Neuropsychologia 50 (2012) 862–868

Contents lists available at SciVerse ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Similar spatial patterns of neural coding of category selectivity in FFA and VWFA under different attention conditions

Guifang Xu [a,b], Yi Jiang [c,d], Lifei Ma [a,b], Zhi Yang [a,d,*], Xuchu Weng [a,e,**]

[a] Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China
[b] Graduate University of Chinese Academy of Sciences, Beijing 100039, China
[c] Key Laboratory of Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China
[d] Magnetic Resonance Imaging Research Center, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China
[e] Center for Cognition and Brain Disorders, Hangzhou Normal University, Hangzhou 310003, China

ARTICLE INFO

Article history:
Received 5 September 2011
Received in revised form 20 January 2012
Accepted 20 January 2012
Available online 28 January 2012

Keywords:
fMRI
Fusiform face area
Visual word form area
Multi-voxel pattern analysis (MVPA)
Contrast-to-noise ratio

ABSTRACT

It has long been debated whether attention alters the categorical selectivity in regions such as the fusiform face area (FFA) and the visual word form area (VWFA). We addressed this issue by examining whether the spatial pattern of neural representations for certain stimulus categories in these regions would change under different attention conditions. Faces, Chinese characters, and textures were presented in a block design fMRI experiment where participants in different runs attended to the stimuli under different conditions of attention. After localizing regions of interest (ROIs) in FFA and VWFA using general linear models, we performed spatial pattern analyses to examine both within- and cross-condition classification in these ROIs. The within-condition results replicated previous findings showing significant classification accuracy reduction when there was less attention compared with more attention. Critically, cross-condition classification in both FFA and VWFA revealed significantly above-chance accuracy for all stimulus categories, suggesting similar spatial neural representations across different attention conditions. Further strengthening this conclusion, when the contrast-to-noise ratio (CNR) of the signals was adjusted to increase signal strength, cross-condition classification accuracy for faces in FFA and for Chinese characters in VWFA improved significantly, even approaching within-condition accuracy. This indicates that attention does not modulate the spatial pattern of neural representations involved in category selectivity, but only changes the signal strength relative to the noise level.

1. Introduction

Attention plays an important role in visual information processing. One important question about attention is to understand how it modulates neural representations in visual regions demonstrating category-selectivity. Functional magnetic resonance imaging (fMRI) studies have revealed attentional modulation of the magnitude of the blood oxygen level dependence (BOLD) responses showing enhanced activations for attended than less-attended visual stimuli (Murray & Wojciulik, 2004; Wojciulik, Kanwisher, & Driver, 1998). More recently, several studies adopted a new approach of Multi-Voxel Pattern Analysis (MVPA, Haynes & Rees, 2005; Kamitani & Tong, 2005) which is sensitive to the spatial pattern in response to the stimulus content (Norman, Polyn, Detre, & Haxby, 2006). These studies demonstrated that the within-condition classification accuracies in the fusiform face area (FFA) and the parahippocampal place area (PPA) were critically dependent on attention levels (MacEvoy & Epstein, 2009; Reddy & Kanwisher, 2007; Reddy, Kanwisher, & VanRullen, 2009; Sterzer, Haynes, & Rees, 2008). Nevertheless, the classification performance for the preferred stimulus categories (i.e., faces for the FFA and houses for the PPA) were still significantly above chance even when the stimuli were not attended, suggesting that the category selectivity was somehow preserved.

Although these studies indicate that attention modulates category specificity, the underlying mechanism of such modulation remains to be explored. One possible computational model where attention affects category specificity is that the population-coded neural representation may be "diffuse" under limited attention but gets sharpened (becoming more "sparse") when there is more attention to the stimuli. In support of such a model, some fMRI studies showed that orientation coding under attended and unattended conditions engaged different spatial patterns of activation in lateral occipital cortex and primary visual cortex (Fischer & Whitney, 2009; Murray & Wojciulik, 2004). In contrast, some other

* Corresponding author at: 4A Datun Road, Chaoyang District, Beijing 100101, China. Tel.: +86 10 64888632; fax: +86 10 64888632.
** Corresponding author at: 4A Datun Road, Chaoyang District, Beijing 100101, China. Tel.: +86 10 64854476; fax: +86 10 64872070.
E-mail addresses: yangz@psych.ac.cn (Z. Yang), wengxc@psych.ac.cn (X. Weng).

