

Investigating the Brain Neural Mechanism when Signature Objects were Masked during a Scene Categorization Task using Functional MRI

Qiaomu Miao,^a Gaoyan Zhang,^a Weiran Yan^a and Baolin Liu^{a,b,*}

^a School of Computer Science and Technology, Tianjin Key Laboratory of Cognitive Computing and Application, Tianjin University, Tianjin 300350, PR China

^b State Key Laboratory of Intelligent Technology and Systems, National Laboratory for Information Science and Technology, Tsinghua University, Beijing 100084, PR China

Abstract—Objects play vital roles in scene categorization. Although a number of studies have researched on the neural responses during object and object-based scene recognition, few studies have investigated the neural mechanism underlying object-masked scene categorization. Here, we used functional magnetic resonance imaging (fMRI) to measure the changes in brain activations and functional connectivity (FC) while subjects performed a visual scene-categorization task with different numbers of ‘signature objects’ masked. The object-selective region in the lateral occipital complex (LOC) showed a decrease in activations and changes in FC with the default mode network (DMN), indicating changes in object attention after the masking of signature objects. Changes in top-down modulation effect were revealed in the FC from the dorsolateral prefrontal cortex (DLPFC) to LOC and the extrastriate visual cortex, possibly participating in conscious object recognition. The whole-brain analyses showed the participation of fronto-parietal network (FPN) in scene categorization judgment, and right DLPFC served as the core hub in this network. Another core hub was found in left middle temporal gyrus (MTG) and its connection with middle cingulate cortex (MCC), supramarginal gyrus (SMG) and insula might serve in the processing of motor response and the semantic relations between objects and scenes. Brain-behavior correlation analysis substantiated the contributions of the FC to the different processes in the object-masked scene-categorization tasks. Altogether, the results suggest that masking of objects significantly affected the object attention, cognitive demand, top-down modulation effect, and semantic judgment. © 2018 IBRO. Published by Elsevier Ltd. All rights reserved.

Key words: scene categorization, signature object masked, lateral occipital complex, functional connectivity, functional magnetic resonance imaging.

INTRODUCTION

Humans interact with scenes at all times, during which they receive complex information of the environment around themselves, including numbers of salient objects and scene’s background information. When navigating through different scenes, they usually categorize scenes as different categories subconsciously. Obviously, human’s scene categorization largely depends on the objects presented in scenes (Biederman et al., 1982; Biederman, 1987; Graef et al., 1990), which provide a reliable source of information for scene’s categories, as some objects typically appear in particular scene categories. In addition, behavioral studies have demonstrated that human’s scene categorization performance can be significantly disrupted by removal of ‘signature objects’

or input of incongruent objects (Davenport, 2007; Macevoy and Epstein, 2011). However, besides the information provided by objects, scene categorization may also refer to other aspects of information resources. The spatial properties of scenes, including spatial openness and depth, have also been proved to serve vital roles in scene categorization (Feifei et al., 2007; Greene and Oliva, 2009; McCotter et al., 2005; Oliva and Torralba, 2001). These research suggest the reliance of scene categorization on multiple sources of information.

Besides the behavioral studies, neuroimaging studies have provided more insight into the underlying neural mechanisms of scene categorization. Functional regions of parahippocampal place area (PPA), and retrosplenial cortex (RSC) were typically identified in scene processing (Epstein and Kanwisher, 1998; O’Craven and Kanwisher, 2014; Epstein et al., 2007). These regions were involved in scene recognition and navigation, and showed significantly higher activations for scenes compared to other stimuli (Maguire, 2001; Epstein

*Correspondence to: B. Liu, School of Computer Science and Technology, Tianjin University, Tianjin 300350, PR China. Fax: +86-10-62781789.

E-mail address: liubaolin@tsinghua.edu.cn (B. Liu).

et al., 2014). Multi-voxel pattern analyses (MVPA) have also revealed category-specific neural patterns of activities in response to different scene categories in PPA, RSC, as well as object-selective region of lateral occipital complex (LOC) (Kourtzi and Kanwisher, 2000; Johnson and Johnson, 2014; Grillspector et al., 1998), confirming the neural representations of scene categories in these regions (Walther et al., 2009). LOC has later been found showing similar activity patterns of ‘signature objects’ in scenes and their corresponding scene categories (Macevoy and Epstein, 2011), and was suggested to be explicitly involved in object-based scene categorization (Linsley and MacEvoy, 2014; Wang et al., 2017). On the other hand, the scene-selective PPA and RSC were sensitive to the spatial properties of scenes (Epstein, 2008; Kravitz et al., 2011; Harel et al., 2013; Park et al., 2014). The PPA was shown to have distinct neural patterns to scenes in different spatial expanse, and encode the spatial layout of scene (Epstein and Kanwisher, 1998; Kravitz et al., 2011), while the RSC was typically involved in spatial imagery and navigation (Maguire, 2001; Vass and Epstein, 2016). The above studies indicated the significant roles of spatial properties in the neural representation of scenes, and implied diverse neural pathways underlying scene processing.

Apart from the category-selective regions of interest (ROI), the collaborations between different brain regions were also important in visual cognition. Baldassano et al. have revealed the distinct networks including the scene-selective regions by resting-state functional connectivity (FC) analysis (Baldassano et al., 2015). Other studies have suggested the involvement of high-level regions in visual recognition (Beck et al., 2001; Dehaene and Naccache, 2001). FC between prefrontal regions and visual cortices has been revealed in conscious visual perception (Lumer and Rees, 1999; Imamoglu et al., 2012). Studies have also identified top-down modulation between the prefrontal and early visual cortex in visual object recognition (Bar, 2003; Gazzaley et al., 2005). This top-down modulation might extend to scene recognition based on objects. In addition, objects usually introduce selective attention in visual scene perception (Logan, 1996; Arrington et al., 2000; Greenberg et al., 2015). The selective attention to salient objects might be vital in building the object-scene relations in rapid scene categorization (Wu et al., 2014). Studies have revealed posterior parietal and medial frontal regions in visual object attention (Shomstein and Behrmann, 2006; Arrington et al., 2000; Serences et al., 2004) and interactions between these regions and LOC were observed in visual search (Pantazatos et al., 2012). Objects also carry the semantic information associated with scene categories (Goto et al., 2010; Wang et al., 2017), which might be processed by brain regions associated with semantic processing, such as the middle temporal gyrus and supramarginal gyrus (Binder and Desai, 2011; Mummery et al., 2000; Binder et al., 2009).

Although the above studies have provided much insight into the corresponding brain regions and/or networks in object and scene recognition, few studies

have investigated neural mechanisms in scene categorization when the signature objects are masked. If the signature objects are masked in scene categorization, subjects will lose most of the object-related sources of information and may only refer to other information resources (e.g. spatial properties). Despite the fact that one behavioral study showed that masking of signature objects could significantly disrupt scene categorization performance (Macevoy and Epstein 2011), the underlying neural mechanisms are largely unclear. In addition, there is still a lack of knowledge of the role of FC in scene categorization. In this study, we explicitly placed masks on the ‘signature objects’ in scenes which were strongly associated with the corresponding scene categories. Four categories of visual scenes were presented (kitchen, bathroom, intersection, playground), each containing two ‘signature objects’ (kitchen: refrigerator, microwave oven; bathroom: toilet, bathtub; playground: swing, slide; intersection: car, traffic light). Neural activities were recorded using functional magnetic resonance imaging (fMRI) when subjects performed a scene categorization task with different numbers of signature objects masked. Both the brain regional signal changes and the whole-brain FC changes were analyzed during the object-masked scene categorization tasks. We infer that top-down cognitive control, attention modulations and the semantic information carried by signature objects would change during the object-masked scene categorization tasks. These changes in neural mechanisms might also indicate the role of objects in scene categorization from another angle.

EXPERIMENTAL PROCEDURES

Participants

Fourteen right-handed healthy subjects (age: 18–32, 7 females) with normal or corrected-to-normal vision participated in this study. This study was carried out in accordance with the recommendations of Institutional Review Board (IRB) of Tianjin Key Laboratory of Cognitive Computing and Application, Tianjin University. The protocol was approved by the IRB of Tianjin Key Laboratory of Cognitive Computing and Application, Tianjin University. All subjects gave written informed consent in accordance with the Declaration of Helsinki. All subjects were compensated for their time after the experiment.

Visual stimuli

Visual stimuli were colored images from four categories of common visual scenes: kitchen, bathroom, intersection, and playground selected from the stimuli in (Macevoy and Epstein, 2011). The first two were indoor scenes, and the last two were outdoor scenes. Each scene category contained 2 ‘signature objects’ (kitchen: refrigerator, microwave oven; bathroom: toilet, bathtub; playground: swing, slide; intersection: car, traffic light). Another 20 participants rated the obviousness of the signature objects in each scene image (4 points: both of the two signature objects were obvious; 2 points: only one of the

signature object is obvious; 1 points: both of the signature objects were not obvious). Then these images were sorted according to their average score, and the first 24 images were selected as test stimuli of the original scene images (scores: kitchen: 3.59 ± 0.26 ; bathroom: 3.60 ± 0.25 ; intersection: 3.55 ± 0.20 ; playground: 2.86 ± 0.25). Each image had a resolution of 400×400 pixels. To investigate the neural dynamics in the absence of signature objects within scenes, a new set of images were generated by placing masks on the signature objects in the original images. The masking procedure was conducted similarly as in the behavioral study in (Macevoy and Epstein, 2011). A Fourier transformation and phase random were performed on the original image, and the part of the phase-scrambled image corresponding to the position of signature objects in the original image was picked as the mask, thus preserving the low-level features of the original scene image as much as possible. As a result, we obtained four versions of images for each kind of scene. (NM, images in which no signature object was masked; M1(A): images in which signature object A was masked; M1(B): images in which signature object B was masked; M2, images in which both signature objects were masked). Examples of the different versions of images of each scene category are shown in Fig. 1A. Visual stimuli were back-projected onto a translucent screen viewed by subjects through a mirror fixed on the head coil.

