RESEARCH ARTICLE

Connectivity of the ventral visual cortex is necessary for object recognition in patients

Ye Li[1,2]* | **Yuxing Fang**[1,3]* | **Xiaoying Wang**[1,3] | **Luping Song**[4,5] |
Ruiwang Huang[6] | **Zaizhu Han**[1,3] | **Gaolang Gong**[1,3] | **Yanchao Bi**[1,3]

[1]State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing 100875, China

[2]School of Psychology, Beijing Normal University, Beijing 100875, China

[3]Beijing Key Laboratory of Brain Imaging and Connectomics, Beijing Normal University, Beijing 100875, China

[4]Rehabilitation Medical College of Capital Medical University, Beijing 100068, China

[5]Department of Neurorehabilitation, China Rehabilitation Research Center, Beijing 100068, China

[6]Center for the Study of Applied Psychology, Key Laboratory of Mental Health and Cognitive Science of Guangdong Province, School of Psychology, South China Normal University, Guangzhou 510631, China

Correspondence
Yanchao Bi, State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain

domain-specific distribution for object processing in the VTC is tightly associated with its structural and functional connectivity patterns, especially with those regions beyond VTC that process nonvisual object properties (He et al., 2013; Hutchison, Culham, Everling, Flanagan, & Gallivan, 2014; Mahon & Caramazza, 2011; Osher et al., 2016; Saygin et al., 2012; Wang et al., 2015). Evidence supporting this hypothesis includes data showing that functional selectivity strength for scenes or faces can be predicted by whole-brain white matter (WM) connectivity patterns (Osher et al., 2016; Saygin et al., 2012) and that the tripartite structure of the stimulus-evoked organization in the VTC has distinctive resting-state functional connectivity (rsFC) networks, with routes through both the early visual cortex and longer range connections (Konkle & Caramazza, 2016). These studies showing the association between VTC functionality and its connectivity properties are correlational in nature and do not address whether these connections are necessary for object perception/recognition behavior or whether the VTC itself is sufficient.

Does successful object recognition require the orchestration between the VTC and other regions via intact connections, or do the VTC connections only support downstream object processing and are not themselves required in object recognition? Several studies have focused on the inferior longitudinal fasciculus (ILF)—A major tract that runs through the VTC and links the occipital lobe anteriorly to the anterior temporal lobe and have not obtained consistent results. Thomas et al. (2009) have reported that congenital prosopagnosia patients show significant decreases in fractional anisotropy (FA) of the ILF relative to healthy controls and that the FA values in right ILF significantly correlate with facial recognition scores (Tavor et al., 2014). Gomez et al. (2015), however, did not find such correlation in prosopagnosic individuals, and instead found that in healthy adults face processing performance was correlated with the mean FA value of the WM local to the right fusiform face area (FFA; see also Song et al., 2015 for negative effects of FFA connection bundles in differentiating developmental prosopagnosia from healthy controls). They did observe a significant correlation between place processing performance and the FA values of both local and the whole fiber forward on th0b4.4(n)-.2(7)-.3(h)-21.2(d)a3.9(e)-273.4(p)-.1(a)-wadrw1thele f tfade0-1.9(b)-.1(a)p37.2(w)16.6(c)-6.2(a).7(1)-2.5(b)-TJ0-1.622899D-.0212Tc(ho

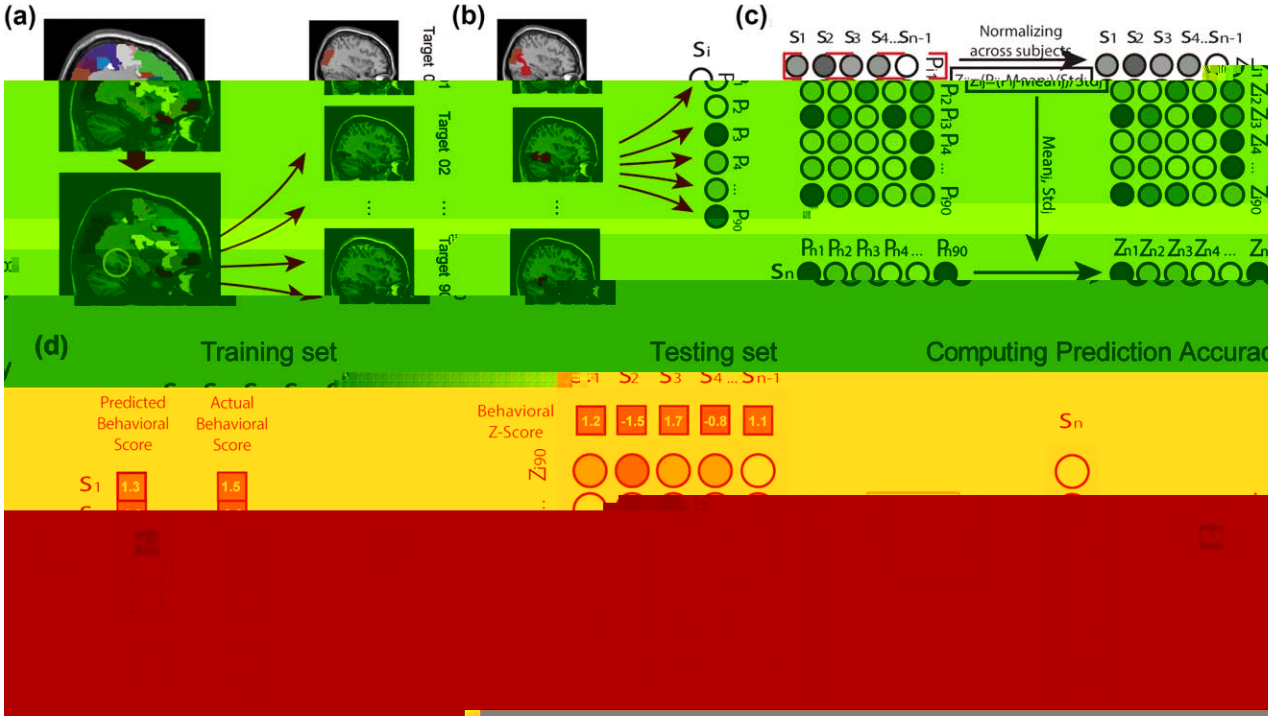

FIGURE 1 Schematic of the processing pipeline. (a) Target region segmentation. The left-top image is an example of the AAL atlas in the native space of a patient. The yellow ring in the left lower image is the edge of the protecting sphere of the animal ROI of this patient; the region inside the protecting sphere has been excluded from the AAL atlas. Then, the AAL atlas was parceled into 90 targets (right row). (b) Fiber tracking. The connection probability of the seed ROI to the other 90 targets was acquired by probabilistic tractography for each patient (left row). Each patient received a vector of probability values (right row). (c) ormalization of the SVR features. For each patient (labeled by i) in the training set, the connection probability values, $_{ij}$, to each target (labeled by j) were Z-normalized across 81 patients (− 1 patients) and transformed into z-scores, $_{ij}$. For the remaining patients in the testing set, the probability values ($_{n1}$ ~ $_{n90}$) were transformed into z-scores (

