3, 33] = 12.57$, $p < .001$, $\eta_p^2 = 0.53$), with the response to part relations numerically highest in each region (see Figure 4). However, post-hoc comparisons (Holms-Bonferroni corrected) revealed that activation to part relations was statistically highest only in the right hemisphere parietal regions, but not the left hemisphere parietal regions. Namely, in the right hemisphere, the activation to part relations was significantly higher than distance (rpIPS: $t[11] = 4.66$, $p < .001$, $d = 1.34$; raIPS: $t[11] = 4.18$, $p < .001$, $d = 1.21$), 3D shape (rIPS: $t[11] = 3.47$, $p = .006$, $d = 1.00$; raIPS: $t[11] = 5.77$, $p < .001$, $d = 1.67$), and tools (rpIPS: $t[11] = 4.05$, $p = .001$, $d = 1.17$; raIPS: $t[11] = 4.52$, $p < .001$, $d = 1.31$). By contrast, in the left hemisphere, pIPS responses to part relations were higher than distance ($t[11] = 3.21$, $p = .023$, $d = 1.07$), but not 3D shape or tools ($ts < 2.65$, $ps > .071$, $ds < 0.88$). In left aIPS, responses were higher than distance ($t[11] = 3.51$, $p = .010$, $d = 1.1$) and 3D shape ($t[11] = 2.87$, $p = .039$, $d = 0.91$), but not tools ($t[11] = 2.39$, $p = .097$, $d = 0.75$). In combination with the group statistical map (Figure 3), these results suggest that object-centered part relations may be represented more strongly in the right than left hemisphere parietal regions.

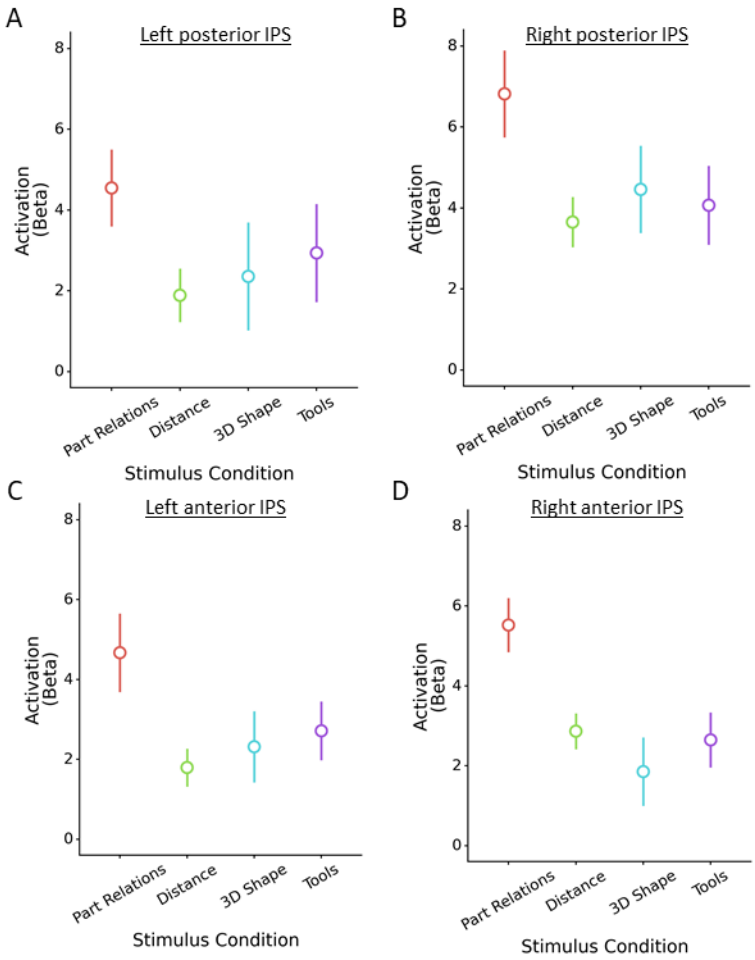


Figure 4. Activation to the part relations (left-out runs), allocentric distance, 3D shape, and tools conditions in (A) left pIPS and (B) right pIPS, (C) left aIPS, and (D) right aIPS. Activation values reflect the standardized parameter estimate. Error bars reflect standard error of the mean.

Conjunction analyses. To explore further the degree to which parietal regions involved in computing part relations overlap with regions computing other dorsal properties, we conducted whole-brain conjunction analyses. First, group-averaged statistical maps were created for every localizer and a cluster-correction threshold applied ($p < .001$; see Figure 3). The resulting statistical maps were consistent with prior research on the neural basis of the allocentric relations (Zachariou et al., 2017) and of tool representations (Chen et al., 2016; Gallivan et al., 2013). No significant clusters were found for the activation profiles on the depth localizer (Georgieva et al., 2008).

Next, we calculated the proportion of independent and overlapping voxels by converting the thresholded statistical map from each group-averaged localizer into binary masks and overlying them with the thresholded statistical map from part relations localizer. Binomial tests revealed that, in right pIPS, there were significantly more independent than overlapping voxels that responded to part relations. Here, the allocentric relations ROI had the greatest amount of overlap with part relations ROI in pIPS (overlapping voxels: 42%, $p < .001$). There were no overlapping voxels from the depth or tool ROIs above the cluster corrected threshold. By contrast, in right aIPS, there were significantly more voxels that overlapped with the allocentric relations ROI than were

independent (overlapping voxels: 65%, $p < .001$). There was also overlap with the tool ROIs (overlapping voxels: 43%, $p < .001$), but there were significantly more independent voxels than overlapping ones. There were no overlapping voxels with the depth localizer (0%). Thus, part relations were represented along a gradient within the dorsal stream, with both distinct and overlapping components.

Finally, to visualize this gradient better, statistical maps were converted into proportions, such that, for each voxel, a value closer to 1 indicates a greater response to part relations and a value closer 0 indicates a greater response to one of the other dorsal properties (e.g., allocentric relations; see Figure 5). Consistent with the analyses above, these maps reveal the least overlap between part relations and other dorsal ROIs in pIPS and the most overlap in aIPS.