Experimental design

The main fMRI experiment presented images in a blocked design to the subjects. It consisted of 3 functional runs. Each run started with a fixation presented for 8 s and followed by 16 blocks. Blocks were separated by an 8-second interval of baseline. Each block lasted for 32 s, including 8 trials, in which different images but belonged to the same scene category with the same number of signature objects masked were presented. With four categories of scenes (kitchen, bathroom, intersection, and playground) and four kinds of masking conditions (NM, M1(A), M1(B) and M2), we got $4 \times 4 = 16$ versions of images in total (Fig. 1A), and each block in a functional run corresponded to one version of images. The presentation orders of the version of images and image samples in each block were randomized. In each trial of one block, a central fixation was first presented for 500 ms to minimize subjects' eye movements, and then the visual image was presented for 150 ms, followed by a mask showed for 350 ms. The masks were created by first dividing every original image (image with no signature object masked) to 400 equal-sized fragments, each having the size of 20×20 pixels, and randomly select 400 fragments every time from the set of fragments acquired from all the original images to generate a new mask image. Each image appeared only once in the entire experiment. The presentation time of the visual scene was short enough to minimize eye movements, while still long enough for scene categorization (Macevoy and Epstein, 2011). After the presentation of the above, subjects had 3 s to perform a four-alternative forced-choice task to indicate the stimuli's scene category by a button press (Fig. 1B). They were

instructed to respond as quickly and accurately as possible. Each experimental run lasted for 10 min 48 s. Subjects' choices and response time (RT) were recorded using the E-prime software.

An additional localizer run was performed after the main experiment. The design of the localizer run refers to the procedures in (Macevoy and Epstein, 2011). Subjects viewed four kinds of color images: scenes, faces, objects, and phase-scrambled objects in a total of 12 blocks. Images have a resolution of 400×400 pixels. Each block presented different images of a single kind, so each kind of stimuli occupied for 3 blocks. Each block consisted of 32 750-ms images. A central fixation cross was superimposed on all images. Blocks were separated by a 12-second interval of baseline. The whole localizer run lasted for 7 min 24 s.

MRI acquisition

All functional data were acquired using a 3.0 T Philips scanner equipped with an 8-channel head coil at Yantai Affiliated Hospital of Binzhou Medical University. T2*-weighted images were acquired using an echo-planar image (EPI) sequence (TR = 2000 ms, TE = 30 ms, voxel size = $2.5 \times 2.5 \times 4.6$ mm³, matrix size = 64×64 , flip angle (FA) = 90° , slices = 33). Foam pads and earplugs were used to reduce the head motion and scanner noise.

Behavioral data analysis

Since all trials in one block belonged to the same scene category, we only used the RT of the first trial in each block in the later analysis of behavioral data to avoid the prediction effect. Therefore, we have 48 samples for each subject. The behavioral RT of M1(A) and M1(B) were averaged to represent the RT in M1. Behavioral data for 2 subjects were not used due to program failure. SPSS23 (SPSS Inc., Chicago, IL) was used for the statistical analysis of behavioral data in the three scene categorization runs, including correct response rate and response time.

fMRI data preprocessing

Data were preprocessed using SPM8 (<http://www.fil.ion.ucl.ac.uk/spm/software/spm8/>). Four volumes at the beginning of each run were discarded before the following data processing. Functional images were corrected in slice timing and motion corrected with respect to the first volume of each run with a six-parameter rigid body transformation. The images were then registered to the Montreal Neurological Institutes (MNI) space, with a resampled voxel size of $3 \times 3 \times 3$ mm³ (Liang et al., 2017, 2018; Yang et al., 2018). Experimental and localizer data were spatially smoothed using an 8-mm full width half maximum (FWHM) Gaussian kernel.

ROI definition

Functional ROIs were defined bilaterally from the localizer run using statistical contrasts in each individual subject.

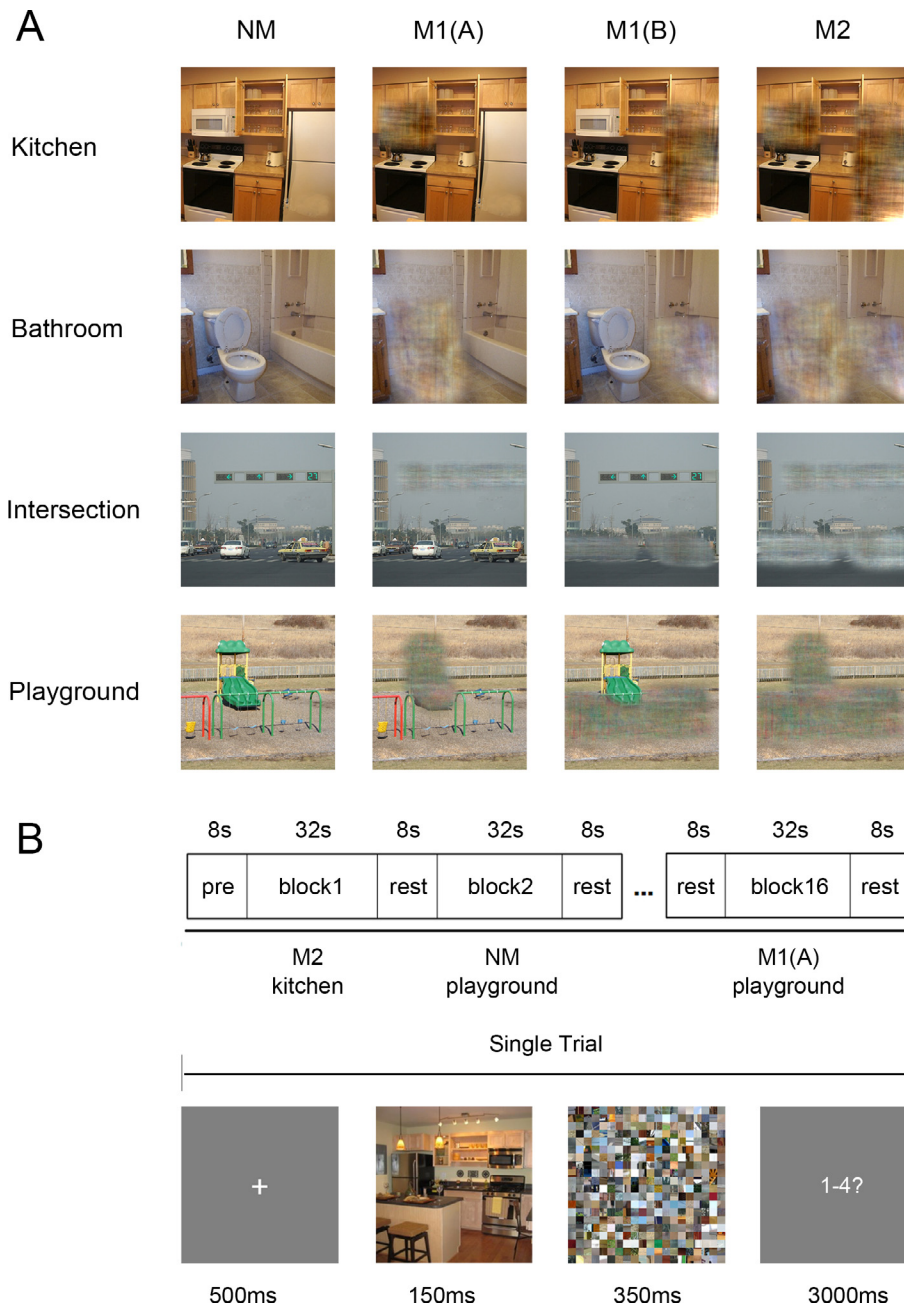


Fig. 1. Examples of experimental stimuli and experimental paradigm. (A) Stimuli belong to four scene categories, each having four conditions of images: NM, images in which no signature object was masked; M1(A): images in which signature object A was masked; M1(B): images in which signature object B was masked; M2, images in which both signature objects were masked. Therefore, we get $4 \times 4 = 16$ versions of images in total. (B) The experimental paradigm of a single run and trial. Each run started with 8 s of fixation and followed by 16 blocks, separated by an 8-second interval of baseline. Each block lasted for 32 s, including 8 trials, in which the same version of images was presented. In a single trial, a central fixation was presented for 500 ms, and then a visual scene image was presented for 150 ms, followed by a mask showed for 350 ms. In the following 3 s, subjects performed a four-alternative forced-choice task to judge the category of the visual scene. The images were selected from the materials in the article “Constructing scenes from objects in human occipitotemporal cortex” published in Nature Neuroscience in 2011, which were permitted by Prof. Sean MacEvoy.

The object-selective region, LOC was defined by selecting the peak voxel of the significant clusters in the lateral occipital cortex from the contrast of objects > phase-scrambled objects, and drawing an 8-mm radius sphere around the peak voxel in each hemisphere.

Scene-selective ROIs were defined similarly, but using the contrast of scenes > faces. The peak voxel of PPA was selected in the posterior parahippocampal-collateral sulcus region, and RSC was selected in the retrosplenial cortex-posterior cingulate-medial parietal region (Fig. 3A).

ROI-based percent signal change analysis

First, to investigate whether the object and scene-selective regions showed changes in activations when the signature objects were masked, percent signal change from the baseline was calculated in the functionally defined LOC, PPA and RSC in different masking degrees for each subject using the marsbar software (<http://marsbar.sourceforge.net/>).

Note that in the data modeling of all the fMRI data analyses in this study, M1(A) and M1(B) conditions were combined to a single condition M1, where only one signature object was masked, to analyze the changes in neural mechanisms with different number of signature objects masked. For simplicity, the different number of signature objects masked was referred to as different ‘masking degrees’ in the analyses below, and there are three masking degrees in total: NM, M1 and M2. The ROIs showing significant changes in neural activations might participate in the object-based scene categorization process, and were used for the following FC analysis.

ROI-based FC analysis

To examine whether the ROIs interact with other regions in the scene categorization process, an ROI-based seed-to-voxel FC analysis was performed using CONN Functional Connectivity Toolbox (version 17.a, www.nitrc.org/projects/conn) (Whitfieldgabrieli and Nietcastanon, 2012). Before analysis, all functional data were filtered using a band-pass filter from

0.008 to 0.09 Hz within the CONN toolbox. Realignment parameters, mean gray matter signal, as well as principal components of white matter and cerebrospinal fluid, were regressed out from the BOLD time series.