Participants were asked to verify whether the two parts were from the same object or different objects and press “Yes” or “ o” on the screen accordingly (see Figure 4a for examples). Sixty object-part items from objects from 5 categories (i.e., animals, fruit and vegetables, small nontool objects, tools, and large manmade objects) were taken as target items, with 12 items from each category to form 24 trials (see Table 1 for the complete large manmade object and animal items). One matching trial and one mismatching trial were constructed for each item respectively by presenting it with a part from the same object or with a part from a different object from the same category, and were allocated to the two separate sessions. Thus, the numbers of matching and mismatching trials were identical across categories and between sessions. One point was given if the participant responded correctly to a target item on both the matching and mismatching trials. Responses after a deadline of 60 s were considered as incorrect and no point was given. The accuracies to large manmade objects and to animals were considered in this study.

2.2.2 | Control tasks

Three control tasks were included as various types of contrast to the main object verification task: oral repetition and number calculation that includes general working memory and executive control without object processing; visual form perception that assesses low-level aspects of visual perception. : This task consisted of 8 words (e.g., “咬” Bite) and 4 short sentences (e.g., “老师帮助孩子们做作业” The teacher helps children to do homework). The participants were asked to repeat the word/sentence after the experimenter, one at a time.

y : This task included seven exact calculation questions: 2 addition questions (e.g., “5 + 2”), 2 subtraction questions (e.g., “9 − 4”), 2 multiplication questions (e.g., “2 × 4”), and 1 division question (“6 ÷ 2”). : This task was adapted from a size match task from the Birmingham Object Recognition Battery (BORB Subtest 3; Riddoch & Humphreys, 1993). This task included 30 trials, each containing two dots of different or same sizes. Patients were required to determine whether the two dots were of the same size and to press the corresponding choice on a touch screen (see Figure 4a for an example).

2.3 | Imaging acquisition and preprocessing

Patients were scanned with a 1.5 T GE SIG A EXCITE scanner at the China Rehabilitation Research Center. Three types of images were

TABLE 1 Stimuli in the object form verification task (large man-made objects and animals)

Category	Session 1		Session 2	
Large manmade objects	Station board	Station board	Station board	Blackboard
	Anchor	Anchor	Anchor	Bathtub
	Parachute	Parachute	Parachute	Monument
	Train	Train	Train	Car
	Radiator	Radiator	Radiator	Fence
	Tent	Tent	Tent	Pavilion
	Monument	Refrigerator	Monument	Monument
	Slide	Elevator	Slide	Slide
	Chimney	Station board	Chimney	Chimney
	Castle	Station board	Castle	Castle
	Tank	Bus	Tank	Tank
	Pavilion	Castle	Pavilion	Pavilion
Animals	Crab	Crab	Crab	Spider
	Rooster	Rooster	Rooster	Chimpanzee
	Monkey	Monkey	Monkey	Kangaroo
	Zebra	Zebra	Zebra	Elephant
	Elephant	Elephant	Elephant	Camel
	Panda	Panda	Panda	Lion
	Frog	Mouse	Frog	Frog
	Camel	deer	Camel	Camel
	Cat	Rooster	Cat	Cat
	Lion	Zebra	Lion	Lion
	Penguin	Swan	Penguin	Penguin
	Squirrel	Monkey	Squirrel	Squirrel

collected: (1) high-resolution 3D T1-weighted MPRAGE images; (2) FLAIR T2-weighted images; and (3) diffusion-weighted images. All the images, except for FLAIR T2 images, were scanned twice to improve the quality of the images.

2.3.1 | Structural MRI data

The 3D images were obtained on the sagittal plane with the following parameters: matrix size = 512 × 512, voxel size = 0.49 × 0.49 × 0.70 mm^3, repetition time = 12.26 ms, echo time = 4.2 ms, inversion time = 400 ms, field of view = 250 × 250 mm^2, flip angle = 15°, and slice number = 248. The FLAIR T2 images were obtained on the axial plane with the following parameters: matrix size = 512 × 512, voxel size = 0.49 × 0.49 × 5 mm^3, repetition time = 8,002 ms, echo time = 127.57 ms, inversion = 2 s, field of view = 250 × 250 mm^2, flip angle = 90°, slice number = 28. Two T1 images were first co-registered by a trilinear interpolation method in SPM5 (http://www.fil.ion.ucl.ac.uk/spm/software/spm5) and then averaged. The FLAIR T2 images were co-registered and resliced to the native space of the averaged T1 images. Two trained personnel manually drew the lesion contour of each patient on the averaged T1 images slice by slice, visually referring to the FLAIR T2 images. The lesion-drawing was supervised and confirmed by an experienced radiologist. The structural images of each patient were resliced into 1 × 1 × 1 mm^3 voxel size and then manually registered into Talairach space via the "3D Volume Tools" in BrainVoyager QX v2.0 (www.brainvoyager.com). The manual registration was completed on the 3D visual interface. We extracted the affine transformation between the native and Talairach spaces, which was further applied to transform the lesion masks into the Talairach space using the "WarpImageMultiTransform" program in the ANTs software package (Advanced Normalization Tools, http://www.picsl.upenn.edu/ANTS/). The lesion masks were then transformed into the Montreal Neurological Institute (MNI) space and quality-checked for each patient.

2.3.2 | Diffusion MRI data

Diffusion-weighted imaging had two separate sequences with different diffusion weighting direction sets with 32 directions being covered in total. The parameters of the first acquisition were as follows: 15 diffusion-weighting directions, matrix size = 128 × 128, voxel size = 1.95 × 1.95 × 2.6 mm^3, repetition time = 13,000 ms, echo time = 69.3 ms, inversion time = 0 s, field of view = 250 × 250 mm^2, flip angle = 90°, slice number = 53. The other acquisition had the same parameters except that it included 17 different directions. The first two volumes were b0 volumes, and the b-value of the other volumes was 1,000 s/mm^2 in each sequence. For the diffusion-weighted imaging data of each subject, we first merged the 15 directions and 17 direction-paired sequences into a single 4D image. Then, we preprocessed the images with PANDA (Cui, Zhong, Xu, He, & Gong, 2013), which incorporates FSL (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/). The preprocessing procedure included (a) BET, skull removal; (b) Eddy correct, correction of eddy current distortion; and (c) DTIFIT, building diffusion tensor models and obtaining the FA maps.