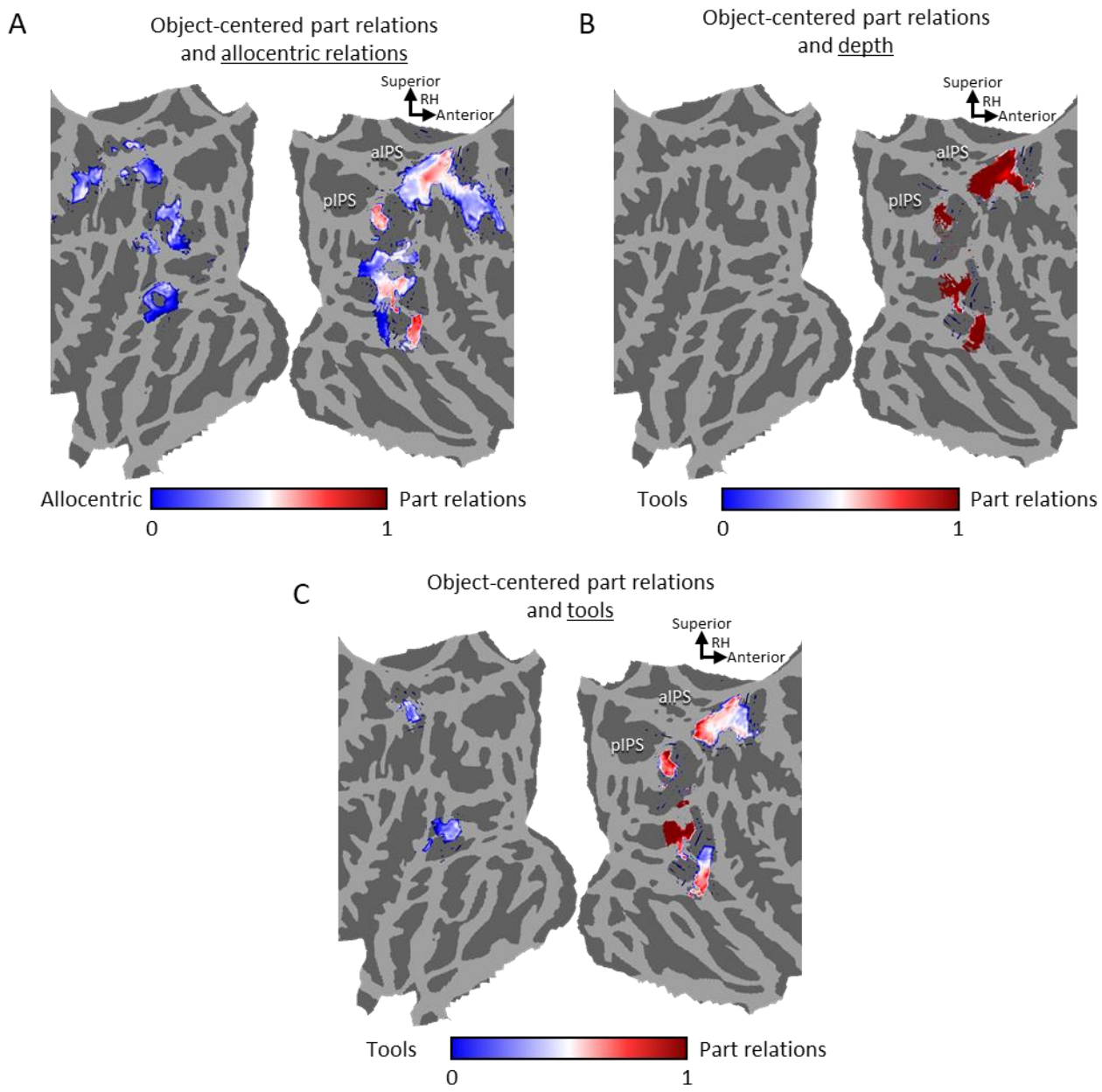


Figure 5. Conjunction maps illustrating areas of distinct and overlapping coding for object-centered part relations and (A) allocentric relations, (B) depth, and (C) tools. A value closer 1 indicates a greater response to part relations; a value closer to 0 indicates a greater response to the control localizer. Maps are zoomed in on the visual cortex for easier inspection.

Task-dependent functional connectivity. If the role of the dorsal pathway in object recognition is to compute object-centered part relations, then a prediction is that pIPS and aIPS will also be functionally connected to the ventral pathway – the nexus of object recognition processing. More specifically, the prediction is that functional connectivity between right and left pIPS or aIPS with ventral cortex will depend on the task demands, such that connectivity would be greatest when perception of part relations is needed, as in the relations, but not feature condition of the localizer. To test this prediction, we conducted PPI analyses to examine whether there was task-dependent

functional connectivity between left and right pIPS and aIPS regions involved in computing object-spatial relations, and ventral regions involved in object recognition (see Materials and Methods).

Examination of the group map (Figure 6) revealed significant connectivity between right hemisphere pIPS and aIPS with bilateral ventral stream regions. Interestingly, there was relatively little connectivity with other dorsal regions, suggesting that the function of right hemisphere pIPS and aIPS may be specifically in the service of object recognition processes in the ventral stream rather than action processes in other dorsal regions. There was no significant connectivity with left pIPS or aIPS that survived FDR correction.

To further examine the specificity of task-dependent connectivity to these regions, we reanalyzed the data from the part relations localizer using the peak voxel from the allocentric relations ROI in the left hemisphere as our seed region. This ROI was chosen because it does not overlap with part relations ROIs, but nevertheless has a conceptually similar representation. These analyses revealed no significant connectivity between allocentric relations ROIs in the left hemisphere and the ventral visual pathway. Moreover, a direct comparison between regions (Holm-Bonferroni corrected), revealed that task-dependent connectivity with LOC, a ventral object region, was significantly stronger with right pIPS (lLOC: $t(11) = 3.41$, $p = .005$, $d = 0.99$; rLOC: $t(11) = 3.28$, $p = .007$, $d = 0.95$) and aIPS (lLOC: $t(11) = 4.36$, $p < .001$, $d = 1.26$; rLOC: $t(11) = 4.56$, $p < .001$, $d = 1.32$) than left allocentric relations ROIs. There were no differences in connectivity between the other pIPS and aIPS regions ($ps > .217$). Together, these findings suggest that dorsal regions involved in computing object-centered part relations, particularly in the right hemisphere are preferentially connected to the ventral stream to support object recognition.