0028-3932/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuropsychologia.2012.01.026

studies support a different model where attention changes the amplitude of neural activity but not the neural population coding specific categories of stimuli (McAdams & Maunsell, 1999; Treue & Martinez Trujillo, 1999; Treue & Maunsell, 1999). Clearly, the two models have opposite predictions as to whether the spatial pattern responsible for coding specific categories remains the same under different attention conditions.

The current study addressed this issue by examining the multi-voxel pattern classification performance under different attention conditions. In one condition, participants were asked to make some fine judgments about the stimuli (e.g., faces, textures, and characters) requiring a high level of attention directed to the stimuli. In the other condition, they paid much less attention to these stimuli as they only need to discriminate these stimuli from some visually very different items. Critically, we performed cross-condition classifications by training classifiers with the responses in one condition and testing its classification performance in the other, in order to examine whether there are similar spatial patterns of neural coding of category selectivity between these two conditions.

Furthermore, we introduced a contrast-to-noise ratio (CNR) manipulation in the cross-condition classifications to test whether the CNR is a factor relevant to classification performance under different attention conditions. Although it could be crucial for the interpretation of MVPA performance, only a few studies have examined this issue (e.g., Smith, Kosillo, & Williams, 2011). Specifically, if attention modulates the magnitudes rather than the spatial patterns of the neural activities, the cross-condition classification accuracy should be well above the chance level, and, more importantly, could be monotonically increased by amplifying the CNR of the data. Alternatively, if attention changes the spatial patterns of the neural responses, one would only expect a close-to-chance accuracy in such cross-condition classification, and the increase of the CNR would not significantly improve the performance of the cross-condition classification.

Other than FFA, we also examined the visual word from area (VWFA), another region in the ventral visual object pathway known to be sensitive to written words (Cohen & Dehaene, 2004; Cohen et al., 2000, 2002; McCandliss, Cohen, & Dehaene, 2003), and presumably at the similar functional level as the FFA in the visual processing hierarchy. Different from FFA, the functional selectivity of the VWFA for visual words is still debated (Baker et al., 2007; Hasson, Levy, Behrmann, Hendler, & Malach, 2002; Indefrey et al., 1997; Reinke, Fernandes, Schwindt, O'Craven, & Grady, 2008; Tagamets, Novick, Chalmers, & Friedman, 2000). Inspection of both the FFA and the VWFA using the MVPA approach may from a different perspective help to inform whether the VWFA has similar functional specialization as the FFA.

2. Methods

2.1. Participants

Eleven native Chinese (Mandarin) speakers participated in the study (6 females, 5 males; mean age = 22.6 years, ranged from 20 to 23). All were right-handed college students (assessed by a Chinese handedness questionnaire described in Li, 1983), with normal or corrected-to-normal vision and free of neurological diseases or psychiatric disorders. None majored in linguistics or related disciplines. Written informed consent was obtained from each participant before the experiment following a protocol approved by the IRB of the Institute of Psychology, Chinese Academy of Sciences.

2.2. Materials and procedures

Three categories of stimuli, including faces, textures and Chinese characters, were used in the experiment. Each category contained 80 images. The faces consisted of 40 female and 40 male faces; the textures consisted of 40 coarse- and 40 fine-grained textures; and the Chinese characters consisted of high frequency characters, 40 with an up-down structure and 40 with a left-right structure.