2.4 | Seed masks and target regions

The overall analysis scheme for WM pattern construction and behavior prediction is shown in Figure 1. To obtain the connection pattern of functional selective regions in the VTC, we first defined category-selective ROIs—large manmade object selective ROI and animal selective ROI. We defined the ROIs by building sphere ROIs with peak coordinates from a previous fMRI study where 16 healthy subjects viewed pictures of large objects, tools, and animals (He et al., 2013). Group-based ROIs from healthy subjects, instead of functional localizers in the patient groups, were used because of the following reasons: (a) pragmatically, it is difficult to conduct task-based fMRI with patients with brain damage, as asking them to lie still for a long time and follow task instructions are very challenging; (b) theoretically, for our current scientific question, it is more appropriate to import a common template from the healthy population. We do not necessarily expect every patient to have normal category-preferring responses in the VTC regions, because if the category-preference response in VTC is modulated by its connectivity profile (Wang et al., 2015), patients whose connectivity profile is compromised by lesion may have altered functional responses here. Instead, the research rationale is to test whether the VTC regions that normally (i.e., in the healthy group) has categorical preference need to work with other downstream regions to support recognition behavior of the corresponding category. Thus, importing the healthy group results is a more direct way for this purpose. Worth noting is that the activation peak coordinates for the domain of interest tend to be consistent across studies (e.g., the peak Talairach coordinates in the left hemisphere for large artifacts/scenes is around −28, −39, −6; e.g., Downing, Chan, Peelen, Dodds, & Kanwisher, 2006; He et al., 2013). Specifically, we obtained the peak MNI coordinates in bilateral parahippocampal gyrus (PHG) and fusiform gyrus (FG) of clusters selective for large manmade objects and animals with the following contrasts (uncorrected $p < 0.001$, cluster size >10 voxels): (a) large manmade objects > animals & tools: left hemisphere −26, −38, −12 and (2) animals > large manmade objects & tools: right hemisphere 35, −49, −18. Then, we generated spheres with 12 mm radii at these peak coordinates to build the ROIs. We calculated the GM/WM makeup of the seed ROIs by overlapping the seed ROIs with a WM and GM mask constructed by T1 segmentation (SPM 8; default threshold), and found that the two seed ROIs did not significantly differ in the WM percentages (27.3% for animal ROI and 25.4% for large manmade object ROI, $\chi^2 = 0.849$, $p = .357$). Given that the neighboring regions of the ROIs may have extremely high connection probabilities in fiber tracking, which may bias the statistical results, two 16-mm-radius protecting spheres with the same centers as the two ROIs were also generated and cut out from the target regions to build stop masks for tractography. To define the target regions, we adopted the Automated Anatomical Labeling atlas (AAL; Tzourio-Mazoyer et al., 2002) an anatomy-based GM parcellation template commonly used in patient studies (Caeyenberghs & Leemans, 2014; Cao, Shu, Cao, Wang, & He, 2014; Lo et al., 2010). We used only the cerebrum parts of the AAL atlas, which had labels from 1 to 90. Note that because the selection of ROI size is usually arbitrary, we have carried out analyses using a range of different sizes (radius 6, 8, 10, and 12 mm) to conduct the Support Vector Regression (SVR) analysis. The result patterns were largely consistent when the radius is bigger than 8 mm. Smaller ROIs may be too small in relation to the scale of the target ROIs (AAL parcellation), and those with a radius larger than 12 mm were not tested because they are too big and tended to reach out of the cerebrum.

To prepare the masks of seeds and targets, we registered two 12-mm-radius ROIs, two 16-mm-radius protecting spheres and the AAL atlas in the MNI space to the FA images in patients' native spaces, so that we could implement probabilistic tractography for each patient in native space. To conduct the registration, we used the T1 image of each patient as intermediate information. Specifically, we first applied the inverse affine transformation obtained in "structural MRI data" to transform the two ROIs, the two protecting spheres and the AAL atlas into the native space of each patient's T1 image. The aligned ROIs in patients' T1 image in their native space of six sample patients were shown in Figure 3b. For each patient, we registered the FA images in the MNI space obtained in "diffusion MRI data" to the T1 images in the native space by using the "FLIRT linear registration" tool of FSL. By applying the "applywarp" command line in FSL to the images (ROIs, protecting spheres and AAL atlas) registered to the T1 images and the FA images registered to the T1 images, we finally registered the two ROIs, the two protecting spheres and the AAL atlas to the FA images in the native diffusion spaces of all patients. The two ROIs in the native space of each patient were used as seed masks in the tractography. For each ROI, we excluded the protecting sphere from the AAL atlas and then segmented the AAL atlas into 90 target regions (Figure 1a).

2.5 | Probabilistic tractography

Probabilistic tractography is an algorithm that probes the fiber orientation probability distributions at each voxel in an ROI, with the advantages of explicitly representing uncertainty in the data (Behrens et al., 2003) and reconstructing crossing fibers more reliably (Behrens, Berg, Jbabdi, Rushworth, & Woolrich, 2007; Berman, Berger, Chung, Nagarajan, & Henry, 2007). In addition, probabilistic tractography is sensitive to structural changes that contribute to clinical disability (Ciccarelli et al., 2006); therefore, we chose this method rather than deterministic tractography to address our data from brain-damaged patients. We carried out probabilistic tractography with patients' diffusion-weighted imaging data to identify regions with a connection with two functional defined ROIs, by using FMRIB's Diffusion Toolbox (FDT v3.0, http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT) in FSL 5.0.0. For each patient, we applied probabilistic tractography to each of the two ROIs in the patient's native diffusion space because the target regions of each ROI differed. The "Bedpostx" program was used to estimate the diffusion parameters at each voxel of seed ROI. Then, we used "Probtrackx" to conduct fiber tracking with 5000 streamline samples in each voxel of seed ROI to each target region, with the connectivity probability of each voxel to a target region equals to the number of connected streamlines divides 5,000 (Figure 1b). We then averaged the connectivity probabilities of all voxels in the seed ROI to the target region. In this way, the connectivity probabilities to 90 AAL regions, a vector with 90 probability values, of each of the two functional ROIs were acquired for all patients.

2.6 | Statistical analysis: Support vector regression

We built an SVR model to test whether the behavioral performance was predicted by the whole-brain structural connectivity pattern. In this study, performances two recognition tasks (animal form verification and large manmade object form verification) and three control tasks (oral repetition, number exact calculation, and visual form perception) were used as behavioral labels.