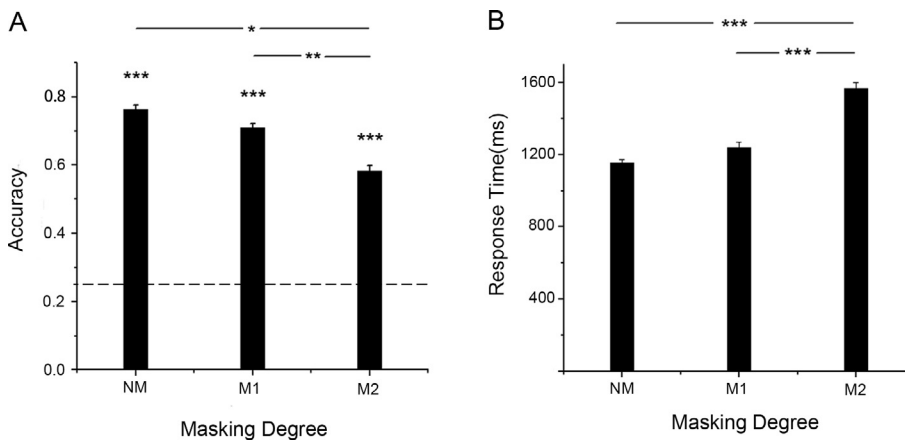


Fig. 2. Behavioral performance averaged across subjects. (A) Average accuracies of scene categorization in all three conditions. (B) Average response time of scene categorization in all three conditions. Dashed line indicates chance level (25%), and error bars indicate standard errors; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

In the FC analysis, the ROIs showing significant changes in activations in the activation analysis were selected as seeds. Mean time series was obtained by averaging the time series of all voxels in the seed region, and Pearson correlation coefficients were calculated between the mean time series of the seed and the time series of every other voxel in the brain. Then the correlation coefficients were Fisher z -transformed, producing a connectivity map for each subject in each masking degree. The first-level connectivity maps were then submitted to a second-level repeated-measure analysis of variance (ANOVA) to identify the brain regions that showed significant changes in FC with the seed region in different masking degrees. Significant results were obtained with a combination of voxel-level uncorrected $p < 0.001$ and a cluster extent family-wise error (FWE)-corrected $p < 0.05$. After the significant clusters were obtained, FC between the seed and each of its resulting clusters was calculated in all 3 masking degrees by averaging the z values of all voxels in the significant clusters for each subject. Paired t -tests were performed on the FC in every two combinations of masking degrees using SPSS23 to examine specifically how the FC changes as more signature objects were masked, corrected at a false discovery rate- (FDR) of $p < 0.05$ for multiple comparisons.

Whole-brain activation analysis

To investigate the neural changes in other cognitive regions during object-masked scene categorization, we performed additional activation and FC analyses in the whole brain.

Based on the preprocessed fMRI data, parameter estimates were calculated using a general linear model (GLM) implemented in SPM8, where different masking degrees were modeled using a boxcar function and convolved with a canonical hemodynamic response

function and the six movement parameters were included as regressors. After the parameter maps were obtained for each subject in each masking degree, a group-level within-subject repeated-measure ANOVA was performed on the parameter maps to identify the brain regions showing significant changes in neural activations in different masking degrees. Statistical maps were obtained after using a threshold of voxel-level FDR-corrected $p < 0.05$, with a cluster extent threshold $k = 20$ voxels. To examine the change in activations specifically, the average percent signal change of each cluster in different masking degrees for each subject was calculated and paired t -tests were performed in every two combinations of masking degrees. The reported contrasts survived an FDR-corrected $p < 0.05$ for multiple comparisons.

Whole-brain FC analysis

As in the ROI-based FC analysis, the whole-brain FC analyses were performed using the CONN toolbox and functional images were preprocessed in the same manner before analysis. In the whole-brain FC analysis, we first searched for the regions which served as hubs in the functional networks significantly affected by the masking of signature objects in scene categorization. To seek for these hubs, we focused on the brain regions showing significant changes in FC to the whole brain in different masking degrees by performing a voxel-to-voxel FC analysis using Intrinsic Connectivity Contrast (ICC) (Martuzzi et al., 2011). This method measures the average connectivity to the whole brain for every voxel by calculating the root mean square of the correlation coefficient value between the time series of the seed voxel and all other voxels in the brain, thus producing a connectivity map for each subject in each masking degree. The values in connectivity maps were converted to z -values using a Fisher z -transformation. First-level connectivity maps were then submitted to second-level repeated-measure ANOVA to identify the brain regions showing significant changes in FC to the whole brain in different masking degrees. Clusters surviving a combination of voxel-level uncorrected $p < 0.001$ and a cluster extent FWE-corrected $p < 0.05$ were identified as core hubs. This voxel-to-voxel analysis overcame the shortcomings of selecting ROIs subjectively (Martuzzi et al., 2011; Zhang et al., 2016), and provided a direct way to locate the core hubs in the functional networks possibly involved in object-masked scene categorization.

The core hubs obtained in the voxel-to-voxel analysis were then used as seeds in another seed-to-voxel analysis to find out which specific brain regions showed changes in FC with them. The analysis procedure was

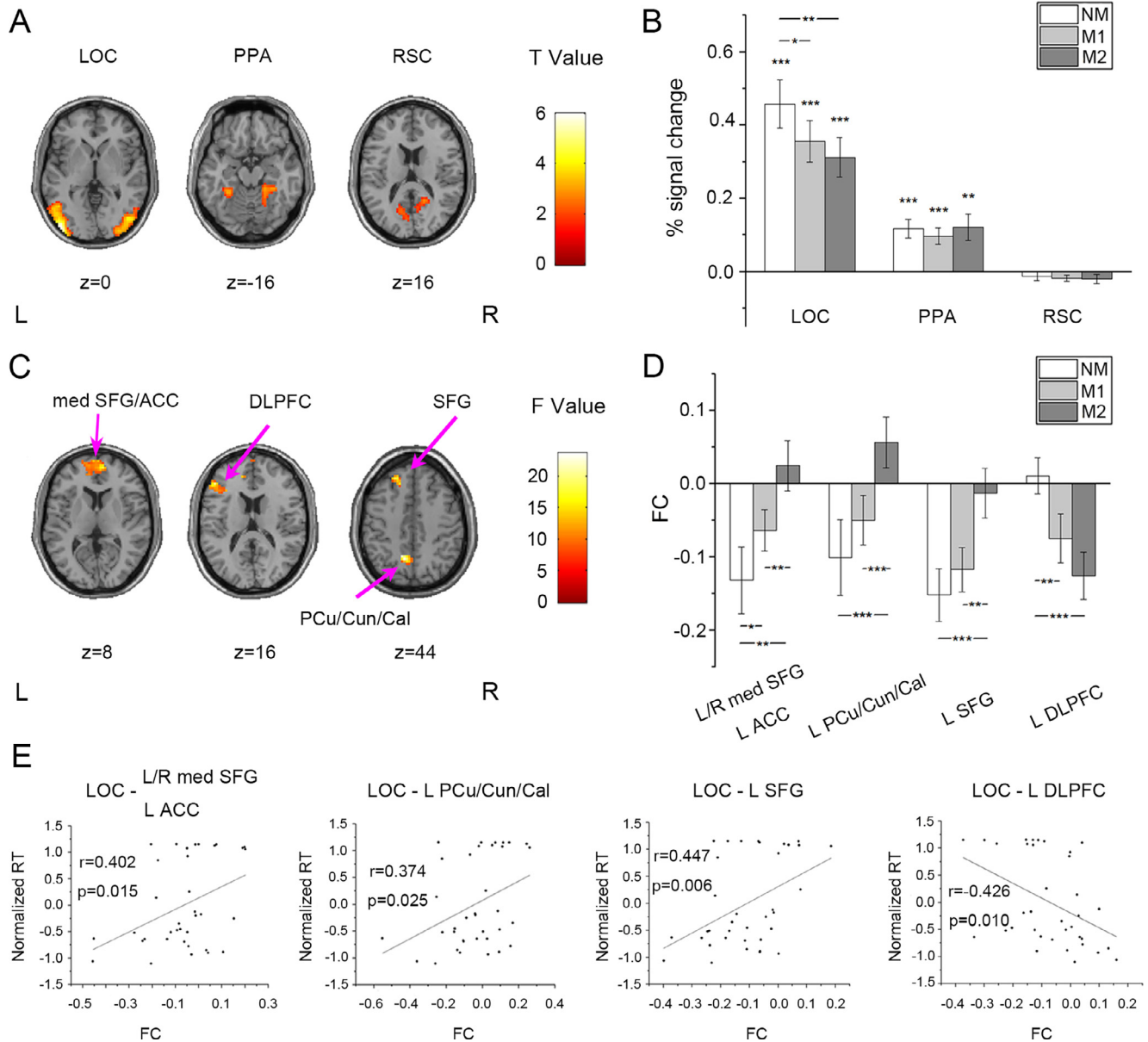


Fig. 3. ROI analysis results. (A) Statistical maps of the significant clusters of LOC, PPA, RSC of a representative subject (uncorrected $p < 0.05$ with a cluster size > 20 voxels). Peak voxel was selected in each cluster and each ROI was created by drawing an 8 mm sphere around that voxel. (B) Percent signal change of LOC, PPA, RSC averaged across all subjects in each masking degree. (C) Group-level statistical maps of clusters showing significant changes in FC in different masking degrees (NM, M1, and M2) with LOC in the seed-to-voxel analysis, thresholded at a combination of voxel-level $p < 0.001$ uncorrected and a cluster extent FWE-corrected $p < 0.05$. (D) FC between LOC and each significant cluster averaged across all subjects. (E) Correlations between FC and normalized behavioral RT. Significant correlations with behavioral RT were observed in all FC. SFG Superior Frontal Gyrus, ACC, Anterior Cingulate Cortex; PCu, precuneus; Cun, cuneus; Cal, calcarine; DLPFC, dorsolateral prefrontal cortex; med, medial; L, left; R, right; Error bars indicate standard errors; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

identical to the ROI-based seed-to-voxel analysis. Significant results were obtained with the same threshold method (a combination of voxel-level $p < 0.001$ uncorrected and a cluster extent FWE-corrected $p < 0.05$). As in the ROI-based seed-to-voxel analysis, FC between the core hub and each of its resulting clusters was calculated in all 3 masking degrees for all subjects, and paired t-tests were performed on the FC in every two combinations of masking degrees (FDR-corrected $p < 0.05$ for multiple comparisons).