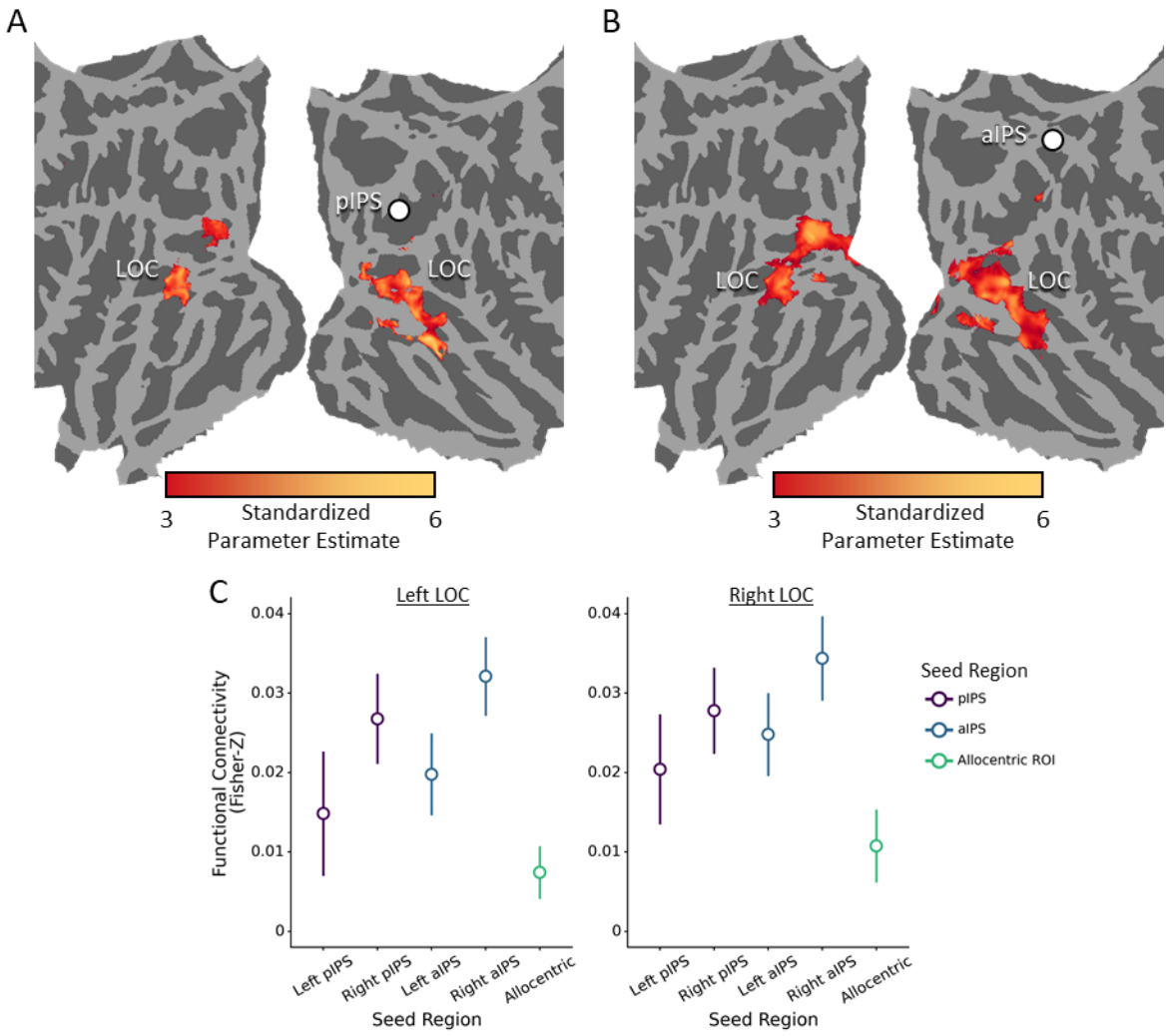


Figure 6. Task-based functional connectivity results. (A-B) Functional connectivity map (zoomed in on the visual cortex) for (A) right pIPS and (B) right aIPS. Seed regions are displayed as white circles. There was no functional connectivity above the cluster corrected threshold in left pIPS, left aIPS, or the left allocentric ROI. (C) Plots comparing the connectivity between pIPS, aIPS, and the other ROIs in left LOC and right LOC ROIs. Error bars reflect standard error of the mean.

Experiment 2: Dorsal contributions to object recognition

Category decoding. To test whether dorsal regions that compute object-centered part relations contribute to object recognition, we examined whether multivariate pattern within these regions could be used to classify objects (see Figure 2). Using a 20-fold cross-validation procedure, an SVM classifier was trained on the multivariate pattern for three exemplars from each category, and then predicted the category of the two left out exemplars.

One-sample comparisons to chance (0.20) revealed that category decoding was significantly above chance in right pIPS, M = 32.7%, $t(11)$ = 3.15, p = .009, d = 0.91, but not in right aIPS, left pIPS or left aIPS defined on the basis of part relation (Ms < 23.4%, ps > .110, ds < 0.72; Figure 7). To further examine the specificity of category decoding the dorsal stream, we also tested how well a left hemisphere allocentric relations ROI can decode object categories. These analyses revealed that

decoding was not above chance in the left allocentric ROI, M = 18.7%, $t(11)$ = -0.82, p = .780, d = 0.23. Direct comparisons between right pIPS and the other regions (Holm-Bonferroni corrected) further confirmed that, categorization accuracy was significantly higher in right pIPS than left allocentric regions ($t[11]$ = 3.88, p = .004, d = 1.23 and laIPS ($t[11]$ = 4.32, p = .001, d = 1.37), though not raIPS ($t[11]$ = 2.65, p = .096, d = 0.837) nor lpIPS ($t[11]$ = 2.48, p = .127, d = 0.78). Next, we examined how category decoding in the dorsal stream compares to ventral stream object recognition regions, namely LOC. As would be expected, categorization accuracy was above chance in left and right LOC, (lLOC: M = 26.0% , $t[11]$ = 3.56, p = .004, d = 1.03; rLOC: M = 26.7% $t[11]$ = 2.30, p = .042, d = 0.66), with the neither region differing significantly from right pIPS (ts < 1.62, ps > .357, ds < 0.42). Thus, regions in right posterior IPS involved in computing object-centered part relations can support categorization of object exemplars.

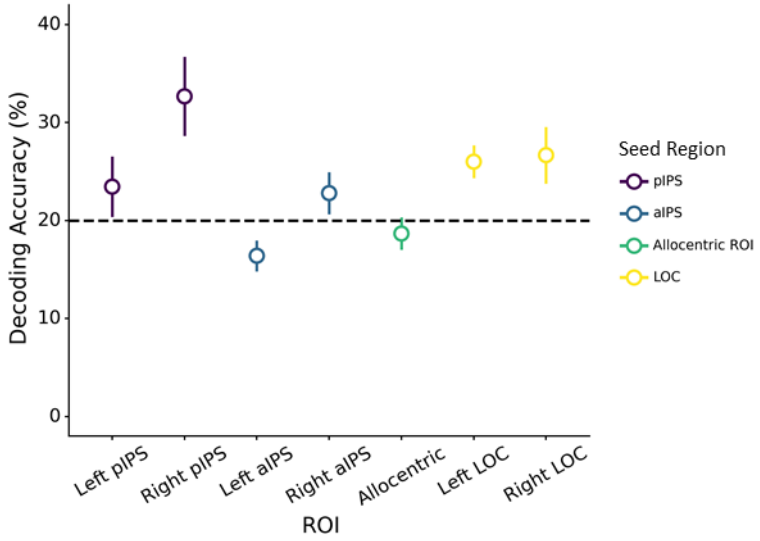


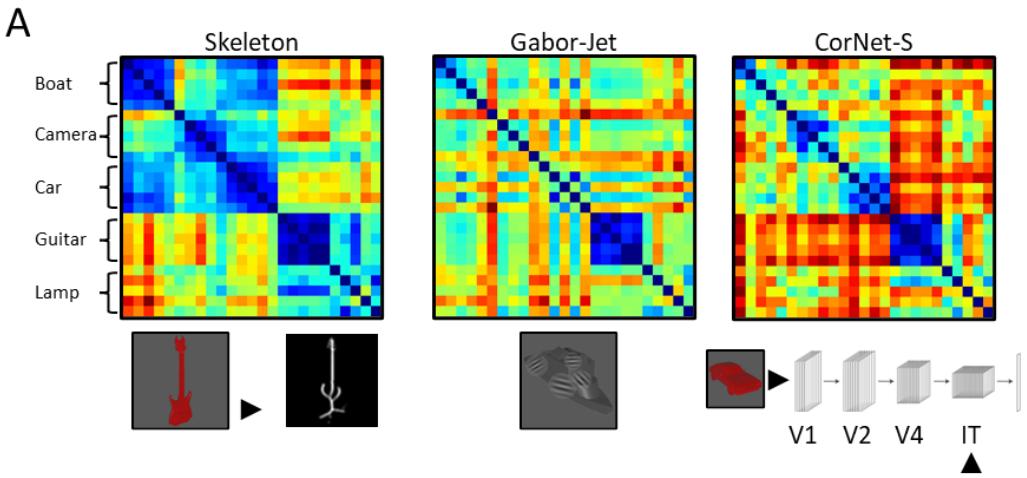
Figure 7. Object categorization accuracy for pIPS, aIPS, the left allocentric ROI, and LOC. Error bars reflect standard error of the mean.

Representational content of dorsal ROIs. The results above show that a region in pIPS defined on the basis of part relations can be used to decode the category of objects. Yet, despite the fact that this region was defined using a part relations localizer, it is possible that categorization was accomplished using other visual properties. Indeed, it is well known that pIPS retains a retinotopic organization (Wang et al., 2014) and is tightly connected to early visual cortex (Greenberg et al., 2012). Thus, it is possible that the categorization performance of right pIPS may have been achieved on the basis of low-level image-level similarity. Moreover, it is unclear to what degree categorization in right pIPS is accomplished using high-level visual representations distinct from those in the ventral stream.

To examine whether right pIPS accomplished object categorization on the basis of object-centered part relations, we used representational similarity analyses (RSA). Specifically, we tested whether a skeletal model, which approximates object-centered part relations, explains unique variance in pIPS over and above other models of vision (see Materials and Methods). Skeletal models describe the spatial arrangement of object parts via a low-dimensional symmetry axis (see Figure 8A). For comparison, we also tested Gabor-jet (GBJ), a model of low-level image similarity (Margalit et al., 2016; see Figure 8A), as well as CorNet-S a neural network model whose upper layers approximate

the response profile of high-level ventral regions in monkeys (Kubilius et al., 2019; Schrimpf et al., 2018; see Figure 8A).