The experiment consisted of two runs, one for the High Attention condition (HighAtt) condition, and one for the Low Attention condition (LowAtt). The run order was counter-balanced across-subjects. For the first task, participants were asked to press a button on a button-box when they saw some stimuli (male faces, fine grained textures, or up-down structured characters) but not to respond otherwise. For the second task, 20% of the stimuli were randomly replaced with nonsense geometrical shapes, and participants were asked to respond to these oddball images with a button press but not to respond otherwise. Both speed and accuracy were emphasized. As shown in Fig. 1, each run consisted of 4 replications of 3 blocks, with one block for each of the three stimulus categories. The block order for the three categories was pseudo-randomized. Each block lasted 20 s with a 20 s fixation interval between successive blocks. Each block involved the presentation of 20 images (each for 200 ms), interleaved with a central fixation cross shown for 800 ms.

2.3. MRI data acquisition

All images were acquired with a GE 3.0 T Signa MRI scanner (Milwaukee, Wisconsin, USA). During the scan, a tight cushion was used to immobilize the head and reduce head motion. A single-shot T2* weighted gradient echo planar sequence was used to acquire the functional images with the following parameters: TR/TE/flip angle = 2000 ms/40 ms/90°, FOV = 240 mm × 240 mm, matrix = 64 × 64, slice thickness = 4 mm, gap = 0 mm. Twenty contiguous axial slices covering the occipital and temporal lobes were acquired. After the functional scans, high-resolution 3D images were acquired using a spoiled gradient echo sequence (TR = 6.8 ms, minimal TE, flip angle = 12°, FOV = 280 mm × 280 mm, matrix = 256 × 256, slice thickness = 1 mm).

2.4. Data analysis

2.4.1. Functional localizer

The functional data was preprocessed using AFNI (Analysis of Functional Neural Image, Cox, 1996). The two functional runs were preprocessed separately, starting from head motion corrections, and followed by co-registration of the anatomical image to the functional images, linear trend correction, and time course standardization involving subtraction of the mean signal and division by the standard deviation of the entire time course yielding a Z-score for each time point. Finally, the data were spatially smoothed with a Gaussian kernel of 5 mm FWHM (full width at half magnitude). Note that the data were not transformed into the Talairach coordinate system.

A multiple regression analysis was then conducted for each individual participant, with three regressors modeling BOLD responses to the three different stimuli categories (faces, characters, and textures), using the 'BLOCK' model in AFNI. Six head-motion parameters derived from motion corrections and four parameters used to estimate the polynomial signal baseline were included as nuisance regressors. The 20-s fixation periods were modeled as the baseline condition. The HighAtt and LowAtt runs were first concatenated in the regression analysis to localize the VWFA and FFA. The VWFA was defined within the left fusiform gyrus based on a contrast of the characters against the other two stimulus categories (i.e., character − 0.5 × face − 0.5 × texture), and the FFA was defined within the right fusiform gyrus based on a contrast of the faces against the other two categories (i.e., face − 0.5 × texture − 0.5 × character). The two ROIs, referred to as VWFA_H_L and FFA_H_L, were localized individually for each participant using a statistical criterion $q < 0.05$ (false discovery rate), and used in the within-condition classification analyses to replicate findings in previous studies. Two more ROIs, referred to as VWFA_H and FFA_H were defined following the same procedure except using only the HighAtt run data to make a relative independent localizer (Kriegeskorte, Simmons, Bellgowan, & Baker, 2009; Vul, Harris, Winkielman, & Pashler, 2009) for further cross-condition classification analyses. The mean coordinates and volumes (x, y, z, voxels) were (37.9, −53.7, −14.3, 52) and (37.9, −52.3, −14.8, 46) for FFA_H_L and FFA_H, and (−42.5, −57.6, −7.4, 44) and (−43.8, −55.6, −8.8, 48) for VWFA_H_L and VWFA_H. Fig. 2 shows the location and extent of such ROIs for a representative participant (see Supplementary Table S1 for the same information of all participants).

2.4.2. Multi-voxel pattern analysis (MVPA)

Preprocessing for the MVPA was the same as in the GLM analysis except that no spatial smoothing was conducted. The MVPA analysis involved three parts.