For each seed ROI, we adopted a leave-one-subject-out-cross-validation (LOOCV) approach. Specifically, the SVR model was first trained with the connection probabilities of each seed ROI (features) and the behavioral scores from each task (labels) for the 81 (i.e., subjects) as SVR features to avoid including false connections. Note that this step is not "feature selection" in the conventional sense of machine learning as it is not related to the training labels (behavioral responses), and thus does not increase the risk of overfitting. The final number of features being included this way is 26 for the animal-ROI seed and 17 for the large manmade object-ROI seed, a relatively small number compared to the 82 samples. We did not perform any further feature selection step. Before training, the features and the labels of the training group were Z-normalized across 81 patients in the training group. The data from the remaining patients were transformed into Z-scores on the basis of the mean and standard deviation of the data from the 81 patients. Linear classifiers are sensitive to the way features are scaled and the accuracy of an support vector classifier can severely degrade if the data is not normalized (Ben-Hur & Weston, 2010; Chang & Lin, 2011). The advantages of normalizing include avoiding attributes in greater numeric ranges to dominate those in smaller numeric ranges and avoiding numerical difficulties during the calculation, according to parameter selection guide from LIBSVM (https://www.csie.ntu.edu.tw/~cjlin/libsvm/). We used z-normalization (also known as standardization or zero-mean normalization), a method widely used in SVR. It is commonly done within features (i.e., across subjects; Li & Liu, 2011; Stolcke, Kajarekar, & Ferrer, 2008), which makes the feature weights (magnitude of contribution) more interpretable because the features have uniform scaling. A further reason for doing normalization across subjects instead of within patients was because the signals of interest come from variations across patients for connectivity integrity due to brain damage, and the impact of such variations on patients' behavioral performances. Normalization within patient erases such general effects of brain damage across patients. Then a score was predicted by the trained model with the remaining patients' features (Figure 1c). This procedure was done iteratively for all patients, such that we obtained a predicted score for each patient.

We measured the model performance by computing the Pearson correlation coefficient between the predicted scores from the SVR model and the actual behavioral scores (Figure 1d), and estimated the statistical significance using a permutation test. We built models using the same data but with shuffled patient labels, thus uncoupling the relationship between the behavioral labels and actual connectional features. By repeating this process 10,000 times, we generated a null distribution of accuracies. The significance of the model was computed by using the following formula:

$$= \frac{\text{ranking}+1}{10,000+1}$$

where ranking is the number of s in the null distribution of the permutation test that was greater than the actual prediction accuracy from the SVR model. A threshold of $< .05$ was applied.

2.7 | Voxel-wise validation analysis

To avoid any potential bias using the group-level functional-defined ROIs, we conducted a validation voxel-wise analysis without being constrained by functional ROIs and attempt to visualize the potentially significant voxel distributions. In this analysis, probabilistic tractography was done and SVR models were built for each voxel in two anatomically defined regions: bilateral FG and PHG. We excluded patients who have any lesion in these regions, with 72 remaining patients (59 males; mean age = 46, SD = 14; mean number of years of education = 12.99, SD = 3.29; the mean score of MMSE test = 21.22, SD = 8.18). We resliced the bilateral FG and PHG from AAL atlas into 2 × 2 × 2 mm^3 voxels with AFNI (Analysis of Functional NeuroImages, http://afni.nimh.nih.gov/) to correspond to the FA image resolution and registered them to the FA images in the native diffusion spaces of all patients. For each voxel, an SVR model was built predicting behavioral performances with the voxel's WM connectivity patterns using the same method as described above, except for the statistical testing method. Due to the heavy computation burden in the voxel-wise level analysis, we used 1,000 times permutation test to estimate the statistical significance of SVR models. The formula to compute the significance of the model was

$$= \frac{\text{ranking}+1}{1,000+1}$$

and a threshold of $< .05$ was applied.

3 | RESULTS

3.1 | Lesion distribution

The lesion distribution of the 82 patients in this study is shown in Figure 2. All patients' lesions spared the two functional defined ROIs and distributed into the WM structures.

3.2 | Behavioral results

Patients' behavioral performances on object form verification tasks were as follows: object form verification of large manmade objects, 68% ± 18% (mean ± SD); object form verification of animals, 82% ± 15%. For control tasks their performances were: oral repetition, 77% ± 27%; number exact calculation, 74% ± 29%; and visual form perception, 84% ± 12%.

3.3 | White-matter network for large manmade object- and animal-selective regions

Results of probabilistic tractography seeding from the two ROI spheres (large manmade object ROI and animal ROI, Figure 3a) to each of the AAL parcellations for each patient are shown in Figure 3b. The regions having connection probabilities higher than zero were mostly ipsilateral to the two ROIs, covering most of the temporal, parietal, and extending to the frontal lobe (for animal ROI), with a few contralateral connections in the occipital lobe and temporal lobe (Figure 3c). A threshold was applied to avoid spurious connections: connections that were obtained in more than 10% of subjects (8 patients) were considered and their probability values were used as features in the following SVR model (Supporting Information, Table 2).

3.4 | Prediction performance of VTC white-matter pattern on object recognition behavior using SVR: Large manmade objects

To examine whether the integrity of the specific WM pattern of the VTC seed regions was necessary for object recognition behavior, we conducted SVR analysis, using the WM patterns (a vector with probability values to 90 AAL cortical regions) for each patient as features and the patient’