To test whether the skeletal model explained unique variance in right pIPS, we conducted linear regression analyses with the neural RDM from pIPS as the dependent variable and the different models of visual similarity as predictors (Skeleton ∪ GBJ ∪ CorNet-S; see Figure 8B). These analyses revealed that only skeletal model explained unique variance in right pIPS (β = 0.33, p < .001), not the other models (GBJ: β = 0.04, p = .493; CorNet-S: β = -0.02, p = .839). To examine the specificity of this effect to rpIPS, we also examined whether the skeletal model explained unique variance in the left pIPS, as well as left and right aIPS ROI defined on the basis of object-centered relations. The skeletal model did not explain significant unique variance in any of these ROIs (β < 0.14, ps > .068; see Figure 8B-C). These results suggest that categorization in right pIPS was accomplished by representing the object-centered relations, not other low- or high-level visual properties.



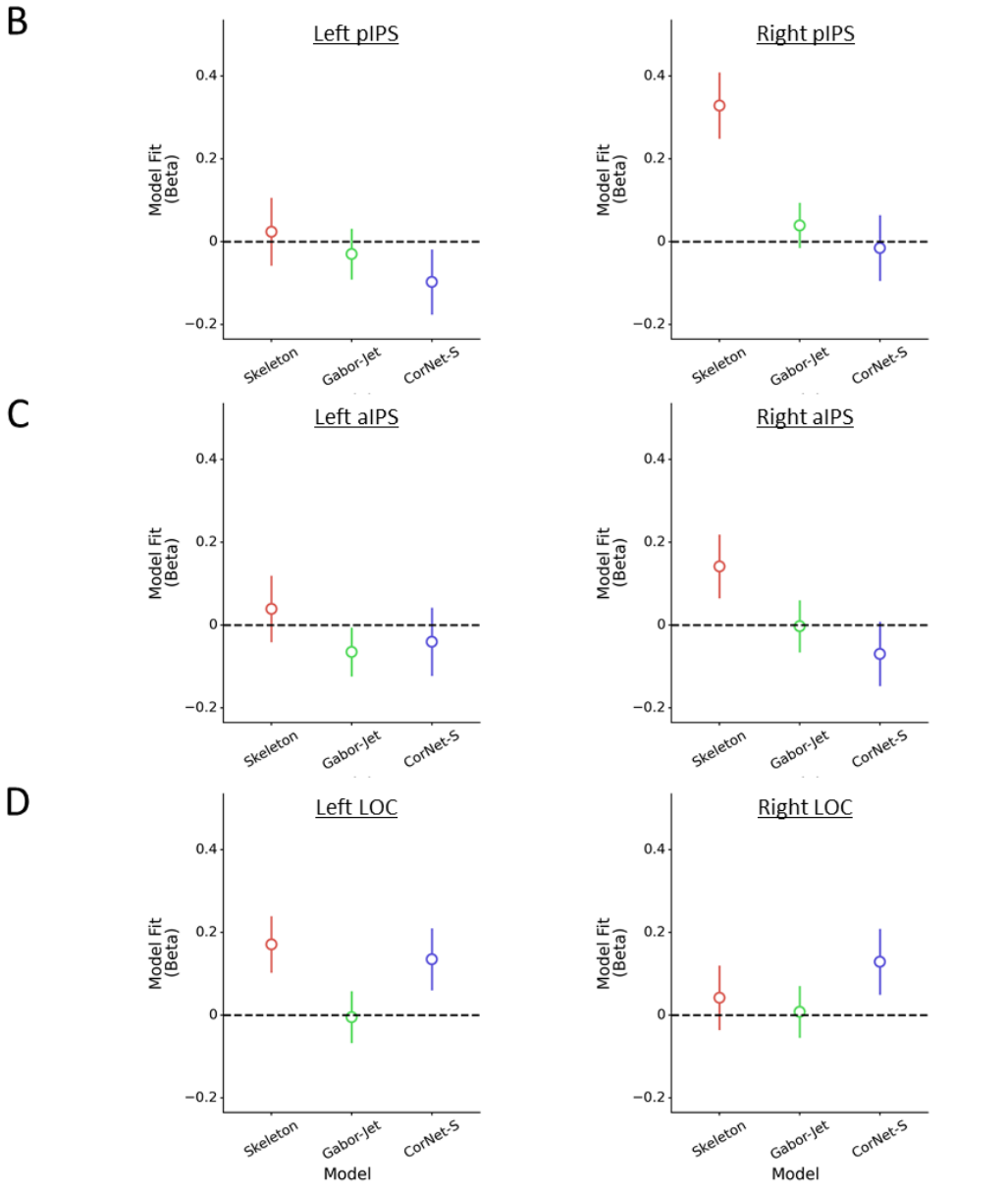


Figure 8. Models and results of the representational similarity analyses (RSA). (A) Representational dissimilarity matrices (RDMs) and a schematic illustration of the (left) the skeletal model, (middle) Gabor-jet model, and (right) CorNet-S. (B-D) Standardized coefficients (Betas) from the linear regression analyses examining the fit of the skeletal, Gabor-jet, and CorNet-S models for left and right (B) pIPS, (C) aIPS), and (D) LOC.

Unique contributions of dorsal ROIs to ventral stream processing. Next, we examined whether right pIPS represents distinct visual information from ventral object regions such as LOC. We repeated the linear regression analyses, except here we used neural RDMs from left and right LOC as the dependent variable. These analyses revealed that, the skeletal model explained unique variance in left (β = 0.17, p = .023), but not right LOC (β = 0.04, p = .582; see Figure 8D).

This finding raises the question: do regions of dorsal cortex compute object-centered part relations and then transmit that information to ventral cortex for object recognition? Or, are part relations computed in the ventral pathway, as previously proposed (Ayzenberg et al., accepted; Behrmann et al., 2006) and transmitted to dorsal regions such as right pIPS? Alternatively, part relations may be coded in parallel in both pathways. To investigate these possibilities, we examined whether multivariate response in pIPS mediates the relation between the skeletal model and the neural RDM in LOC. In other words, we tested whether skeletal coding in LOC is represented independently or by way of right pIPS.

To test these possibilities, we first repeated the linear regression analyses in left LOC, but this time we included the neural RDM from right pIPS in addition to the skeleton, GBJ, and CorNet-S models. With pIPS included as a predictor, the skeletal model no longer explained unique variance in left LOC (β = 0.07, p = .345), only right pIPS (β = 0.31, p < .001) and CorNet-S (β = 0.14, p = .053) explained unique variance. By contrast, when linear regression analyses are conducted on right pIPS with the left LOC RDM as a predictor in addition to the skeleton, GBJ, and CorNet-S models, both the skeleton model (β = 0.28, p < .001) and left LOC RDM (β = 0.30, p < .001) explain unique variance. Finally, a mediation analysis (with GBJ and CorNet-S as covariates) confirmed that right pIPS fully mediated the relation between the skeletal model and left LOC (b = 0.10, 95% CI [.05, .16]). There was no direct relation otherwise (b = 0.07, 95% CI [-.074, .21]). By contrast, when left LOC is used as a mediator between the skeletal model and right pIPS, there continues to be a direct relation between the skeletal model and right pIPS (b = 0.28, 95% CI [0.14, 0.42]). Here, left LOC acts as only a partial mediator (b = 0.05, 95% CI [0.00, 0.10]). Together these results suggest that object-centered part relations, as approximated by a skeletal model, are computed in right pIPS independently of ventral regions. Moreover, representations of part relation in ventral regions such as left LOC may arise via input from right pIPS.