(1) The data from the HighAtt and LowAtt runs were analyzed separately to obtain within-condition classification accuracies in the predefined FFA_H_L and VWFA_H_L ROIs. To avoid effects of signal ramping at block onsets and offsets, the stimuli blocks (sequences of 0 s and 1 s) were first convolved with a hemodynamic response model and then thresholded at a magnitude at 0.8 (Coutanche, Thompson-Schill, & Schultz, 2011). Nine time points in the plateau of each block (from 7 s to 24 s after the block-onset), were selected and then used in the following MVPA analysis. Back-propagation neural network implemented in the MVPA toolbox (Detre et al., 2006) was used to classify the stimulus categories. To control for voxel number effects (Reddy & Kanwisher, 2007), 18 voxels from each ROI were randomly sampled to perform a four-fold (leave one block out) cross-validation analysis. Specifically, the signals from the 18 voxels were arranged into an 18-dimension vector (pattern) with each element being a Z-score of a voxel. Thus, a total of 36 such patterns (9 in each block × 4 blocks) were obtained for each of the stimuli categories. For the four-fold cross-validation, 81 (27 × 3) of the patterns (from nine blocks, three for each category) were used as the training

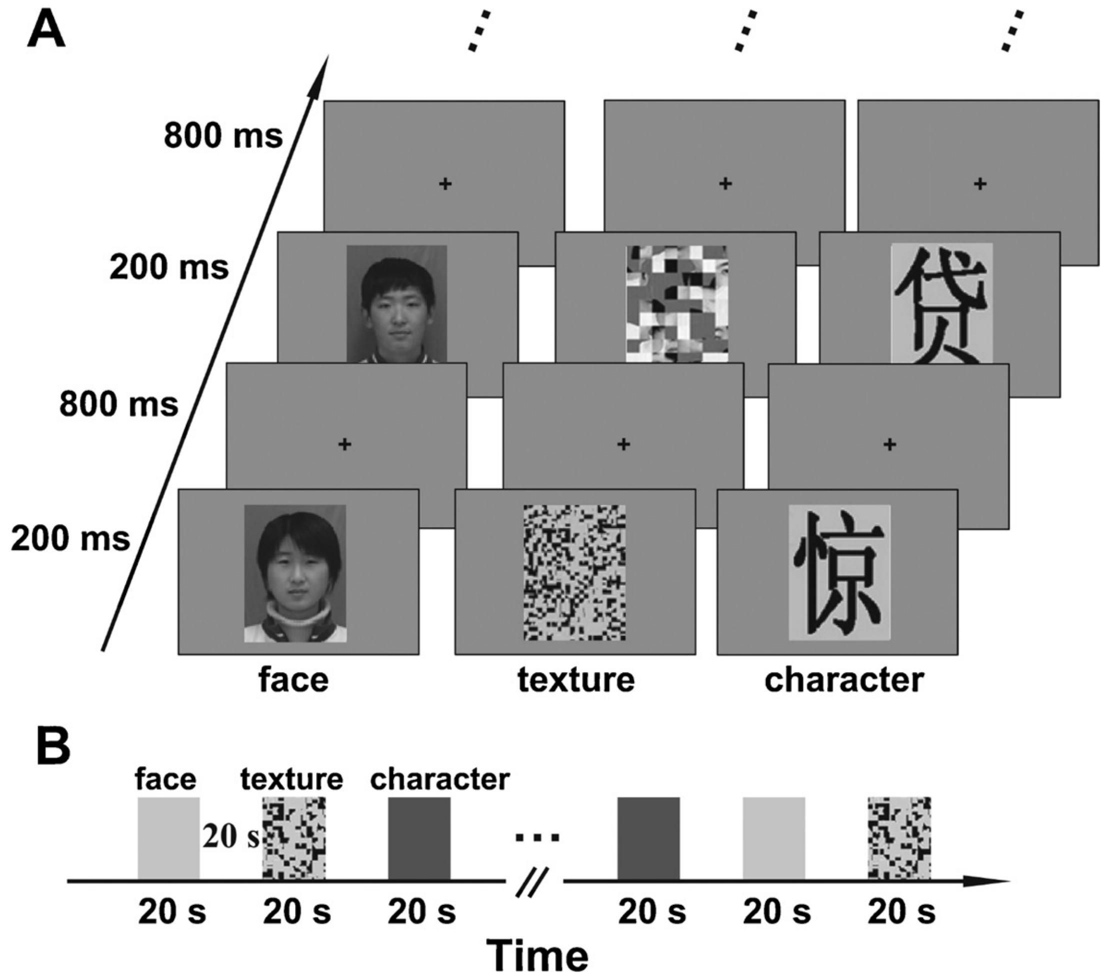

Fig. 1. The experimental paradigm and example stimuli. Attention was manipulated to include a High Attention condition (HighAtt) and a Low Attention condition (LowAtt). The figure shows the HighAtt condition. There were 4 replications of 3 block groups, each block for one of the three stimulus categories. The length of each block was 20 s with a 20 s fixation between blocks. Each image was presented for 200 ms, separated by an 800 ms fixation. The setup of the LowAtt condition was the same except that 20% of the stimuli were randomly replaced with nonsense geometric images serving as the detection targets.

set, while the remaining 27 (9×3) patterns (from the other three blocks, one for each category) were used as the test set. The back-propagation classifiers were then trained to classify the patterns into three categories by learning the relationships between the patterns and their category labels. The classification accuracies were obtained by applying the trained classifiers to the test set. Quantitatively, the accuracy for each category was defined as the proportion of the patterns that were classified into the correct category (including true positive and true negative cases). By this definition, a non-informative pattern would have one third of chance to be classified into each of the three stimulus categories, yielding a chance-level classification accuracy at 33.3%. The training and test sets were then reconstructed with different combinations of blocks, and the above classification