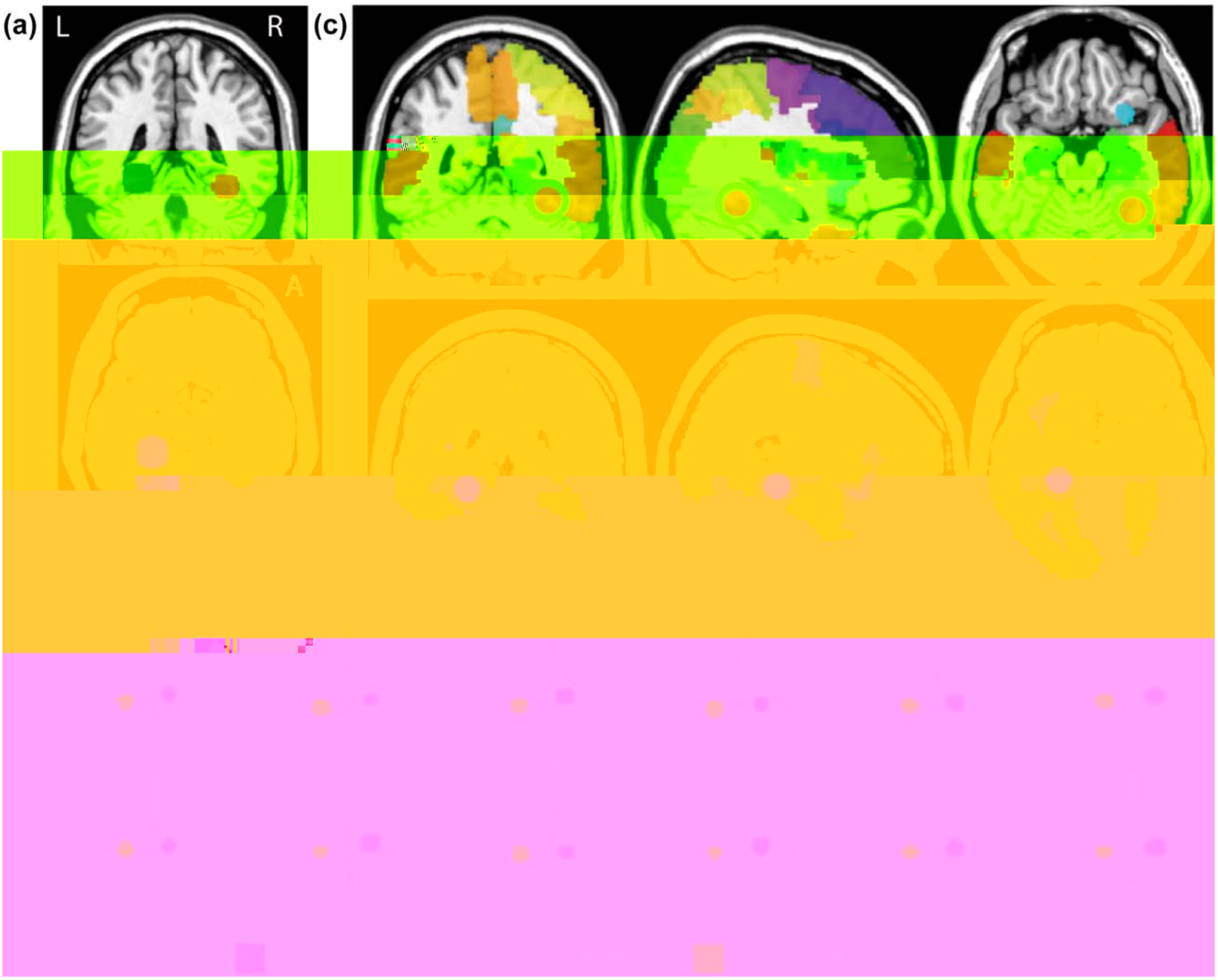

FIGURE 3 Functional defined ROIs and their whole-brain connectivity. (a) Spheres of 12 mm radius centered at the peak voxel of animals > large manmade objects + tools (red, uncorrected < .001, around the right lateral posterior fusiform gyrus), large manmade objects > animals + tools (blue, uncorrected < .001, around the left parahippocampal gyrus and the medial fusiform gyrus). (b) The registration results of two ROIs in the native space of six sample patients. (c) AAL regions that have connection probabilities higher than zero in at least one patient to the animal ROI (top) and to the left large manmade object ROI (bottom). The two ROIs are circled in gray rings. L, left hemisphere; R, right hemisphere

Brain et viewer (Xia, Wang, & He, 2013) (http://www.nitrc.org/projects/bnv) and shown in Figure 5. This finding was not a general effect of cognitive state, as the prediction accuracies of the SVR models for the control tasks were at chance (oral repetition: $= -.021$, $= .887$; number exact calculation: $= .007$, $= .950$; and visual form perception: $= -.014$, $= .891$). Furthermore, this observation was not a general effect of recognition for all types of objects, as the prediction accuracy of the SVR model for animal verification performances was not statistically different from chance ($= .089$, $= .236$). These results indicated that the effects of the WM connection of the large object ROI (left PHG/med FG) were relatively specific for large object recognition.

3.4.1 | Excluding the early visual processing stream

To test whether the selective WM connection effect on large object recognition was driven by the early visual processing streams, we further conducted an SVR model excluding the connections to the early visual cortex. We first excluded the connection probability of the left calcarine sulcus, where the primary visual cortex (V1) is concentrated, and used the remaining features to train the SVR model. The prediction accuracy was significantly higher than baseline ($= .278$, $= .010$). Then, we further excluded the connection probability of the left lingual gyrus, the cuneus, the superior occipital gyrus, and the middle occipital gyrus, and the remaining features still significantly predicted the actual behavioral scores of large object verification ($= .270$, $= .011$).

3.4.2 | The effects of lesion pattern in the cortical regions

To test whether indeed it was the ROI-related WM connectivity pattern that was necessary, not the target cortical regions themselves (e.g., through connections from earlier visual cortex), we used the lesion volume percentages of the nonvisual AAL grey matter regions (excluding the whole occipital lobe) as features to train the

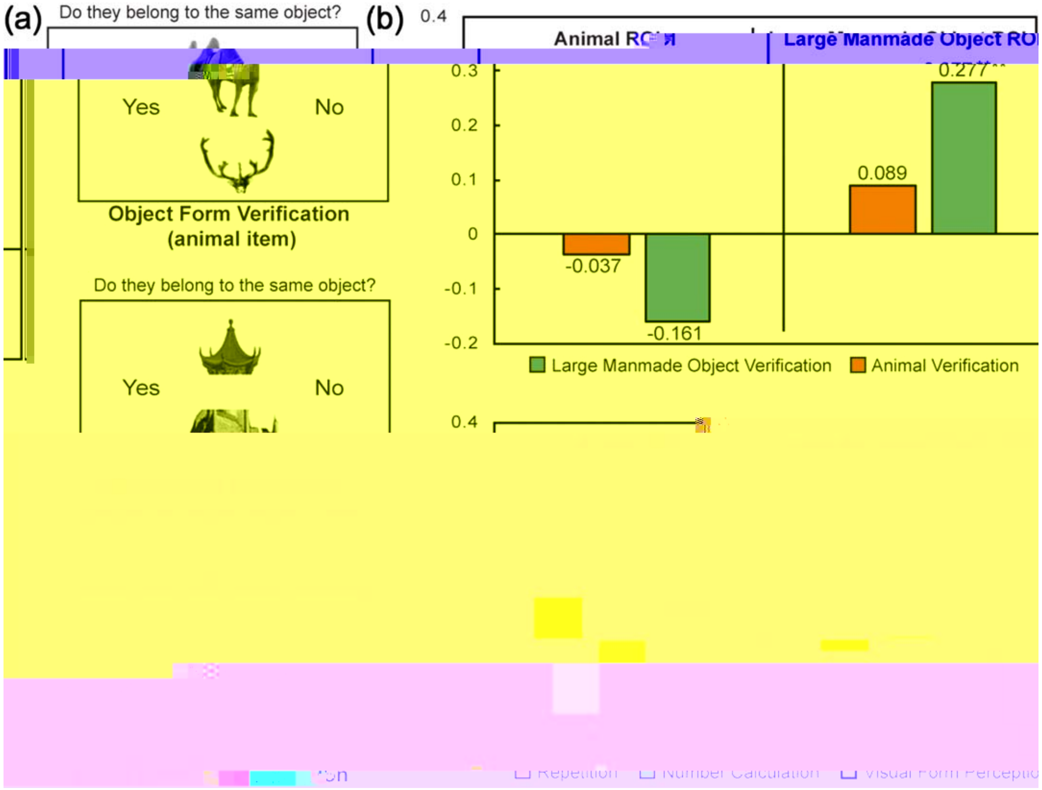

FIGURE 4 Prediction performance of the SVR models. (a) Examples of object form verification (top: an animal item; middle: a large manmade object item) and visual form perception (bottom). (b) Prediction performance of VTC WM connection pattern on object form verification (top) and three control tasks (bottom). The value of each bar refers to the value of correlation between the predicted scores from the SVR model (trained by the connection pattern of the ROI in each header in the LOOCV method) and the actual behavioral scores. "**" refers to a significant correlation at < .01 uncorrected, 10,000-times permutation

SVR model and to predict the behavioral tasks performances. This SVR model was unable to predict any of our behavioral tasks (large manmade object verification: = .049, = .363; animal verification: = −.041, = .653; oral repetition: = .151, = .120; number exact calculation: = −.035, = .636; visual form perception: = .042, = .389).