Multivariate connectivity. Thus far, we have documented that a ROI in pIPS, particularly in the right hemisphere, is sensitive to object-centered part relations, able to categorize objects, and account for the representation of part relations in the ventral stream. Together, these results suggest that this region interacts with ventral regions in support of object recognition. To provide converging evidence for this result, we used multivariate pattern dependence (MVPD) analyses to test whether right pIPS also exhibits functional connectivity with ventral stream regions during object viewing (see Materials and Methods).

Examination of the group map (Figure 9B) revealed broad connectivity between both right pIPS with bilateral dorsal and ventral regions. To examine the specificity of this interaction between right pIPS and ventral regions, we also examined the multivariate connectivity patterns of left pIPS and bilateral aIPS defined on the basis of part relations. Like right pIPS, these regions also showed broad connectivity with bilateral dorsal and ventral regions (see Figure 9). Direct comparisons between these ROIs (Holm-Bonferroni corrected), revealed that connectivity between right pIPS and bilateral LOC was stronger than both left aIPS (lLOC: $t(11) = 3.09$, $p = .028$, $d = 0.97$; rLOC: $t(11) = 3.77$, $p = .005$, $d = 1.19$) and right aIPS (lLOC: $t(11) = 2.62$, $p = .072$, $d = 0.83$; rLOC: $t(11) = 3.16$, $p = .019$, $d = 1.00$). There were no differences between left and right IPS ($ps > .312$, $ds < 0.70$), nor among the other ROIs ($ps > .130$, $ds < 0.71$) Together, these findings suggest that right pIPS regions involved in computing object-centered part relations are connected to the ventral pathway.

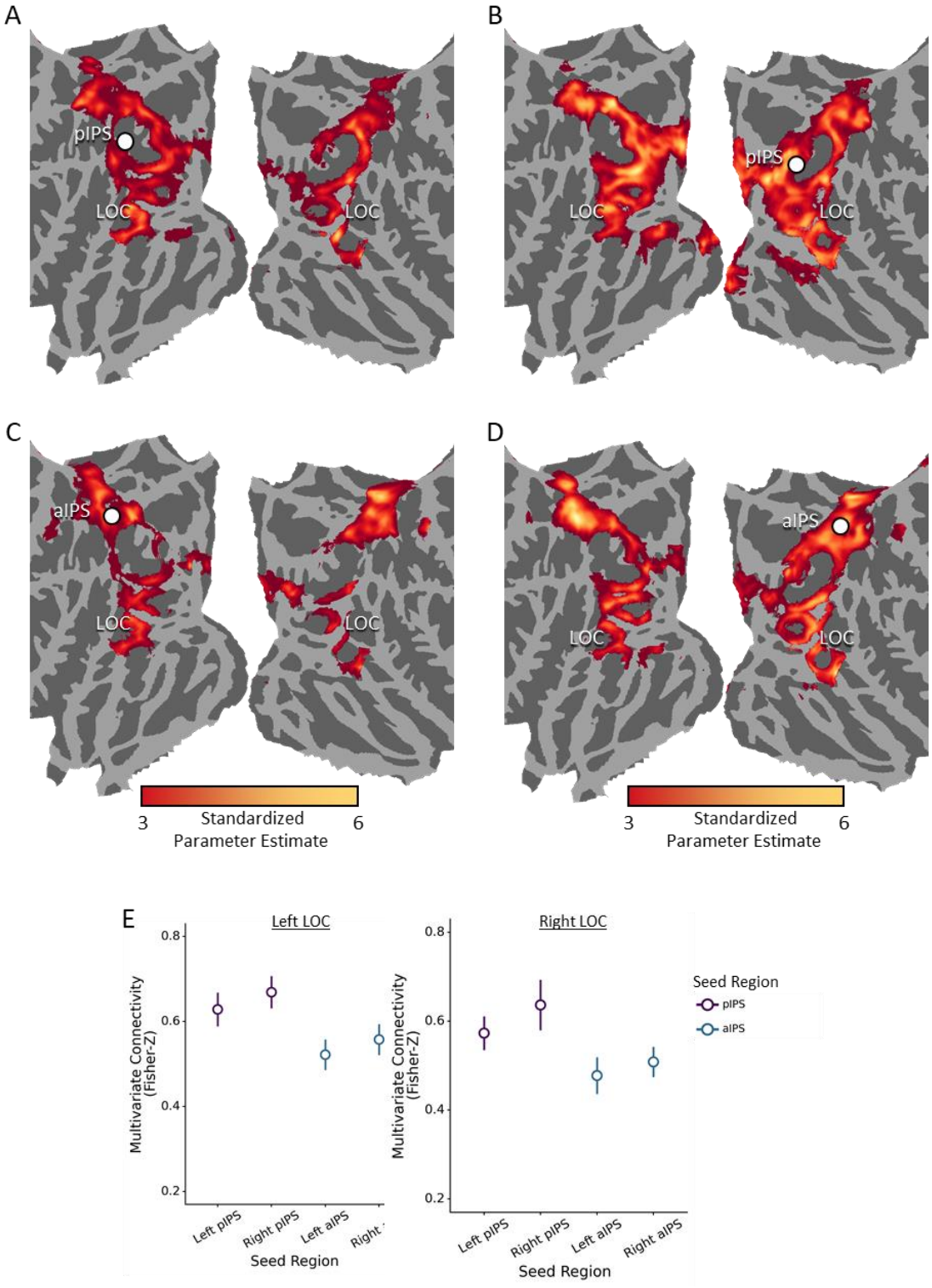


Figure 9. Multivariate functional connectivity results. (A-D) Functional connectivity map for (A) left pIPS, (B) right pIPS, (C) left aIPS, and (D) right aIPS. Seed regions are displayed as a white circle. (C) Plots comparing the connectivity between ROIs in left LOC and right LOC ROIs. Error bars reflect standard error of the mean.

General Discussion

Here, we examined the contribution of the dorsal visual pathway to object recognition. Given its sensitivity to spatial information and its contribution to object perception (Freud et al., 2020), we hypothesized that dorsal cortex may compute the relations among an object's parts and transmit this information to ventral cortex to support object recognition. We found that regions in posterior and anterior IPS, particularly in the right hemisphere, displayed selectivity for part relations independent of allocentric spatial relations and other dorsal object representations, such as 3D shape and tools. Importantly, these regions also exhibited task-dependent functional connectivity with ventral regions, such that connectivity increased when part relations differed.

Next, we found that object category could be decoded successfully in right pIPS, with categorization performance comparable to ventral object regions. Similarity analyses further confirmed that decoding in right pIPS was supported by a representation of part relations, as approximated by a skeletal model, and not by low- or high-level image properties. Crucially, we found that the multivariate response in right pIPS mediated representations of part relations in ventral cortex, with pIPS also exhibiting higher multivariate connectivity with ventral cortex than several other parietal regions. Together, these findings highlight how object-centered part relations, a property crucial for object recognition, are represented neurally, and validate the strong link between dorsal and ventral visual cortex in accomplishing object recognition.

Neural representations of object-centered part relations

Many studies have examined how allocentric spatial information is represented neurally, but few have explored the representations of object-centered part relations. Lescroart and Biederman (2012) decoded the spatial arrangements of object parts in both ventral and dorsal cortices, but did not test whether these were independent of other dorsal representations nor whether other visual properties influenced the decoding. Ayzenberg et al. (accepted) identified ventral regions that coded for part relations (as approximated by a skeletal model) independent of other visual properties, with strongest coding in left LOC – a finding consistent with the RSA results of the current study. However, they did not investigate whether such representations also exist in dorsal cortex and could account for their effects. Finally, Behrmann et al. (2006) reported that patients with LOC damage and object recognition deficits were impaired in perceiving part relations, but not the features of object parts, suggesting a ventral locus for object-centered relations.