was repeated four times. The final classification accuracies were obtained by averaging the accuracies from the four cross-validations. This procedure was repeated for 100 times with different sets of voxels, and the resulting performance accuracies were further averaged (Reddy & Kanwisher, 2007). Paired-sample *t*-tests were then conducted to examine group-level performance differences between the HighAtt and LowAtt runs for each stimulus category.

(2) A cross-condition classification was performed between the HighAtt and LowAtt runs using the FFA_H and VWFA_H ROIs to avoid artificial enhancement of classification performance when using information from both runs. To match with the within-condition classification, three blocks from the HighAtt run were used to train the classifiers, while the other block from the LowAtt run as the test dataset. The same algorithm as described above was applied in the classification except the four-fold cross-validation. The roles of the two runs were then switched. Classification accuracies from the two iterations were averaged to represent the cross-condition classification performance (two-fold cross-validation).

(3) In this analysis, cross-condition classifications were performed at different levels of CNR. The fitted signal for the LowAtt condition (from the GLM analysis) was multiplied with an amplifying factor (ranging from 1 to 3.8 with a step of 0.4) and then added with the residuals (see Fig. S1 in the Supplementary Materials for a flowchart) to produce a signal with different levels of CNR. Cross-condition classification was performed using the CNR-adjusted LowAtt runs as the training dataset and the HighAtt runs as the test dataset. As in the second analysis described above, only three blocks from each CNR-adjusted LowAtt run were used in training the classifiers, and the analysis was repeated four times to cover all possible combinations when selecting three blocks from four.

3. Results

3.1. Behavior results

For all participants in the LowAtt condition, the hit rate of the oddball detection was above 97%, and the false alarm rate was below 4%. In the HighAtt condition, the mean hit and false alarm rates were 96.6% and 4.3% for faces, 99.5% and 2.3% for textures, and 100% and 3.6% for characters. The hit rate was significantly lower for faces than for characters ($t(10) = -4.33$, $p < 0.02$). The reaction times (RTs) to faces, textures, and characters were 435 ± 41 ms, 382 ± 50 ms, 463 ± 41 ms, respectively, showing a significant effect of stimulus category ($F(2, 30) = 8.39$, $p < 0.001$). Paired *t*-tests revealed that response to textures was significantly faster than to faces ($t(10) = 4.29$, $p < 0.002$) and characters ($t(10) = 6.36$, $p < 0.001$), and response to faces was faster than to characters ($t(10) = 2.87$, $p < 0.02$).