Brain-behavior correlation analysis

Correlation analysis of the regional signal changes and FC with behavioral responses was conducted to investigate whether these neural changes contributed to scene categorization behaviors.

For all subjects together, Pearson correlation coefficients were calculated between the behavioral RT and percent signal change in each significant changed ROI and cluster in different masking degrees. To eliminate individual difference of general RT, each

subject's RT was z-score normalized with respect to his/her mean RT in all masking degrees. Correlation coefficients were statistically tested from 0 to reveal the contribution of different regions in object-masked scene categorization behaviors. For the ROI-based and voxel-based FC analysis, similar correlation analysis was conducted.

RESULTS

Behavioral performance

In all three masking degrees, the categorization accuracies were significantly higher than the chance level of 25% (NM: $76.4 \pm 13.21\%$, $t(11) = 13.47$, $p < 0.001$; M1: $70.8 \pm 16.38\%$, $t(11) = 9.693$, $p < 0.001$; M2: $58.3 \pm 19.14\%$, $t(11) = 6.034$, $p < 0.001$). Paired t-tests were performed on the accuracies and RTs in every two combinations of the three masking degrees. The accuracies of NM and M1 were significantly higher than that of M2 (NM vs. M2: $t(11) = 3.120$, $p = 0.01$; M1 vs. M2: $t(11) = 3.717$, $p = 0.003$) (Fig. 2A), and the RT of M2 (1566 ± 402 ms) was significantly longer than that of NM (1153 ± 246 ms; $t(11) = 5.828$, $p < 0.001$) and M1 (1240 ± 342 ms; $t(11) = 5.800$, $p < 0.001$) (Fig. 2B). The behavioral results suggest severely degraded performance of scene categorization when the signature objects were masked in scenes.

ROI-based percent signal change results

Percent signal change was calculated in the predefined LOC, PPA, and RSC in all masking degrees for each subject. Significant activations compared to baseline were found in all masking degrees in LOC (NM: $t(13) = 6.874$, $p < 0.001$; M1: $t(13) = 6.366$, $p < 0.001$; M2: $t(13) = 5.754$, $p < 0.001$) and PPA (NM: $t(13) = 4.573$, $p < 0.001$; M1: $t(13) = 4.275$, $p < 0.001$; M2: $t(13) = 3.340$, $p = 0.005$) in the scene categorization task. Deactivations were observed in RSC, but not significant. Paired t-tests revealed that LOC showed significant higher activations in NM compared to M2 ($t(13) = 3.661$, $p = 0.003$), which survived an FDR-corrected for multiple comparisons of $p < 0.05$, and M1 ($t(13) = 2.824$, $p = 0.014$), of which the FDR-corrected p-value was slightly higher than 0.05 (Fig. 3B). No significant difference in activations was observed in PPA and RSC in different masking degrees.

ROI-based FC results

Due to that only LOC in the predefined ROIs showed a significant change in activations in different masking degrees, LOC was defined as the seed region and a seed-to-voxel FC analysis was performed to study whether it interacted with other regions in object-masked scene categorization process. Significant clusters were observed in bilateral medial superior frontal gyrus (SFG) and left anterior cingulate cortex (ACC), left precuneus/cuneus/calcarine, left SFG, and left DLPFC (Fig. 3C). All clusters survived a combination of voxel-level uncorrected $p < 0.001$ and a cluster extent FWE-corrected $p < 0.05$. Anatomical

details are shown in Table 1. To further determine the location of these clusters in intrinsic functional networks, we employed a previously published brain template (Yeo et al., 2011), and found that all of the clusters except left DLPFC belonged to the default mode network (DMN). FC between LOC and these clusters were further calculated in each masking degree for each subject. Since conversions from negative FC to weak positive FC and from weak positive FC to negative FC were observed in these clusters, one-sample t-tests were performed on the positive FC to confirm their significance (M2 in bilateral medial SFG & left ACC and left SFG, NM in left DLPFC). None of these FC was significantly different from 0 (all $p > 0.05$). Therefore, the trend of changes in FC could all be interpreted as decrease or increase in negative FC between LOC and these clusters. Paired t-tests revealed a significant decrease in negative FC between LOC and bilateral medial SFG & left ACC in M2 compared to NM ($t(13) = 4.067$, $p = 0.001$) and M1 ($t(13) = 3.260$, $p = 0.006$), and in M1 compared to NM ($t(13) = 2.418$, $p = 0.031$). Significant decrease in negative FC in M2 compared to NM and M1 was also observed in left precuneus/cuneus/calcarine (M2 vs. NM: $t(13) = 5.092$, $p < 0.001$; M2 vs. M1: $t(13) = 4.312$, $p < 0.001$) and left SFG (M2 vs. NM: $t(13) = 4.782$, $p < 0.001$; M2 vs. M1: $t(13) = 3.845$, $p = 0.002$). On the contrary, significant increase in negative FC was observed in M1 and M2 compared to NM in left DLPFC (M1 vs. NM: $t(13) = -3.310$, $p = 0.006$; M2 vs. NM: $t(13) = -5.329$, $p < 0.001$). All contrasts survived an FDR-corrected for multiple comparisons of $p < 0.05$. FC averaged across all subjects was calculated and visualized in Fig. 3D.

Whole-brain activation analysis results

The group-level repeated-measure ANOVA on the activation maps revealed that activations in the bilateral medial orbitofrontal cortex (OFC), right dorsolateral prefrontal cortex (DLPFC), left precuneus and left inferior parietal lobule (IPL) showed a significant main effect of masking degree (Fig. 4A). All clusters survived a threshold of voxel-level FDR-corrected $p < 0.05$, with a cluster threshold of 20 voxels. Anatomical details of these clusters are shown in Table 2. A cluster was found in the left lateral occipital cortex with a more relaxed threshold of voxel-level uncorrected $p < 0.001$, but not survived an FDR-corrected $p < 0.05$.

Paired t-tests were performed on the percent signal change in each significant cluster in every two combinations of masking degrees. Significant higher activations was observed in M2 compared to M1 ($t(13) = 6.774$, $p < 0.001$) and NM ($t(13) = 3.378$, $p = 0.005$) in right DLPFC, and in NM and M2 compared to M1 in left IPL (NM vs. M1: $t(13) = 2.655$, $p = 0.02$; M2 vs. M1: $t(13) = 5.305$, $p < 0.001$). On the other hand, deactivations were observed in bilateral medial OFC and left precuneus. Significant decreased activations were observed in M2 compared to NM ($t(13) = -5.166$, $p < 0.001$) and M1 ($t(13) = -3.708$, $p = 0.003$) in bilateral medial OFC, and in M1 and M2 compared to NM in left precuneus (M1 vs. NM: $t(13) = -3.951$, $p = 0.002$; M2 vs. NM: $t(13) = -4.702$, $p < 0.001$). All

Table 1. Clusters showing significant changes in FC with LOC in different masking degrees

Brain Regions	Number of Voxels	Brodmann's Area	Peak MNI-coordinates			Peak F-score
			x	y	z	
L/R med SFG	171	10/32	9	56	7	17.43
L ACC						
L PCu/Cun/Cal	87	7/31	-9	-58	43	21.94
L SFG	68	8/9	-18	35	46	23.66
L DLPFC	68	46	-48	35	16	19.7

Note: SFG Superior Frontal Gyrus, ACC, Anterior Cingulate Cortex; PCu, precuneus; Cun, cuneus; Cal, calcarine; DLPFC, dorsolateral prefrontal cortex; med, medial; L, left; R, right.

contrasts survived an FDR correction for multiple comparisons at $p < 0.05$. Percent signal change averaged across all subjects in all 3 masking degrees was calculated and visualized in Fig. 4B.

Whole-brain FC analysis results

Voxel-to-voxel analysis results. To locate the core hubs in the brain networks affected by the masking of signature objects in scene categorization, we first performed a voxel-to-voxel analysis of FC using ICC. One-way repeated measure ANOVA of the ICC connectivity map revealed that right dorsolateral prefrontal cortex (DLPFC) and left medial temporal gyrus (MTG) showed a significant main effect of masking degree in their FC to the whole brain (Fig. 5A). Anatomical details of these two core hubs are shown in Table 3(a).

Seed-to-voxel analysis results. After obtaining the two “core hubs” in the voxel-to-voxel analysis, we further performed another seed-to-voxel FC analysis using these two core hubs as the seed regions. When Right DLPFC was selected as the seed region, significant clusters were found in bilateral cuneus/calcarine, medial SFG, left triangular part of inferior frontal gyrus (IFG), inferior temporal gyrus (ITG), inferior parietal lobule (IPL), middle frontal gyrus (MFG) and opercular part of inferior frontal gyrus (IFG) (Fig. 5B). By employing the previous brain template (Yeo et al., 2011), we confirmed that these regions largely overlapped with the fronto-parietal network (FPN). Anatomical details of these significant clusters are shown in Table 3(b). Average FC between right DLPFC and each significant cluster in all 3 masking degrees was calculated for each subject and paired t-tests were performed in every two masking conditions. Results reveal significant increase in negative FC in bilateral cuneus/calcarine, with significant higher negative FC in M2 than NM ($t(13) = -7.907$, $p < 0.001$) and M1($t(13) = -7.770$, $p < 0.001$). Significant higher FC was observed in M1 and M2 compared to NM in bilateral medial SFG (M1 vs. NM: $t(13) = 3.723$, $p = 0.003$; M2 vs. NM $t(13) = 4.687$, $p < 0.001$), left triangular part of IFG (M1 vs. NM: $t(13) = 4.131$, $p = 0.001$; M2 vs. NM: $t(13) = 3.893$, $p = 0.002$), left ITG (M1 vs. NM: $t(13) = 3.868$, $p = 0.002$; M2 vs. NM: $t(13) = 4.215$, $p = 0.001$), left IPL (M1 vs. NM: $t(13) = 4.464$, $p < 0.001$; M2 vs. NM: $t(13) = 3.965$, $p = 0.002$), left MFG/opercular part of IFG (M1 vs. NM: $t(13) = 4.333$, $p < 0.001$; M2 vs. NM: $t(13) = 3.569$, $p = 0.003$). All contrasts survived

an FDR corrected for multiple comparisons of $p < 0.05$. FC averaged across all subjects is shown in Fig. 5D.