Consistent with these studies, we too found that part relations are represented in ventral cortex. However, our data suggest that this information arises via input from dorsal cortex. We documented functional connectivity between IPS and LOC, and show that right pIPS mediates the representation of part relations in ventral regions, and not the other way around. This finding is compatible with research showing that visual object information reaches posterior parietal cortex earlier than ventral regions (Regev et al., 2018), that topological object properties may only become represented in the ventral pathway through feedback connections (Bar et al., 2006; Wang et al., 2020), and that temporary inactivation of posterior parietal regions impairs ventral object processing (Van Dromme et al., 2016; Zachariou et al., 2017). Altogether, these studies suggest a causal role for dorsal cortex in ventral object processing.

Our results also uncovered a posterior-to-anterior gradient, especially evident in Experiment 2. Although selectivity for part relations was found in both pIPS and aIPS, only right pIPS was able to decode object category. Moreover, right pIPS exhibited the highest multivariate functional

connectivity with LOC, and its representation of object similarity was most consistent with a model of part relations (i.e., medial axis skeleton). This gradient may reflect a common organizing principle of the dorsal pathway. Regions of posterior parietal cortex exhibit greater sensitivity for object properties in the service of recognition (Freud et al., 2017; Gillebert et al., 2015; Van Dromme et al., 2016), and greater connectivity to ventral object regions (Janssen et al., 2018; Takemura et al., 2016; Webster et al., 1994). By contrast, anterior parietal cortex shows greater sensitivity to properties that afford action (Chao & Martin, 2000; Chen et al., 2017; Chen et al., 2016; Culham et al., 2003). Whereas right pIPS may be more involved in computing part relations for the purpose of recognition, right aIPS may be more involved in computing part relations to help coordinate grasping behaviors. Relatedly, we found greater overlap between right aIPS and regions involved in representing allocentric relations and tools – which are both critical for coordinating action.

Object-centered relations and other dorsal representations

We found that IPS regions responded more to object-centered part relations than allocentric relations, 3D shape, and tools, suggesting selectivity in these regions. However, our conjunction analyses also revealed that object-centered relations may be represented along a continuum in parietal cortex, with varying degrees of overlap with other dorsal properties, particularly, with allocentric spatial relations. The overlap between object-centered and allocentric relations in parietal cortex may reflect a broader organizing principle for spatial coding in dorsal cortex in which reference frames are organized topographically. Recent evidence suggests that the dorsal pathway represents visual information at different spatial scales ranging from single objects to large, multi-object perspectives (Josephs & Konkle, 2020). This possibility is also consistent with a rich literature on hemi-spatial neglect, in which right parietal damage impairs object perception on the left side of space (Caramazza & Hillis, 1990; Corbetta & Shulman, 2011; Heilman & Valenstein, 1979), and that, depending on the scope of the damage, multiple reference frames are often affected simultaneously, suggesting that the representations overlap or abut (Halligan et al., 2003; Medina et al., 2009). Thus, although our results show selectivity for part relations, we suggest that such representations are situated within a broader topographic map for spatial coding.

We found relatively little overlap between regions involved in representing part relations and those involved in representing tools – with overlap occurring exclusively in aIPS. This finding is consistent with the hypothesis formulated earlier, that coding of part relations in aIPS may be in support of coordinating grasping behaviors. It is important to note that here we used a particularly stringent definition of tool ROIs, wherein tools were contrasted with other manipulable objects (Chen et al., 2018), and this decision may have minimized the degree to which we observed activity related to object action affordances (since all stimuli afforded action). Moreover, by using objects with elongated axes in the part-relations localizer (an important indicator of action affordance; Chen et al., 2017), we may have further suppressed the degree to which regions representing part relations overlapped with those representing tools. Future work may use a more direct object affordance localizer (Freud et al., 2018; Snow et al., 2011) and a more variable stimulus set to localize part relations.

Finally, extensive pilot work (Ayzenberg et al., unpublished data) suggested that depth regions in parietal cortex could be reliably localized with the 3D and 2D shape stimuli used here. However, we were unable to do so in current study – precluding conjunction analyses. Two runs of the depth localizer may have been insufficient to identify regions involved in processing 3D shape, and/or

depth from shading (as used here) may be less consistently represented than depth from texture or disparity (Georgieva et al., 2008). Given that the computation of depth structure in the dorsal pathway is critical for object recognition (Farivar, 2009; Freud et al., 2020; Van Dromme et al., 2016; Welchman, 2016), future work is required to explore the link between regions subserving part relations and 3D shape.

The role of object-centered part relations in object recognition

Representations of object-centered part relations are thought to be critical for object recognition because they describe an object's global shape structure – a key organizing feature of most basic-level categories (Barenholtz & Tarr, 2006; Hummel, 2000; Mervis & Rosch, 1981). Such a representation may even support rapid object learning in infancy when experience with objects is minimal (Ayzenberg & Lourenco, 2021; Kraebel & Gerhardstein, 2006; Rakison & Butterworth, 1998). Yet, ANNs, the current best models of human object recognition, are largely insensitive to the relations among object parts and require extensive object experience to categorize novel objects (Baker et al., 2018; Baker et al., 2020). One potential reason for this deficit is that most current ANNs exclusively model ventral stream processes (Blauch et al., 2021; Schrimpf et al., 2020; Yamins et al., 2014). Indeed, the few ANNs that model dorsal cortex focus on action or motion related processes (Güçlü & van Gerven, 2017; Mineault et al., 2021). Here, we propose that the dorsal pathway may play a key role in object recognition by computing the object-centered part relations and propagating these signals to ventral object regions. Right pIPS, in particular, may be important for object recognition, in that its multivariate response was sufficient to decode object category and it was well explained by an object recognition model that computes the part relations (i.e., a skeletal model). Importantly, we consistently found connectivity between right pIPS regions and regions in ventral cortex, with evidence that right pIPS may even mediate the representation of part relations in LOC. Thus, by incorporating the dorsal pathway with the ventral pathway, we may gain a better understanding of the broader network that supports object recognition and the relative contributions of each pathway.

References

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., . . . Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics, 8*(14). doi:10.3389/fninf.2014.00014

Anzellotti, S., Caramazza, A., & Saxe, R. (2017). Multivariate pattern dependence. *PLOS Computational Biology, 13*(11), e1005799.

Ayzenberg, V., Chen, Y., Yousif, S. R., & Lourenco, S. F. (2019). Skeletal representations of shape in human vision: Evidence for a pruned medial axis model. *Journal of Vision, 19*(6), 1-21. doi:10.1167/19.6.6

Ayzenberg, V., Kamps, F. S., Dilks, D. D., & Lourenco, S. F. (accepted). Skeletal representations of shape in the human visual cortex. *Neuropsychologia*, 799650.

Ayzenberg, V., Kubert, J., Dilks, D. D., & Lourenco, S. F. (unpublished data). The dorsal stream faciliatates viewpoint-invariant object recognition.

Ayzenberg, V., & Lourenco, S. F. (2019). Skeletal descriptions of shape provide unique perceptual information for object recognition. *Scientific Reports, 9*(1), 1-13. doi:10.1038/s41598-019-45268-y

Ayzenberg, V., & Lourenco, S. F. (2021). The shape skeleton supports one-shot categorization in human infants: Behavioral and computational evidence. *PsyArxiv*.

Baker, N., Lu, H., Erlikhman, G., & Kellman, P. J. (2018). Deep convolutional networks do not classify based on global object shape. *PLOS Computational Biology, 14*(12), e1006613. doi:10.1371/journal.pcbi.1006613

Baker, N., Lu, H., Erlikhman, G., & Kellman, P. J. (2020). Local features and global shape information in object classification by deep convolutional neural networks. *Vision Research, 172*, 46-61.