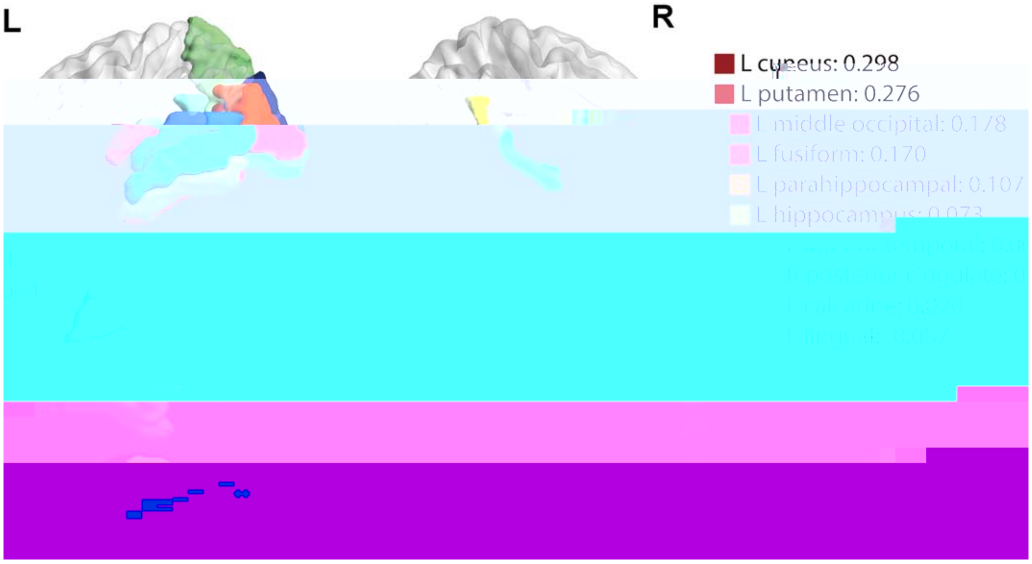

FIGURE 5 Connection weights of the significant SVR model. Weights of target regions are color-coded from dark blue to dark red, reflecting a negative or positive weight coefficient, respectively, and are projected onto surface of a standard brain. The exact weight values are presented in the right column. Transparent regions represent the regions that were not included in the SVR model. L, left hemisphere; R, right hemisphere

3.4.3 | The effects of the VTC ROI local property

While the patients did not have lesion in the local VTC ROI, it might be argued that there are still subtle individual differences here (due to lesion elsewhere or premorbid differences) that contribute to the current results. We further performed a validation analysis, and found that the mean FA value within the large object ROI—extracted by averaging the FA values of all voxels in the 12 mm ROI mask—was not significantly correlated with large object recognition ($= .192$, $= .083$).

3.5 | Prediction performance of VTC white-matter pattern on object recognition behavior using SVR: Animals

The prediction accuracy of the SVR model built from the WM pattern of the animal ROI for animal verification performances was at chance ($_{\text{between predicted value and actual value}} = -.037$, $= .642$). The connection pattern of this ROI was unable to predict the large manmade object verification scores ($= -.161$, $= .920$) or the control tasks scores (oral repetition: $= .079$, $= .519$; number exact calculation: $= -.137$, $= .218$; visual form perception: $= .044$, $= .738$).

3.5.1 | Excluding the early visual processing stream

To test whether the WM connection effect of the animal ROI was masked by any potential effect of early visual processing streams, we also tested a model that excluded the connections with the early visual cortex. We first excluded the connection probability of the right calcarine sulcus and used the remaining features to train the SVR model, and the prediction was still not successful after we excluded the right calcarine sulcus ($= -.021$, $= .570$) or after we further excluded the right lingual gyrus, the right cuneus, the bilateral superior occipital gyrus, the right middle occipital gyrus, and the right inferior occipital gyrus ($= -.107$, $= .833$).

3.5.2 | The effects of the VTC ROI local property

Simialrly, we tested whether the individual differences of the local VTC region will affect the recognition behavior of animal domain. We found that the mean FA value across voxels within the animal ROI mask was not significantly correlated with animal verification scores ($= .112$, $= .315$).

3.6 | Voxel-wise validation analysis: Prediction performance of white-matter pattern of VTC voxels on object recognition

On the voxel level, Figure 6a displays the parahippocampal gyrus and fusiform gyrus voxels whose whole-brain WM connectivity predicted the recognition performance (large manmade objects: $s \geq .181$, $s < .05$ uncorrected; animals: $s \geq .172$, $s < .05$ uncorrected). When plotting the number of significant (uncorrected) voxels (as shown in Figure 6c) and the prediction strength (averaged prediction accuracy of significant voxels, shown in Figure 6b) along the y-axis, a lateral-medial trend could be seen for the effects of animals versus large manmade objects, aligned with the lateral-medial distribution of animate-inanimate distribution in the literature (Chao et al., 1999; Downing et al., 2006; He et al., 2013; Mahon & Caramazza, 2009). When we statistically compared the effects of large object versus animal effects between the medial and lateral (left 1/3 vs. right 1/3 on the y axis, leaving out the middle 1/3 to avoid overlapping effects), nonparametric