When left MTG was selected as the seed region, three clusters were found showing significant changes in FC, including right middle cingulate cortex (MCC) and left supplementary motor area (SMA)/paracentral lobule (PCL), right supramarginal gyrus (SMG)/superior temporal gyrus (STG), and right insula/putamen (Fig. 5C). Anatomical details of significant clusters are shown in Table 3(c). As the procedures for right DLPFC, the average FC in each masking condition for each subject was calculated and paired t-test was performed in every two masking conditions. Results reveal significant decrease in negative FC in M1 and M2 compared to NM in all three clusters: left SMA/PCL (M1 vs. NM: $t(13) = 7.77$, $p < 0.001$, M2 vs. NM: $t(13) = 7.907$, $p < 0.001$), right SMG/STG (M1 vs. NM: $t(13) = 4.214$, $p = 0.001$; M2 vs. NM: $t(13) = 5.571$, $p < 0.001$;) and right insula/putamen (M1 vs. NM: $t(13) = 5.742$, $p < 0.001$; M2 vs. NM: $t(13) = 5.421$, $p < 0.001$), except for right SMG/STG, which also reveals significant lower negative FC in M2 compared to M1 ($t(13) = 3.728$, $p = 0.003$). All contrasts survived an FDR corrected for multiple comparisons of $p < 0.05$. Average FC between left MTG and each significant cluster is shown in Fig. 5E.

Brain-behavioral correlation results. First, the percent signal change in the predefined ROIs-LOC, PPA, and RSC was correlated with subjects' normalized behavioral RT in different masking degrees. No significant correlation was found in any of these ROIs. However, significant correlations with behavioral RT were found in all FC between LOC and its significant clusters in the seed-to-voxel analysis, including bilateral medial SFG & left ACC ($r = 0.402$, $p = 0.015$), left precuneus/cuneus/calcarine ($r = 0.374$, $p = 0.025$), left SFG ($r = 0.447$, $p = 0.006$) and left DLPFC ($r = -0.426$, $p = 0.010$) (Fig. 3E).

In the whole-brain analysis, percent signal change in the significant clusters in the whole-brain activation analysis was correlated with the subjects' behavioral RT in the scene categorization task. Significant correlations were found in right DLPFC ($r = 0.665$, $p < 0.001$) and bilateral medial OFC ($r = -0.43$, $p = 0.009$) (Fig. 4C). Marginal correlation with behavioral RT was found in left IPL ($r = 0.289$, $p = 0.087$). In the whole-brain FC analysis, significant correlations with behavioral RT were observed in FC between right DLPFC and bilateral cuneus/calcarine ($r(12) = -0.3888$, $p = 0.019$),

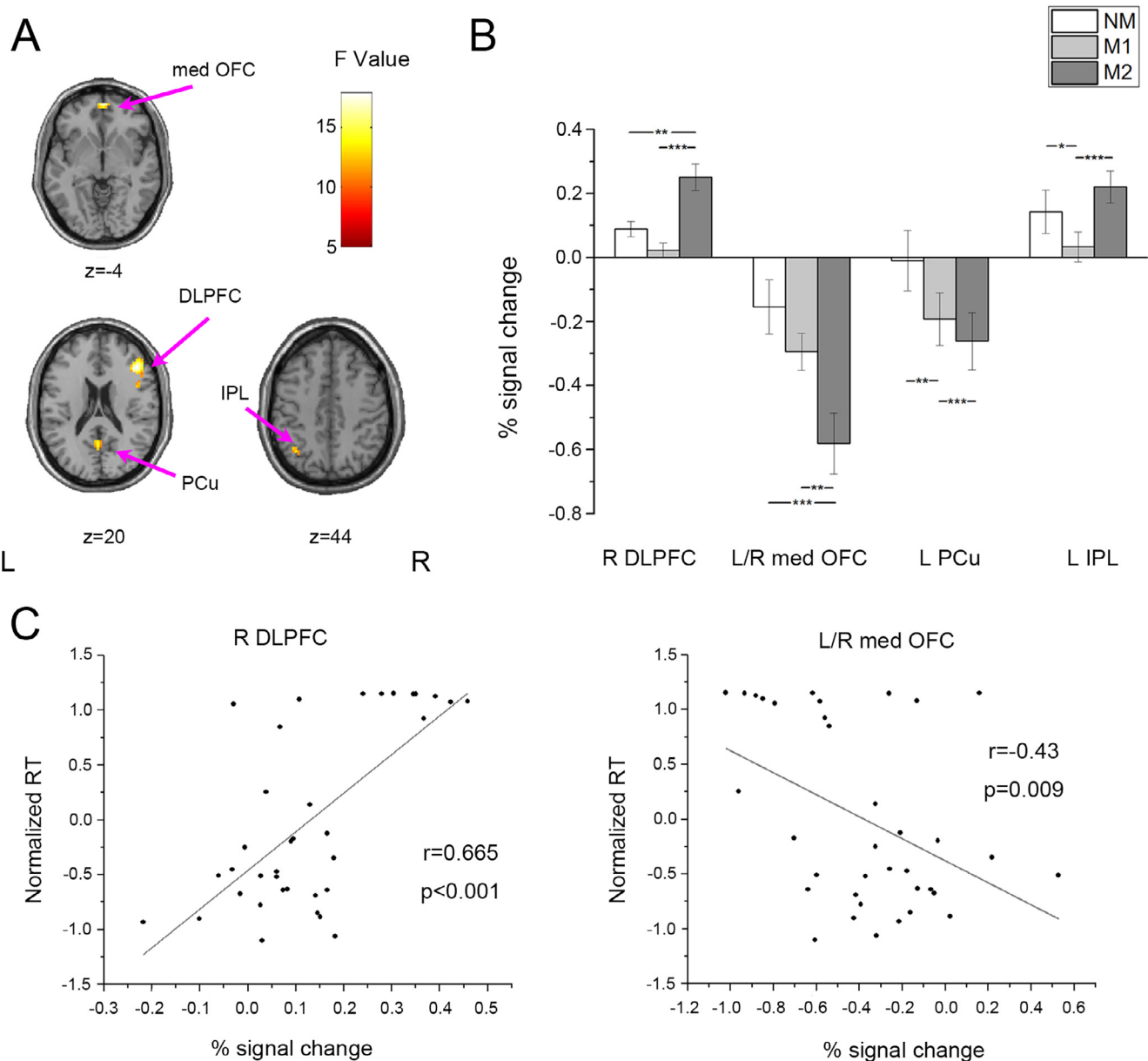


Fig. 4. Significant clusters in the whole-brain activation analysis. (A) Group-level statistical maps of clusters showing significant changes in activations in different masking degrees (NM, M1, and M2), thresholded at $p < 0.05$ FDR-corrected with a cluster extent threshold $k = 20$ voxels. (B) Percent signal change calculated in each significant cluster averaged across all subjects. (C) Correlations between the percent signal change in each cluster and normalized behavioral RT. DLPFC, dorsolateral prefrontal cortex; OFC, orbitofrontal cortex; PCu, precuneus; IPL, inferior frontal lobule; med, medial; L, left; R, right; Error bars indicate standard errors; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 2. Clusters showing significant changes in activations in different masking degrees

Brain Regions	Number of Voxels	Brodmann's Area	MNI coordinates			Peak F-score
			x	y	z	
R DLPFC	175	46/9/45	45	29	19	17.92
L/R med OFC	36	10	3	59	-5	17.19
L PCu	20	31/23	0	-58	19	14.17
L IPL	22	7/40	-36	-61	49	13.41

DLPFC, dorsolateral prefrontal cortex; OFC, orbitofrontal cortex; PCu, precuneus; IPL, inferior frontal lobule; med, medial; R, right; L, left.

between left MTG and right MCC & left SMA/PCL $r(12) = 0.450$, $p = 0.006$) and between left MTG and right SMG/STG ($r(12) = 0.429$, $p = 0.009$) (Fig. 5F).

DISCUSSION

In the present study, the predefined ROI-based analysis showed that the object-selective region in LOC was