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., . . . Halgren, E. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences of the United States of America, 103*(2), 449-454. doi:10.1073/pnas.0507062103

Barenholtz, E., & Tarr, M. J. (2006). Reconsidering the Role of Structure in Vision. In *Psychology of learning and motivation* (Vol. 47, pp. 157-180): Academic Press.

Behrmann, M., Peterson, M. A., Moscovitch, M., & Suzuki, S. (2006). Independent representation of parts and the relations between them: evidence from integrative agnosia. *Journal of Experimental Psychology: Human Perception and Performance, 32*(5), 1169-1184.

Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. *Psychological Review, 94*(2), 115-147.

Blauch, N. M., Behrmann, M., & Plaut, D. C. (2021). A connectivity-constrained computational account of topographic organization in high-level visual cortex. *PNAS*.

Blum, H. (1973). Biological shape and visual science (Part I). *Journal of Theoretical Biology, 38*(2), 205-287.

Bracci, S., & Op de Beeck, H. (2016). Dissociations and associations between shape and category representations in the two visual pathways. *Journal of Neuroscience, 36*(2), 432-444.

Caramazza, A., & Hillis, A. E. (1990). Levels of representation, co-ordinate frames, and unilateral neglect. *Cognitive Neuropsychology, 7*(5-6), 391-445.

Chang, A. X., Funkhouser, T., Guibas, L., Hanrahan, P., Huang, Q., Li, Z., . . . Su, H. (2015). Shapenet: An information-rich 3d model repository. *arXiv preprint arXiv:1512.03012*.

Chao, L. L., & Martin, A. (2000). Representation of manipulable man-made objects in the dorsal stream. *Neuroimage, 12*(4), 478-484.

Chen, J., Snow, J. C., Culham, J. C., & Goodale, M. A. (2017). What Role Does "Elongation" Play in "Tool-Specific" Activation and Connectivity in the Dorsal and Ventral Visual Streams? *Cerebral Cortex, 28*(4), 1117-1131. doi:10.1093/cercor/bhx017

Chen, Q., Garcea, F. E., Jacobs, R. A., & Mahon, B. Z. (2018). Abstract representations of object-directed action in the left inferior parietal lobule. *Cerebral Cortex, 28*(6), 2162-2174.

Chen, Q., Garcea, F. E., & Mahon, B. Z. (2016). The representation of object-directed action and function knowledge in the human brain. *Cerebral Cortex, 26*(4), 1609-1618.

Corbetta, M., & Shulman, G. L. (2011). Spatial neglect and attention networks. *Annual review of neuroscience, 34*, 569-599.

Culham, J. C., Danckert, S. L., De Souza, J. F., Gati, J. S., Menon, R. S., & Goodale, M. A. (2003). Visually guided grasping produces fMRI activation in dorsal but not ventral stream brain areas. *Experimental Brain Research, 153*(2), 180-189.

Dimitrov, P., Damon, J. N., & Siddiqi, K. (2003). *Flux invariants for shape.* Paper presented at the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.

Farivar, R. (2009). Dorsal–ventral integration in object recognition. *Brain research reviews, 61*(2), 144-153.

Feldman, J., & Singh, M. (2006). Bayesian estimation of the shape skeleton. *Proceedings of the National Academy of Sciences, 103*(47), 18014-18019.

Freud, E., Behrmann, M., & Snow, J. C. (2020). What does dorsal cortex contribute to perception? *Open Mind, 4*, 40-56.

Freud, E., Culham, J. C., Plaut, D. C., & Behrmann, M. (2017). The large-scale organization of shape processing in the ventral and dorsal pathways. *eLife, 6*, e27576.

Freud, E., Ganel, T., Shelef, I., Hammer, M. D., Avidan, G., & Behrmann, M. (2015). Three-Dimensional Representations of Objects in Dorsal Cortex are Dissociable from Those in Ventral Cortex. *Cerebral Cortex, 27*(1), 422-434. doi:10.1093/cercor/bhv229

Freud, E., Macdonald, S. N., Chen, J., Quinlan, D. J., Goodale, M. A., & Culham, J. C. (2018). Getting a grip on reality: Grasping movements directed to real objects and images rely on dissociable neural representations. *Cortex, 98*, 34-48.

Freud, E., Plaut, D. C., & Behrmann, M. (2016). 'What'is happening in the dorsal visual pathway. *Trends in Cognitive Sciences, 20*(10), 773-784.

Friston, K., Buechel, C., Fink, G., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage, 6*(3), 218-229.

Gallivan, J. P., McLean, D. A., Valyear, K. F., & Culham, J. C. (2013). Decoding the neural mechanisms of human tool use. *eLife, 2*, e00425.

Gauthier, I., & Tarr, M. J. (2016). Visual Object Recognition: Do We (Finally) Know More Now Than We Did? *Annual Review of Vision Science, 2*(1), 377-396. doi:10.1146/annurev-vision-111815-114621

Georgieva, S. S., Todd, J. T., Peeters, R., & Orban, G. A. (2008). The Extraction of 3D Shape from Texture and Shading in the Human Brain. *Cerebral Cortex, 18*(10), 2416-2438. doi:10.1093/cercor/bhn002

Gillebert, C. R., Schaeverbeke, J., Bastin, C., Neyens, V., Bruffaerts, R., De Weer, A.-S., . . . Vandenberghe, R. (2015). 3D Shape Perception in Posterior Cortical Atrophy: A Visual Neuroscience Perspective. *The Journal of Neuroscience, 35*(37), 12673-12692. doi:10.1523/jneurosci.3651-14.2015

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences, 15*(1), 20-25.

Greenberg, A. S., Verstynen, T., Chiu, Y.-C., Yantis, S., Schneider, W., & Behrmann, M. (2012). Visuotopic cortical connectivity underlying attention revealed with white-matter tractography. *Journal of Neuroscience, 32*(8), 2773-2782.

Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research, 41*(10), 1409-1422.

Güçlü, U., & van Gerven, M. A. (2017). Increasingly complex representations of natural movies across the dorsal stream are shared between subjects. *Neuroimage, 145*, 329-336.

Halligan, P. W., Fink, G. R., Marshall, J. C., & Vallar, G. (2003). Spatial cognition: evidence from visual neglect. *Trends in Cognitive Sciences, 7*(3), 125-133. doi:https://doi.org/10.1016/S1364-6613(03)00032-9

Haxby, J. V., Grady, C. L., Horwitz, B., Ungerleider, L. G., Mishkin, M., Carson, R. E., . . . Rapoport, S. I. (1991). Dissociation of object and spatial visual processing pathways in human extrastriate cortex. *Proceedings of the National Academy of Sciences, 88*(5), 1621-1625.

Heilman, K. M., & Valenstein, E. (1979). Mechanisms underlying hemispatial neglect. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society, 5*(2), 166-170.

Holler, D. E., Behrmann, M., & Snow, J. C. (2019). Real-world size coding of solid objects, but not 2-D or 3-D images, in visual agnosia patients with bilateral ventral lesions. *Cortex, 119*, 555-568.

Hummel, J. E. (2000). Where view-based theories break down: The role of structure in shape perception and object recognition. In E. Dietrich & A. Markman (Eds.), *Cognitive dynamics: Conceptual change in humans and machines* (pp. 157-185). Hillsdale, NJ: Erlbaum.

Hummel, J. E., & Stankiewicz, B. J. (1996). Categorical relations in shape perception. *Spatial Vision, 10*(3), 201-236.

Janssen, P., Verhoef, B.-E., & Premereur, E. (2018). Functional interactions between the macaque dorsal and ventral visual pathways during three-dimensional object vision. *Cortex, 98*, 218-227.