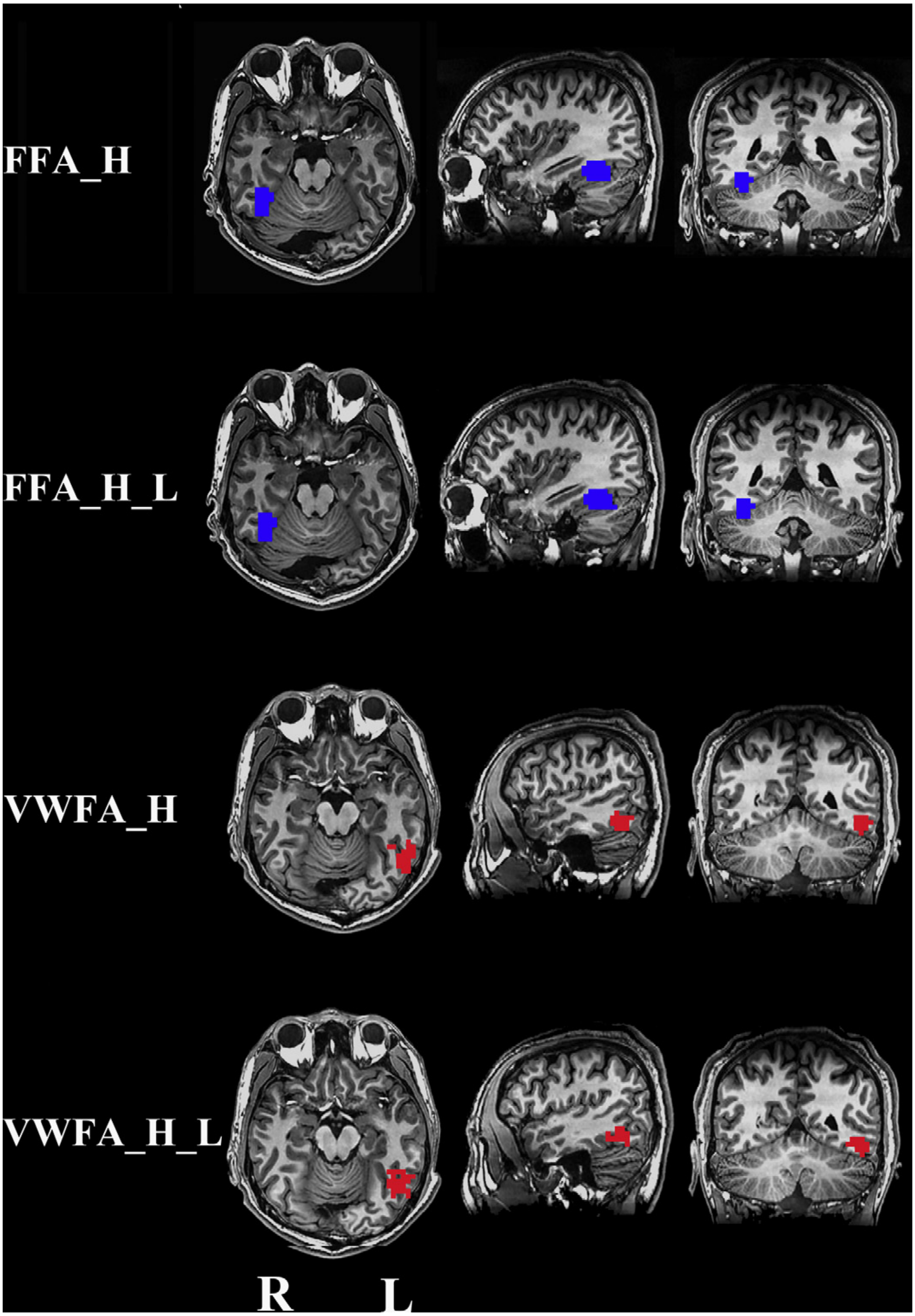

Fig. 2. The location and extent of the FFA (blue) and VWFA (red) ROIs (regions of interest) in a representative participant, defined in two ways. The FFA_H and VWFA_H ROIs were defined only using the HighAtt run data, while the FFA_H_L and VWFA_H_L ROIs were defined using the concatenation of the HighAtt and LowAtt runs. VWFAs were based on a contrast between characters and faces and textures (character – 0.5 × face – 0.5 × texture) within the left fusiform gyrus. FFAs were based on a contrast between faces and characters and textures (face – 0.5 × texture – 0.5 × character) within the right fusiform gyrus. The statistical criteria were $q < 0.05$ (false discovery rate) with a cluster size from 25 to 100 voxels. (The reader is referred to the web version of the article for a colored figure.)

3.2. Comparison of signal changes

At group level, signal changes between the HighAtt and LowAtt conditions were compared. For both the FFA_H_L and VWFA_H_L, their preferred stimulus categories (faces for FFA, characters for VWFA) showed enhanced activations in the HighAtt condition than the LowAtt condition (FFA: $t(10) = 6.05$, $p < 0.001$; VWFA: $t(10) = 5.28$, $p < 0.001$), whereas their non-preferred categories showed no different activations across condition (FFA: character, $t(10) = 1.43$, $p > 0.18$; texture, $t(10) = 0.55$, $p > 0.50$; VWFA: face, $t(10) = 0.73$, $p > 0.49$; texture, $t(10) = -0.19$, $p > 0.85$, see Fig. 3).