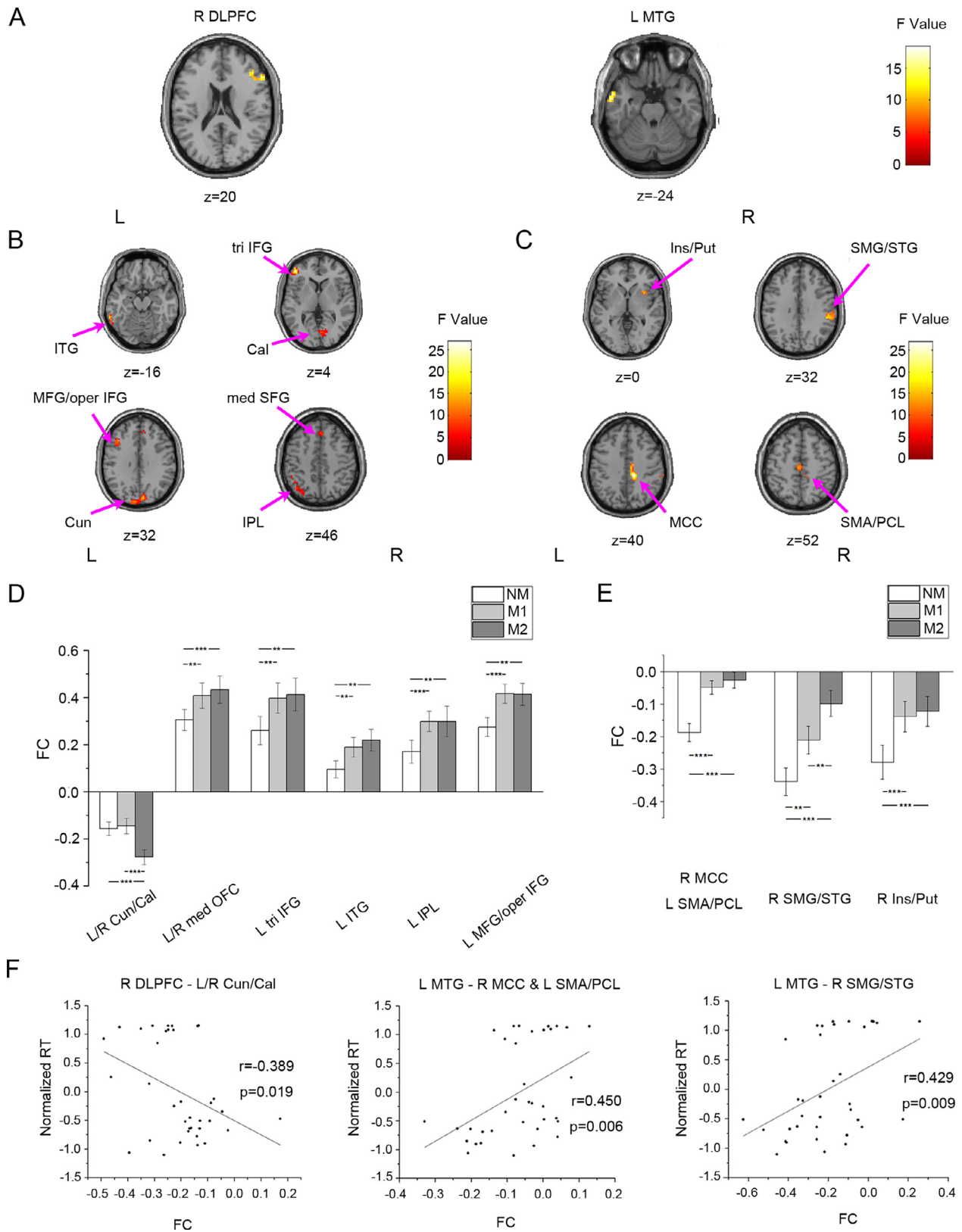


Table 3. Clusters showing significant changes in FC in the whole-brain FC analysis

Analysis	Brain Regions	Number of Voxels	Brodmann's Area	Peak MNI-coordinates			Peak F-score
				x	y	z	
(a) Voxel-to-voxel	R DLPFC	93	46/45/9/8	57	26	16	18.39
	L MTG	34	21/20	−57	2	−23	16.52
(b) Seed-to-voxel (R DLPFC)	L/R Cun/Cal	294	19/18/17/23/30/7	9	−82	31	16.11
	L/R med SFG	63	8/9/6	0	32	37	19.39
	L tri IFG	51	46/10	−45	44	4	26.86
	L ITG	45	20/21/37	−63	−46	−17	19.04
	L IPL	44	7/40	−42	−52	46	11.64
	L MFG/oper IFG	41	9	−39	20	34	15.16
(c) Seed-to-voxel (L MTG)	R MCC	164	31/6/24/5	12	−31	40	25.42
	L SMA/PCL						
	R SMG/STG	110	40/42/2	60	−34	28	16.00
	R Ins/Put	46	13	36	11	−2	14.00

DLPFC, dorsolateral prefrontal cortex; MTG, middle temporal gyrus; Cun, cuneus; Cal, calcarine; SFG, superior frontal gyrus; IFG, inferior frontal gyrus; ITG, inferior temporal gyrus; IPL, inferior parietal lobule; MFG, middle frontal gyrus; MCC, middle cingulate cortex; SMA, supplementary motor area; PCL, paracentral lobule; SMG: supramarginal gyrus; STG: superior temporal gyrus; Ins, insula; Put, putamen; med, medial part; tri, triangular part; oper, opercular part; L, left; R, right.

significantly affected when the signature objects were masked in scene categorization, both reflected in its percent signal change and in FC with the DMN and DLPFC. These changes may reflect the changes in object attention and top-down modulation effect. Further whole-brain analyses were performed to investigate the neural effects or FC on other regions by the mask of signature objects in scene categorization, and we found changes in FC in the FPN, MTG, and other regions involved in semantic information processing. The contributions of the brain activations and FCs to object-masked scene categorization behaviors were further revealed by the behavioral correlation analysis. All these results together suggest that masking of signature objects significantly affected the processing of top-down modulation of attention, cognitive demand, and semantic processing in the human brain.

Changes in FC between LOC and high-level cognitive regions

The LOC showed significant decreases in neural activations after the signature objects were masked, in correspondence with its role in object recognition and object-based scene categorization (Kourtzi and Kanwisher, 2000; Linsley and MacEvoy, 2014; Malach et al., 1995), and consistent with a previous study showing

higher activations in LOC for scenes containing objects than scenes without objects (Harel et al., 2013). No significant change in activations was observed in PPA, possibly because of the basically unchanged spatial properties of scenes (spatial boundary, distance, etc.). In respect of RSC, no significant activation was found in the entire experiment, which might be due to the relatively short presentation time of the visual scenes. Furthermore, RSC might be more involved in spatial navigation and imagination (Epstein, 2008; Vass and Epstein, 2016), which is not much invoked in this scene categorization task.

In the seed-to-voxel analysis, LOC was found showing changes in FC with attention-related regions in the DMN. Brain regions in the DMN show suppression in neural activations in attention-demanding cognitive tasks, and showed anti-correlations with the task-positive regions (Singh and Fawcett, 2008; Raichle et al., 2001; Qiushi et al., 2015; Xu et al., 2017; Geng et al., 2018). This might cause its negative correlations with LOC, which was activated in the task. The medial SFG and ACC were implicated in selective attention to task-relevant stimuli (Weissman et al., 2003, 2005; Boorman et al., 2009; Siltan et al., 2010), and showed interaction with LOC in visual search of objects (Pantazatos et al., 2012). The precuneus is also involved in the processing of spatial and object attention (Cavanna and Trimble, 2006; Sathian et al., 1999; Serences et al., 2004; Mahayana

Fig. 5. Whole-brain FC analysis results. Group-level statistical maps of clusters showing significant change in FC with (A) the whole-brain in the voxel-to-voxel analysis (core hubs) (B) the core hub of right DLPFC in the seed-to-voxel analysis; (C) the core hub of left MTG in the seed-to-voxel analysis in different masking degrees (NM, M1, and M2), significant under a combination of voxel-level $p < 0.001$ uncorrected and a cluster extent FWE-corrected $p < 0.05$. (D) FC calculated between right DLPFC and each significant cluster averaged across all subjects. (E) FC calculated between left MTG and each significant cluster averaged across all subjects. (F) Correlations between FC and normalized behavioral RT. Significant correlations with behavioral RT were found in FC between right DLPFC and bilateral cuneus/calcarine, between left MTG and right MCC & left SMA/PCL and between left MTG and right SMG/STG. All significant contrasts survived an FDR-corrected for multiple comparisons of $p < 0.05$. DLPFC, dorsolateral prefrontal cortex; MTG, middle temporal gyrus; Cun, cuneus; Cal, calcarine; SFG, superior frontal gyrus; IFG, inferior frontal gyrus; ITG, inferior temporal gyrus; IPL, inferior parietal lobule; MFG, middle frontal gyrus; MCC, middle cingulate cortex; SMA, supplementary motor area; PCL, paracentral lobule; SMG: supramarginal gyrus; STG: superior temporal gyrus; Ins, insula; Put, putamen; med, medial part; tri, triangular part; oper, opercular part; L, left; R, right; Error bars indicate SEM; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

et al., 2014). It showed a decrease in activations when the signature objects were masked in the whole-brain activation analysis, which was consistent with a previous study showing an increase in activations in precuneus when subjects observed more objects in scenes (Wolbers et al., 2008). We infer that LOC maintained negative correlations with the DMN in the processing of object attention to guide scene categorization. Masking of signature objects seriously disrupted object attention and recognition, causing the significant decreases in negative FC. The contributions of the above FC to scene categorization behaviors were further revealed in the behavioral correlation analysis. There was no significant correlation of behavioral RT with the percent signal change in LOC, but significant correlations of RT with LOC-based FC, suggesting that LOC might not operate alone, but was modulated by other high-level brain regions to contribute to object-based scene categorization behaviors.

In addition, significant increases in negative FC were observed between LOC and left DLPFC, and between right DLPFC and extrastriate/striate visual cortices. These increases in negative FC might be caused by top-down modulation effect, and be related to conscious object recognition. A number of previous studies have revealed top-down modulation effect of DLPFC on sensory visual cortices when processing relevant stimuli in tasks (Petrides and Pandya, 2002; Gazzaley et al., 2005, 2007). After the masking of signature objects, subjects might attend more to the non-object sources of information (e.g. spatial properties), causing the increase in negative correlations between DLPFC and LOC. In addition, functional couplings between prefrontal and visual cortices were also observed in conscious visual recognition (Dehaene and Naccache, 2001; Mckeef and Tong, 2007; Imamoglu et al., 2012). Therefore, the increase in negative correlation of LOC and extrastriate/striate visual cortex with DLPFC might also be caused by the underdetermined performance of conscious recognition of signature objects to guide scene categorization. Behavioral correlation analysis demonstrated that the top-down modulation effect in conscious visual recognition also contributed to scene categorization behaviors.

Changes in cognitive demand in the FPN and DMN

Changes in neural responses were also found within the FPN and DMN regions. The FPN regions, including left IPL and prefrontal regions, serve common roles in executive function, decision-making, reward representation, and have also been indicated in the judgment of visual scene categories (Paulus et al., 2001; Barbey et al., 2013; Guitart-Masip et al., 2013; Kauffmann et al., 2015; Zhang et al., 2015). Higher activations in these regions might reflect the increase in cognitive demand, and factors which made the response selection more difficult (Schumacher et al., 2003; Barraclough et al., 2004; Vickery and Jiang, 2009). Similar functions were also found in medial OFC, located in the DMN (Plassmann et al., 2010; Diekhof et al., 2012). On the contrary, the higher cognitive demand might cause its decrease in activations. Apart from the activation changes in FPN, multiple regions, mostly belonging to

the FPN, were found showing changes in FC in the seed-to-voxel FC analysis, and DLPFC served as the core hub in this network. Increases in FC between right DLPFC and other fronto-parietal regions might suggest the higher task difficulty in scene category judgement. No significant correlation with behavioral RT was observed in the FC between the fronto-parietal regions. We think this might be due to that the executive control functions served by FPN is a quite complicated, which involve multiple processes, such as conflict resolution, working memory, planning and decision-making (Callejas et al., 2005; Niendam et al., 2012), and some of the processes are not directly related to scene categorization behaviors. These results further support the involvement of high-level regions in object-based scene recognition.