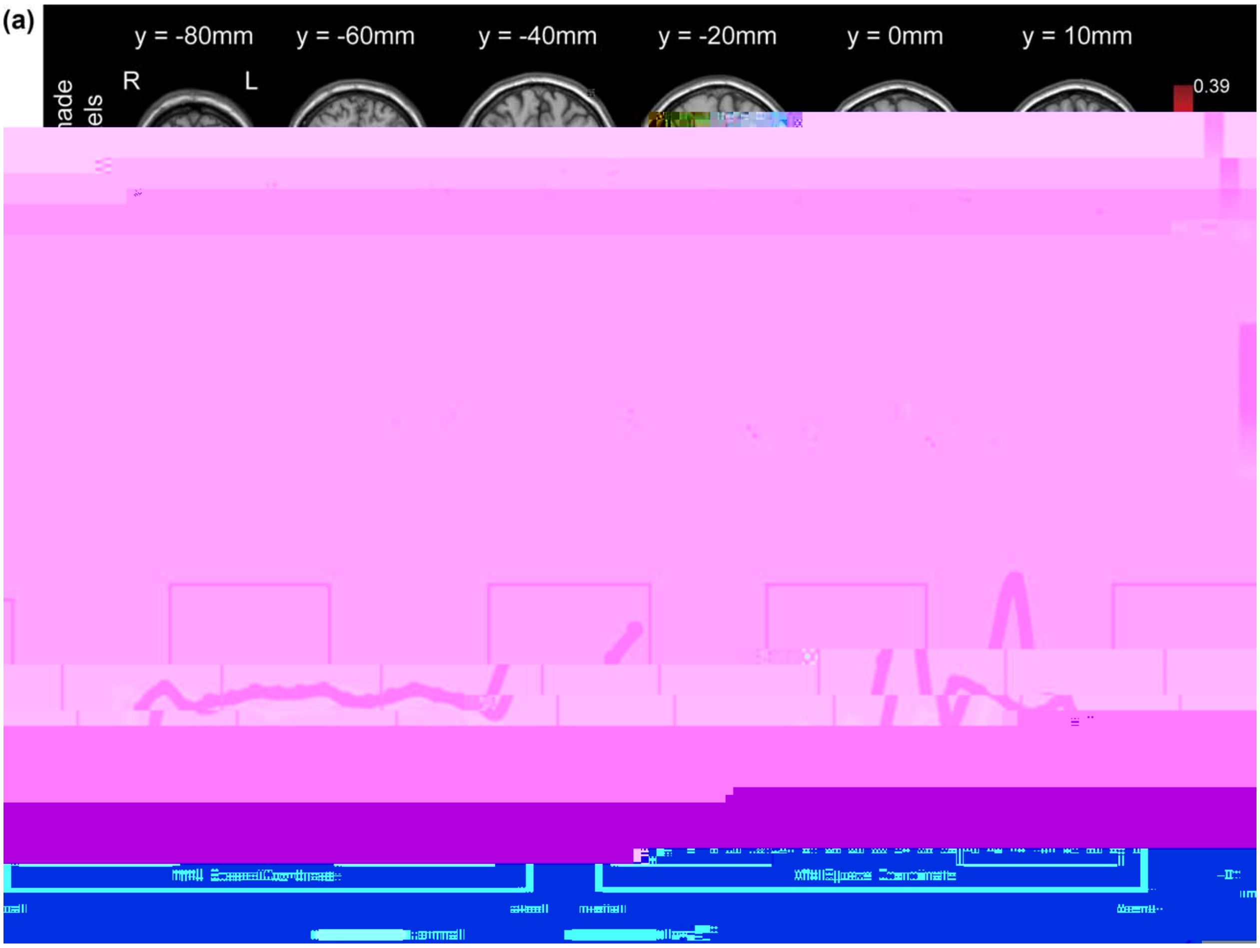

FIGURE 6 Voxels whose whole-brain WM connectivity successfully predicted object recognition in the medial-to-lateral (y axis) direction. (a) Voxels whose whole-brain WM connectivity could predict large manmade object verification (large manmade object-voxels, top) and animal verification (animal-voxels, bottom) performance, respectively. The darker red indicates a higher prediction accuracy of SVR model, according to the color bar in the right. (b) The mean of prediction accuracy of animal-voxels (red line) and large manmade object-voxels (blue line) in the medial-to-lateral direction on the VTC. The VTC here refers to the region comprised of FG and PHG. The voxels in the left hemisphere were flipped to the right hemisphere. The mean prediction accuracy at a given y coordinate was calculated by averaging the SVR prediction accuracy of significant voxels along the anterior–posterior (axis) and superior–inferior (axis) dimensions within the VTC boundary. Gray bars represent the standard deviation (D) of the SVR prediction accuracy of significant voxels on the y axis, with darker grey bars for large manmade object voxels and lighter grey bars for animal voxels. Two orange boxes indicate the medial 1/3 and the lateral 1/3 of the y axis. (c) The number of animal-voxels and large manmade object-voxels in the medial-to-lateral direction on the VTC. The voxels in the left hemisphere were also flipped to the right hemisphere

bottom–up computation in recognition, which is naturally predicted. Importantly, after the removal of the connection probabilities to upstream visual regions from the SVR prediction model, the remaining connections were still capable of predicting large manmade object recognition behavior, thus indicating the necessity of these connections in recognition. Accumulating evidence from neuroimaging and neuropsychological literature suggests that these regions are associated with meaning and knowledge representations (e.g., middle temporal gyrus, Dronkers, Wilkins, Van Valin, Redfern, & Jaeger, 2004; Binder, Desai, Graves, & Conant, 2009; Wei et al., 2012) and specific types of object attributes such as motion (middle temporal gyrus, Hagen et al., 2002; Ricciardi et al., 2011; Van Kemenade et al., 2014; Lingnau & Downing, 2015) or motor features (e.g., putamen, Wymbs, Bassett, Mucha, Porter, & Grafton, 2012; middle and superior temporal gyrus, Grezes & Decety, 2001; Lingnau & Downing, 2015). The involvement of connections between the large object ROI and nonvisual regions highlights the importance of object nonvisual properties in large object recognition, and might relate to interactions with perceivers through mechanisms such as predictive coding (Bar et al., 2006), but the exact functional roles remain unclear. One point worth noting is that fMRI studies have established that large object perception activates a three-region network for scene processing that is commonly assumed to be related to navigation, including the PPA, largely overlapping our large object ROI, along with the retrosplenial complex and the transverse occipital sulcus (Avidan, Levy, Hendler, Zohary, & Malach, 2003; Bar & Aminoff, 2003; Downing et al., 2006; Epstein, Harris, Stanley, & Kanwisher, 1999; He et al., 2013). Additionally, rsFC analyses have shown that the PPA is strongly functionally connected with the bilateral FG, PHG, and

functional neuroimaging studies. *C C (, . : 1991)*, *19*(12), 2767–2796.

Caeyenberghs, K., & Leemans, A. (2014). Hemispheric lateralization of topological organization in structural brain networks. *B* , *35*(9), 4944–4957.

Cao, M., Shu, ., Cao, Q., Wang, Y., & He, Y. (2014). Imaging functional and structural brain connectomics in attention-deficit/hyperactivity disorder. , *50*(3), 1111–1123.

Caramazza, A., & Shelton, J. R. (1998). Domain-specific knowledge systems in the brain: The animate-inanimate distinction. *C* - , *10*(1), 1–34.

Chang, C., & Lin, C. (2011). LIBSVM *. AC* - , *2*(3), 1–39.

Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. , *2*(10), 913–919.

Chao, L. L., & Martin, A. (1999). Cortical regions associated with perceiving, naming, and knowing about colors. *C* - , *11*(1), 25–35.

Ciccarelli, O., Behrens, T. E., Altmann, D. R., Orrell, R. W., Howard, R. S., Johansen-Berg, H., . . . Thompson, A. J. (2006). Probabilistic diffusion tractography: A potential tool to assess the rate of disease progression in amyotrophic lateral sclerosis. *B* , *129*(Pt 7), 1859–1871.