Jeong, S. K., & Xu, Y. (2016). Behaviorally Relevant Abstract Object Identity Representation in the Human Parietal Cortex. *The Journal of Neuroscience, 36*(5), 1607-1619. doi:10.1523/jneurosci.1016-15.2016

Josephs, E. L., & Konkle, T. (2020). Large-scale dissociations between views of objects, scenes, and reachable-scale environments in visual cortex. *Proceedings of the National Academy of Sciences, 117*(47), 29354-29362.

Julian, J. B., Fedorenko, E., Webster, J., & Kanwisher, N. (2012). An algorithmic method for functionally defining regions of interest in the ventral visual pathway. *Neuroimage, 60*(4), 2357-2364. doi:10.1016/j.neuroimage.2012.02.055

Konen, C. S., & Kastner, S. (2008). Two hierarchically organized neural systems for object information in human visual cortex. *Nature Neuroscience, 11*(2), 224-231.

Kourtzi, Z., & Kanwisher, N. (2001). Representation of perceived object shape by the human lateral occipital complex. *Science, 293*(5534), 1506-1509.

Kraebel, K. S., & Gerhardstein, P. C. (2006). Three-month-old infants' object recognition across changes in viewpoint using an operant learning procedure. *Infant Behavior and Development, 29*(1), 11-23.

Kravitz, D. J., Saleem, K. S., Baker, C. I., & Mishkin, M. (2011). A new neural framework for visuospatial processing. *Nature Reviews Neuroscience, 12*, 217. doi:10.1038/nrn3008

Kubilius, J., Schrimpf, M., Kar, K., Rajalingham, R., Hong, H., Majaj, N., . . . Schmidt, K. (2019). *Brain-like object recognition with high-performing shallow recurrent ANNs.* Paper presented at the Advances in Neural Information Processing Systems.

Kumar, M., Anderson, M. J., Antony, J. W., Baldassano, C., Brooks, P. P., Cai, M. B., . . . Huberdeau, D. (2020). BrainIAK: The brain imaging analysis kit.

Lescroart, M. D., & Biederman, I. (2012). Cortical representation of medial axis structure. *Cerebral Cortex, 23*(3), 629-637.

Mahon, B. Z., Milleville, S. C., Negri, G. A., Rumiati, R. I., Caramazza, A., & Martin, A. (2007). Action-related properties shape object representations in the ventral stream. *Neuron, 55*(3), 507-520.

Margalit, E., Biederman, I., Herald, S. B., Yue, X., & von der Malsburg, C. (2016). An applet for the Gabor similarity scaling of the differences between complex stimuli. *Attention, Perception, & Psychophysics, 78*(8), 2298-2306. doi:10.3758/s13414-016-1191-7

Medina, J., Kannan, V., Pawlak, M. A., Kleinman, J. T., Newhart, M., Davis, C., . . . Hillis, A. E. (2009). Neural substrates of visuospatial processing in distinct reference frames: evidence from unilateral spatial neglect. *Journal of Cognitive Neuroscience, 21*(11), 2073-2084.

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology, 32*(1), 89-115.

Mineault, P. J., Bhaktiari, S., Richards, B. A., & Pack, C. C. (2021). Your head is there to move you around: Goal-driven models of the primate dorsal pathway. *bioRxiv*.

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: two cortical pathways. *Trends in Neurosciences, 6*(0), 414-417. doi:http://dx.doi.org/10.1016/0166-2236(83)90190-X

Rakison, D. H., & Butterworth, G. E. (1998). Infants' attention to object structure in early categorization. *Developmental Psychology, 34*(6), 1310-1325. doi:10.1037/0012-1649.34.6.1310

Regev, T. I., Winawer, J., Gerber, E. M., Knight, R. T., & Deouell, L. Y. (2018). Human posterior parietal cortex responds to visual stimuli as early as peristriate occipital cortex. *European Journal of Neuroscience, 48*(12), 3567-3582.

Rezanejad, M., & Siddiqi, K. (2013). Flux graphs for 2D shape analysis. In *Shape perception in human and computer vision* (pp. 41-54): Springer.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*(3), 382-439.

Schrimpf, M., Kubilius, J., Hong, H., Majaj, N. J., Rajalingham, R., Issa, E. B., . . . DiCarlo, J. J. (2018). Brain-Score: Which Artificial Neural Network for Object Recognition is most Brain-Like? *bioRxiv*. doi:10.1101/407007

Schrimpf, M., Kubilius, J., Lee, M. J., Ratan Murty, N. A., Ajemian, R., & DiCarlo, J. J. (2020). Integrative Benchmarking to Advance Neurally Mechanistic Models of Human Intelligence. *Neuron*. doi:10.1016/j.neuron.2020.07.040

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., . . . Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage, 23*, S208-S219. doi:https://doi.org/10.1016/j.neuroimage.2004.07.051

Snow, J. C., Pettypiece, C. E., McAdam, T. D., McLean, A. D., Stroman, P. W., Goodale, M. A., & Culham, J. C. (2011). Bringing the real world into the fMRI scanner: Repetition effects for pictures versus real objects. *Scientific Reports, 1*(1), 130. doi:10.1038/srep00130

Takemura, H., Rokem, A., Winawer, J., Yeatman, J. D., Wandell, B. A., & Pestilli, F. (2016). A major human white matter pathway between dorsal and ventral visual cortex. *Cerebral Cortex, 26*(5), 2205-2214.

Ungerleider, L. G., & Haxby, J. V. (1994). 'What' and 'where' in the human brain. *Current Opinion in Neurobiology, 4*(2), 157-165. doi:http://dx.doi.org/10.1016/0959-4388(94)90066-3

Van Dromme, I. C., Premereur, E., Verhoef, B.-E., Vanduffel, W., & Janssen, P. (2016). Posterior Parietal Cortex Drives Inferotemporal Activations During Three-Dimensional Object Vision. *PLOS Biology, 14*(4), e1002445. doi:10.1371/journal.pbio.1002445

Vaziri-Pashkam, M., & Xu, Y. (2019). An information-driven 2-pathway characterization of occipitotemporal and posterior parietal visual object representations. *Cerebral Cortex, 29*(5), 2034-2050.

Wang, L., Mruczek, R. E. B., Arcaro, M. J., & Kastner, S. (2014). Probabilistic Maps of Visual Topography in Human Cortex. *Cerebral Cortex, 25*(10), 3911-3931. doi:10.1093/cercor/bhu277

Wang, W., Zhou, T., Zhuo, Y., Chen, L., & Huang, Y. (2020). Subcortical magnocellular visual system facilities object recognition by processing topological property. *bioRxiv*.

Webster, M. J., Bachevalier, J., & Ungerleider, L. G. (1994). Connections of inferior temporal areas TEO and TE with parietal and frontal cortex in macaque monkeys. *Cerebral Cortex, 4*(5), 470-483.

Welchman, A. E. (2016). The human brain in depth: how we see in 3D. *Annual Review of Vision Science, 2*, 345-376.

Yamins, D. L., Hong, H., Cadieu, C. F., Solomon, E. A., Seibert, D., & DiCarlo, J. J. (2014). Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proceedings of the National Academy of Sciences, 111*(23), 8619-8624.

Zachariou, V., Nikas, C. V., Safiullah, Z. N., Gotts, S. J., & Ungerleider, L. G. (2017). Spatial mechanisms within the dorsal visual pathway contribute to the configural processing of faces. *Cerebral Cortex, 27*(8), 4124-4138.

Zhuang, C., Yan, S., Nayebi, A., Schrimpf, M., Frank, M. C., DiCarlo, J. J., & Yamins, D. L. (2021). Unsupervised neural network models of the ventral visual stream. *Proceedings of the National Academy of Sciences, 118*(3).