Changes in semantic processing after masking of signature objects

Another core hub in the whole-brain FC analysis was left MTG, which was possibly responsible for the semantic processing of signature objects and scenes in this study, because previous studies showed that this region was involved in multimodal semantic processing and retrieval of semantic information about objects (Mummary et al., 2000; Binder et al., 2009; Visser et al., 2014). It also located in the lateral temporal cortex, which is part of the DMN, causing the negative FC observed. Regions showing significant changes in FC with the core hub in left MTG included MCC, SMA, PCL, SMG, STG, insula, and putamen. Due to involvement of SMA and adjacent MCC in movement control and action-based decision making (Shima and Tanji, 1998; Wunderlich et al., 2009), the changes in FC between left MTG and SMA/MCC might reflect the use of semantic knowledge of signature objects in the button press action of decision making in scene categorization. On the other hand, the SMG/STG and insula/putamen were previously reported in object recognition and in particular, object naming (Bookheimer et al., 1995; Price et al., 2005; Ellis et al., 2006), and showed coactivations with left MTG when subjects naming tools compared to naming animals or viewing nonsense objects (Martin et al., 1996). These regions were also implicated in semantic processing (Binder and Desai, 2011; Vandenberghe et al., 1996), and showed cross-modal representations of object categories in an MVPA study (Devereux et al., 2013). We suggest that the negative FC between MTG and SMG/STG might be responsible for the processing of semantic information of signature objects to guide scene categorization. In addition, a recent study has also shown cross-decoding of scene categories across pictures and words in the left MTG and IPL, extending the semantic encoding in these regions to visual scenes (Kumar et al., 2017). Therefore, our results also provide further neural support for the semantic information relation between signature objects and scenes (Wang et al., 2017).

All together, we suggest that this brain network, including left MTG, inferior SMG/STG and insula/putamen, play vital roles in processing semantic information and motor response, and left MTG serves as

a core hub in this network as the center of semantic information processing. The IPL (SMG) was once again observed, but unlike the left IPL serves for decision-making in the FPN, the right IPL might serve its role in the semantic retrieval of object and scene category information. The decrease in the negative FC of these connections might indicate the reduction of semantic information processed due to the mask of signature objects. The connectivity between left MTG and right MCC, left SMA/PCL and right SMG/STG also showed significant contributions to scene categorization behaviors.

To our knowledge, this is the first study investigating the brain mechanisms underlying object-masked scene categorization revealed by changes in neural activation and FC. The role of LOC in object-based scene categorization was verified in the ROI analysis. Seed-to-voxel analysis and brain-behavioral correlation analysis further implicated that it might contribute to scene categorization by interacting with the brain regions in the DMN, rather than operating alone. Top-down modulation was also found from DLPFC to LOC and other extrastriate/striate visual cortices, and might also reflect conscious recognition of objects in scenes. In the whole-brain analysis, FC changes were found within the FPN, indicating the FPN in the judgment of scene categories, and right DLPFC might serve as the core hub in this network. Another core hub observed in this study was left MTG. Its connections with MCC, SMG/STG and insula/putamen might indicate the processing of semantic information relation between signature objects and scenes. Note that in this scene categorization task, the changes in the neural mechanisms could be related to both scene interpretation and behavioral variables. The brain-behavior correlation analyses were conducted to further dissociate these two effects. The top-down modulation effect between DLPFC and LOC or extrastriate visual cortex, and semantic processing between left MTG and the corresponding regions, were necessary processes in scene interpretation, and also showed significant correlations with task behavior. Therefore, these neural responses could be related to both effects. However, some other neural responses were not directly correlated with task behaviors, such as the signal change in LOC, and FC within the FPN. Due to the evidences of their involvement in scene perception in previous studies (Macevoy and Epstein, 2011; Kauffmann et al., 2015), they might only be related to processes in scene interpretation. Overall, these changes in neural responses suggest that the masking of signature objects significantly affected the object attention, cognitive demand, top-down modulation effect, and semantic judgment.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (No. U1736219, No.61571327 and No. 61503278). The authors would like to thank Prof. Sean MacEvoy (Department of psychology, Boston College, MA, USA) for supplying the stimuli in the experimental procedure.

CONFLICT OF INTEREST

We declare that we have no actual or potential conflict of interest including any financial, personal or other relationships with other people or organizations that can inappropriately influence our work.

REFERENCES

- Arrington CM, Carr TH, Mayer AR, Rao SM (2000) Neural mechanisms of visual attention: object-based selection of a region in space. *J Cogn Neurosci* 12:106–117.
- Baldassano C, Esteva A, Beck D, Fei-Fei L (2015) Two distinct scene processing networks connecting vision and memory. *J Vis* 15:571.
- Bar M (2003) A cortical mechanism for triggering top-down facilitation in visual object recognition. *J Cogn Neurosci* 15:600.
- Barbey AK, Colom R, Grafman J (2013) Dorsolateral prefrontal contributions to human intelligence. *Neuropsychologia* 51:1361–1369.
- Barracough DJ, Conroy ML, Lee D (2004) Prefrontal cortex and decision making in a mixed-strategy game. *Nat Neurosci* 7:404–410.
- Beck DM, Rees G, Frith CD, Lavie N (2001) Neural correlates of change detection and change blindness. *Nat Neurosci* 4:645–650.
- Biederman I (1987) Recognition-by-components: a theory of human image understanding. *Psychol Rev* 94:115.
- Biederman I, Mezzanotte RJ, Rabinowitz JC (1982) Scene perception: detecting and judging objects undergoing relational violations. *Cogn Psychol* 14:143.
- Binder JR, Desai RH (2011) The neurobiology of semantic memory. *Trends Cogn Sci* 15:527.
- Binder JR, Desai RH, Graves WW, Conant LL (2009) Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb Cortex* 19:2767–2796.
- Bookheimer SY, Zeffiro TA, Blaxton T, Gaillard W, Theodore W (1995) Regional cerebral blood flow during object naming and word reading. *Hum Brain Mapp* 3:93–106.
- Boorman E, D, Behrens Timothy EJ, Woolrich Mark W, Rushworth Matthew FS (2009) How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron* 62:733–743.
- Callejas A, Lupiáñez J, Funes MJ, Tudela P (2005) Modulations among the alerting, orienting and executive control networks. *Exp Brain Res* 167:27–37.
- Cavanna AE, Trimble MR (2006) The precuneus: a review of its functional anatomy and behavioural correlates. *Brain A J Neurol* 129:564.
- Davenport JL (2007) Consistency effects between objects in scenes. *M&C* 35:393–401.
- Dehaene S, Naccache L (2001) Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition* 79:1.
- Devereux BJ, Clarke A, Marouchos A, Tyler LK (2013) Representational similarity analysis reveals commonalities and differences in the semantic processing of words and objects. *J Neurosci* 33:18906–18916.
- Diekhof EK, Kaps L, Falkai P, Gruber O (2012) The role of the human ventral striatum and the medial orbitofrontal cortex in the representation of reward magnitude – an activation likelihood estimation meta-analysis of neuroimaging studies of passive reward expectancy and outcome processing. *Neuropsychologia* 50:1252.
- Ellis AW, Burani C, Izura C, Bromiley A, Venneri A (2006) Traces of vocabulary acquisition in the brain: Evidence from covert object naming. *Neuroimage* 33:958–968.
- Epstein R, Kanwisher N (1998) A cortical representation of the local visual environment. *Nature* 392:598–601.
- Epstein RA (2008) Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends Cogn Sci* 12:388–396.