Cui, Z., Zhong, S., Xu, P., He, Y., & Gong, G. (2013). PA DA: A pipeline toolbox for analyzing brain diffusion images. - , *7*, 1–16.

Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., & Damasio, A. R. (1996). A neural basis for lexical retrieval. , *380*(6574), 499–505.

Downing, P. E., Chan, A. W.-Y., Peelen, M. V., Dodds, C. M., & Kanwisher, . (2006). Domain specificity in visual cortex. *C C (, . : 1991)*, *16*(10), 1453–1461.

Dronkers, . F., Wilkins, D. P., Van Valin, R. D., Redfern, B. B., & Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *C* , *92*(1–2), 145–177.

Epstein, R., Harris, A., Stanley, D., & Kanwisher, . (1999). The parahippocampal place area: Recognition, navigation, or encoding? , *23*(1), 115–125.

Fernandino, L., Binder, J. R., Desai, R. H., Pendl, S. L., Humphries, C. J., Gross, W. L., . . . Seidenberg, M. S. (2016). Concept representation reflects multimodal abstraction: A framework for embodied semantics. *C C (, . : 1991)*, *26*(5), 2018–2034.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state": A practical method for grading the cognitive state of patients for the clincian. , *12*(3), 189–198.

Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *B , , & C* , *35*(1), 116–124.

Gainotti, G. (2000). What the locus of brain lesion tells us about the nature of the cognitive defect underlying category-specific disorders: A review. *C* , *36*(4), 539–559.

Gomez, J., Barnett, M. A., atu, V., Mezer, A., Palomero-Gallagher, ., Weiner, K. S., . . . Grill-Spector, K. (2017). Development of face processing expertise. , *355*(6320), 68–71.

Gomez, J., Pestilli, F., Witthoft, ., Golarai, G., Liberman, A., Poltoratski, S., . . . Grill-Spector, K. (2015). Functionally defined white matter reveals segregated pathways in human ventral temporal cortex associated with category-specific processing. , *85*(1), 216–227.

Grezes, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions : A meta-analysis. *B* , *12*(1), 1–19.

Hagen, M. C., Franzén, O., McGlone, F., Essick, G., Dancer, C., & Pardo, J. V. (2002). Tactile motion activates the human middle temporal/V5 (MT/V5) complex. , *16*(5), 957–964.

Han, Z., Ma, Y., Gong, G., He, Y., Caramazza, A., & Bi, Y. (2013). White matter structural connectivity underlying semantic processing: evidence from brain damaged patients. *B* , *136*(10), 2952–2965.

Haynes, J. D. (2015). A primer on pattern-based approaches to fMRI: Principles, pitfalls, and perspectives. , *87*(2), 257–270.

He, C., Peelen, M. V., Han, Z., Lin, ., Caramazza, A., & Bi, Y. (2013). Selectivity for large nonmanipulable objects in scene-selective visual cortex does not require visual experience. , *79*, 1–9.

Hutchison, R. M., Culham, J. C., Everling, S., Flanagan, J. R., & Gallivan, J. P. (2014). Distinct and distributed functional connectivity patterns across cortex reflect the domain-specific constraints of object, face, Phan, T. G., Donnan, G. A., Wright, P. M., & Reutens, D. C. (2005). A digital map of middle cerebral artery infarcts associated with middle cerebral artery trunk and branch occlusion. , *36*(5), 986–991.

Ricciardi, E., Basso, D., Sani, L., Bonino, D., Vecchi, T., Pietrini, P., & Miniussi, C. (2011). Functional inhibition of the human middle temporal cortex affects non-visual motion perception: A repetitive transcranial magnetic stimulation study during tactile speed discrimination. *B* *(* , . .), *236*(2), 138–144.

Riddoch, J. M., & Humphreys, G. W. (1993). *B B: B B* . London: Lawrence Erlbaum Associates.

Saygin, Z. M., Osher, D. E., Koldewyn, K., Reynolds, G., Gabrieli, J. D. E., & Saxe, R. R. (2012). Anatomical connectivity patterns predict face selectivity in the fusiform gyrus. , *15*(2), 321–327.

Song, S., Garrido, L., agy, Z., Mohammadi, S., Steel, A., Driver, J., . . . Furl, . (2015). Local but not long-range microstructural differences of the ventral temporal cortex in developmental prosopagnosia. - , *78*, 195–206.

Stevens, D. W., Tessler, M. H., Peng, C. S., & Martin, A. (2015). Functional connectivity constrains the category-related organization of human ventral occipitotemporal cortex. *B* , *36*(6), 2187–2206.

Stolcke, A., Kajarekar, S., & Ferrer, L. (2008). onparametric feature normalization for SVM-based speaker verification. In: Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on pp 1577–1580.

Tavor, I., Yablonski, M., Mezer, A., Rom, S., Assaf, Y., & Yovel, G. (2014). Separate parts of occipito-temporal white matter fibers are associated with recognition of faces and places. , *86*, 123–130.

Thomas, C., Avidan, G., Humphreys, K., Jung, K., Gao, F., & Behrmann, M. (2009). Reduced structural connectivity in ventral visual cortex in congenital prosopagnosia. , *12*(1), 29–31.

Tzourio-Mazoyer, ., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, ., . . . Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the M I MRI single-subject brain. , *15*(1), 273–289.

Ungerleider, L., & Haxby, J. V. (1994). What" and "where" in the human brain. *C* , *4*(2), 157–165.

Wang, X., Peelen, M. V., Han, Z., He, C., Caramazza, A., & Bi, Y. (2015). How visual is the visual cortex? Comparing connectional and functional fingerprints between congenitally blind and sighted individuals. , *35*(36), 12545–12559.

Wei, T., Liang, X., He, Y., Zang, Y., Han, Z., Caramazza, A., & Bi, Y. (2012). Predicting conceptual processing capacity from spontaneous neuronal activity of the left middle temporal gyrus. - , *32*(2), 481–489.

Wymbs, . F., Bassett, D. S., Mucha, P. J., Porter, M. A., & Grafton, S. T. (2012). Differential recruitment of the sensorimotor putamen and frontoparietal cortex during motor chunking in humans. , *74*(5), 936–946.

Xia, M., Wang, J., & He, Y. (2013). Brain et viewer: A network visualization tool for human brain connectomics. , *8*(7), e68910–e68915.

Zhang, Y., Kimberg, D. Y., Coslett, H. B., Schwartz, M. F., & Wang, Z. (2014). Multivariate lesion-symptom mapping using support vector regression. *B* , *35*(12), 5861–5876.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Li Y, Fang Y, Wang X, et al. Connectivity of the ventral visual cortex is necessary for object recognition in patients. *B* . 2018;39:2786–2799. https://doi.org/10.1002/hbm.24040