- Epstein RA, Bar M, Kveraga K (2014) Neural systems for visual scene recognition. *Scene Vis*:105–134.
- Epstein RA, Higgins JS, Jablonski K, Feiler AM (2007) Visual scene processing in familiar and unfamiliar environments. *J Neurophysiol* 97:3670.
- Feifei L, Iyer A, Koch C, Perona P (2007) What do we perceive in a glance of a real-world scene? *J Vis* 7:10.
- Gazzaley A, Cooney J, Mcevoy K, Knight R (2005) Top-down enhancement and suppression of the magnitude and speed of neural activity. *J Cogn Neurosci* 17:507.
- Gazzaley A, Rissman J, Cooney J, Rutman A, Seibert T, Clapp W, D'Esposito M (2007) Functional interactions between prefrontal and visual association cortex contribute to top-down modulation of visual processing. *Cereb Cortex* 17:1125–1135.
- Geng X, Xu J, Liu B, Shi Y (2018) Multivariate classification of major depressive disorder using the effective connectivity and functional connectivity. *Front Neurosci* 12.
- Goto SG, Ando Y, Huang C, Yee A, Lewis RS (2010) Cultural differences in the visual processing of meaning: detecting incongruities between background and foreground objects using the N400. *Soc Cogn Affect Neurosci* 5:242–253.
- Graef PD, Christiaens D, D'Ydewalle G (1990) Perceptual effects of scene context on object identification. *Psychol Res* 52:317–329.
- Greenberg AS, Rosen M, Cutrone E, Behrmann M (2015) The effects of visual search efficiency on object-based attention. *Atten Perception Psychophys* 77:1544–1557.
- Greene MR, Oliva A (2009) Recognition of natural scenes from global properties: seeing the forest without representing the trees. *Cogn Psychol* 58:137.
- Grillspector K, Kushnir T, Edelman S, Itzhak Y, Malach R (1998) Cue-invariant activation in object-related areas of the human occipital lobe. *Neuron* 21:191.
- Guitart-Masip M, Barnes GR, Horner A, Bauer M, Dolan RJ, Duzel E (2013) Synchronization of medial temporal lobe and prefrontal rhythms in human decision making. *J Neurosci* 33:442–451.
- Harel A, Kravitz DJ, Baker CI (2013) Deconstructing visual scenes in cortex: gradients of object and spatial layout information. *Cereb Cortex* 23:947–957.
- Imamoglu F, Kahnt T, Koch C, Haynes JD (2012) Changes in functional connectivity support conscious object recognition. *Neuroimage* 63:1909.
- Johnson MR, Johnson MK (2014) Decoding individual natural scene representations during perception and imagery. *Front Hum Neurosci* 8:59.
- Kauffmann L, Bourgin J, Guyader N, Peyrin C (2015) The neural bases of the semantic interference of spatial frequency-based information in scenes. *J Cogn Neurosci* 27:2394–2405.
- Kourtzi Z, Kanwisher N (2000) Cortical regions involved in perceiving object shape. *J Neurosci* 20:3310–3318.
- Kravitz DJ, Peng CS, Baker CI (2011) Real-world scene representations in high-level visual cortex: it's the spaces more than the places. *J Neurosci* 31:7322–7333.
- Kumar M, Federmeier KD, Fei-Fei L, Beck DM (2017) Evidence for similar patterns of neural activity elicited by picture- and word-based representations of natural scenes. *Neuroimage*.
- Liang Y, Liu B, Li X, Wang P (2018) Multivariate pattern classification of facial expressions based on large-scale functional connectivity. *Front Hum Neurosci* 12:94.
- Liang Y, Liu B, Xu J, Zhang G, Li X, Wang P, Wang B (2017) Decoding facial expressions based on face-selective and motion-sensitive areas. *Hum Brain Mapp* 38:3113.
- Linsley D, MacEvoy SP (2014) Evidence for participation by object-selective visual cortex in scene category judgments. *J Vis* 14:19.
- Logan GD (1996) The CODE theory of visual attention: an integration of space-based and object-based attention. *Psychol Rev* 103:603–649.
- Lumer ED, Rees G (1999) Covariation of activity in visual and prefrontal cortex associated with subjective visual perception. *Proc Natl Acad Sci U S A* 96:1669–1673.
- Macevoy SP, Epstein RA (2011) Constructing scenes from objects in human occipitotemporal cortex. *Nat Neurosci* 14:1323–1329.
- Maguire E (2001) The retrosplenial contribution to human navigation: a review of lesion and neuroimaging findings. *Scand J Psychol* 42:225–238.
- Mahayana IT, Tcheang L, Chen CY, Juan CH, Muggleton NG (2014) The precuneus and visuospatial attention in near and far space: a transcranial magnetic stimulation study. *Brain Stimul* 7:673–679.
- Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy WA, Ledden PJ, Brady TJ, et al. (1995) Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc Natl Acad Sci U S A* 92:8135.
- Martin A, Wiggs CL, Ungerleider LG, Haxby JV (1996) Neural correlates of category-specific knowledge. *Nature* 379:649–652.
- Martuzzi R, Ramani R, Qiu M, Shen X, Papademetris X, Constable RT (2011) A whole-brain voxel based measure of intrinsic connectivity contrast reveals local changes in tissue connectivity with anesthetic without a priori assumptions on thresholds or regions of interest. *Neuroimage* 58:1044–1050.
- McCotter M, Gosselin F, Sowden P, Schyns P (2005) The use of visual information in natural scenes. *Vis Cogn* 12:938–953.
- McKeeff TJ, Tong F (2007) The timing of perceptual decisions for ambiguous face stimuli in the human ventral visual cortex. *Cereb Cortex* 17:669.
- Mummery CJ, Patterson K, Price CJ, Ashburner J, Frackowiak RS, Hodges JR (2000) A voxel-based morphometry study of semantic dementia: relationship between temporal lobe atrophy and semantic memory. *Ann Neurol* 47:36.
- Niendam TA, Laird AR, Ray KL, Dean YM, Glahn DC, Carter CS (2012) Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cogn Affect Behav Neurosci* 12:241–268.
- O'Craven KM, Kanwisher N (2014) Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *J Cogn Neurosci* 12:1013–1023.
- Oliva A, Torralba A (2001) Modeling the shape of the scene: a holistic representation of the spatial envelope. *Int J Comput Vis* 42:145–175.
- Pantazatos SP, Yanagihara TK, Zhang X, Meitzler T, Hirsch J (2012) Frontal-occipital connectivity during visual search. *Brain Connect* 2:164–175.
- Park S, Konkle T, Oliva A (2014) Parametric coding of the size and clutter of natural scenes in the human brain. *Cereb Cortex* 25:1792.
- Paulus MP, Hozack N, Zauscher B, McDowell JE, Frank L, Brown GG, Braff DL (2001) Prefrontal, parietal, and temporal cortex networks underlie decision-making in the presence of uncertainty. *Neuroimage* 13:91–100.
- Petrides M, Pandya DN (2002) Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. *Eur J Neurosci* 16:291.
- Plassmann H, O'Doherty JP, Rangel A (2010) Appetitive and aversive goal values are encoded in the medial orbitofrontal cortex at the time of decision making. *J Neurosci* 30:10799–10808.
- Price CJ, Devlin JT, Moore CJ, Morton C, Laird AR (2005) Meta-analyses of object naming: effect of baseline. *Hum Brain Mapp* 25:70.
- Qiushi Z, Gaoyan Z, Li Y, Xiaojie Z (2015) Impact of real-time fMRI working memory feedback training on the interactions between three core brain networks. *Front Behav Neurosci* 9:244.
- Raichle ME, Macleod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL (2001) A default mode of brain function. *Proc Natl Acad Sci U S A* 98:676.
- Sathian K, Simon TJ, Peterson S, Patel GA, Hoffman JM, Grafton ST (1999) Neural evidence linking visual object enumeration and attention. *J Cogn Neurosci* 11:36.
- Schumacher E, Elston P, D'Esposito M (2003) Neural evidence for representation-specific response selection. *J Cogn Neurosci* 15:1111–1121.
- Serences JT, Schwarzbach J, Courtney SM, Golay X, Yantis S (2004) Control of object-based attention in human cortex. *Cereb Cortex* 14:1346.

- Shima K, Tanji J (1998) Role for cingulate motor area cells in voluntary movement selection based on reward. *Science* 282:1335–1338.
- Shomstein S, Behrmann M (2006) Cortical systems mediating visual attention to both objects and spatial locations. *Proc Natl Acad Sci U S A* 103:11387–11392.
- Silton RL, Heller W, Towers DN, Engels AS, Spielberg JM, Edgar JC, Sass SM, Stewart JL, et al. (2010) The time course of activity in dorsolateral prefrontal cortex and anterior cingulate cortex during top-down attentional control. *Neuroimage* 50:1292–1302.
- Singh KD, Fawcett IP (2008) Transient and linearly graded deactivation of the human default-mode network by a visual detection task. *Neuroimage* 41:100–112.
- Vandenberghe R, Price C, Wise R, Josephs O, Frackowiak RSJ (1996) Functional anatomy of a common semantic system for words and pictures. *Nature* 383:254–256.
- Vass LK, Epstein RA (2016) Common neural representations for visually guided reorientation and spatial imagery. *Cereb Cortex*.
- Vickery TJ, Jiang YV (2009) Inferior parietal lobule supports decision making under uncertainty in humans. *Cereb Cortex* 19:916.
- Visser M, Jefferies E, Embleton KV, Lambon Ralph MA (2014) Both the middle temporal gyrus and the ventral anterior temporal area are crucial for multimodal semantic processing: distortion-corrected fMRI evidence for a double gradient of information convergence in the temporal lobes. *J Cogn Neurosci* 24:1766–1778.
- Walther DB, Caddigan E, Fei-Fei L, Beck DM (2009) Natural scene categories revealed in distributed patterns of activity in the human brain. *J Neurosci* 29:10573.
- Wang S, Cao L, Xu J, Zhang G, Lou Y, Liu B (2017) Revealing the semantic association between perception of scenes and significant objects by representational similarity analysis. *Neuroscience*.
- Weissman DH, Giesbrecht B, Song AW, Mangun GR, Woldorff MG (2003) Conflict monitoring in the human anterior cingulate cortex during selective attention to global and local object features. *Neuroimage* 19:1361–1368.
- Weissman DH, Gopalakrishnan A, Hazlett CJ, Woldorff MG (2005) Dorsal anterior cingulate cortex resolves conflict from distracting stimuli by boosting attention toward relevant events. *Cereb Cortex* 15:229–237.
- Whitfieldgabrieli S, Nietocastanon A (2012) Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connect* 2:125.
- Wolbers T, Hegarty M, Büchel C, Loomis JM (2008) Spatial updating: how the brain keeps track of changing object locations during observer motion. *Nat Neurosci* 11:1223.
- Wu CC, Ahmed WF, Marc P (2014) Guidance of visual attention by semantic information in real-world scenes. *Front Psychol* 5:54.
- Wunderlich K, Rangel A, O'Doherty JP (2009) Neural computations underlying action-based decision making in the human brain. *Proc Natl Acad Sci U S A* 106:17199–17204.
- Xu J, Yin X, Ge H, Han Y, Pang Z, Liu B, Liu S, Friston K (2017) Heritability of the effective connectivity in the resting-state default mode network. *Cereb Cortex*.
- Yang X, Xu J, Cao L, Li X, Wang P, Wang B, Liu B (2018) Linear representation of emotions in whole persons by combining facial and bodily expressions in the extrastriate body area. *Front Hum Neurosci* 11:653.
- Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, et al. (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol* 106:1125–1165.
- Zhang G, Cheng Y, Liu B (2016) Abnormalities of voxel-based whole-brain functional connectivity patterns predict the progression of hepatic encephalopathy. *Brain Imag Behav*:1–13.
- Zhang G, Yao L, Shen J, Yang Y, Zhao X (2015) Reorganization of functional brain networks mediates the improvement of cognitive performance following real-time neurofeedback training of working memory. *Hum Brain Mapp* 36:1705.

(Received 2 May 2018, Accepted 18 July 2018)
(Available online 26 July 2018)