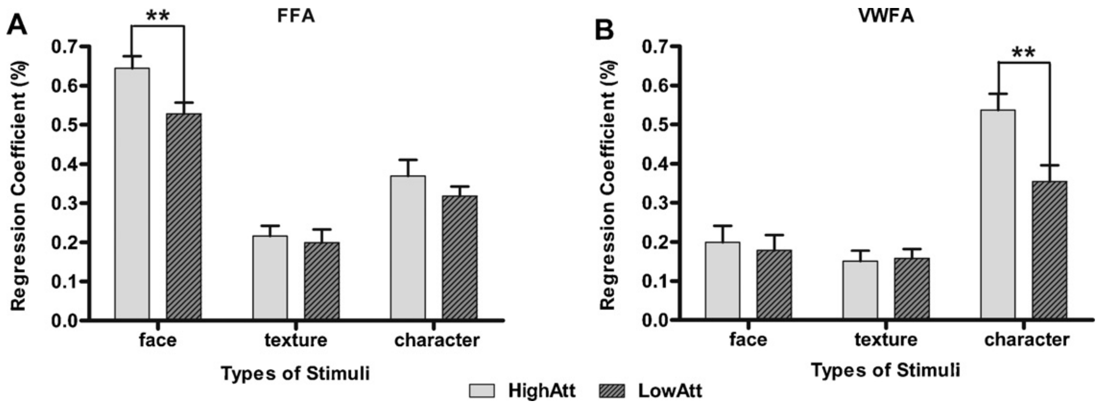

Fig. 3. Signal changes in the two attention conditions. In both the FFA (A) and VWFA (B), only the preferred stimuli categories (faces for FFA, characters for VWFA) showed significant differences across the two conditions.

3.3. MVPA results

For easy descriptions, the within-condition classifications in the HighAtt and LowAtt conditions were referred as the H-H and L-L classifications, the cross-condition classifications as the H-L and L-H classifications, and the cross-condition classification using the CNR-adjusted data of the LowAtt runs for training and the HighAtt runs for testing as the L-H (CNR) classification.

As shown in Fig. 4, in both the FFA_H and the VWFA_H ROIs, classification accuracies for faces and characters were significantly higher in the H-H classification than in the L-L classification (FFA: face, $t(10) = 4.86$, $p < 0.001$; character, $t(10) = 11.57$, $p < 0.001$; VWFA: face, $t(10) = 8.98$, $p < 0.001$; character, $t(10) = 5.84$, $p < 0.001$). Still, the L-L classification accuracies for faces in the FFA ($t(10) = 8.54$, $p < 0.001$) and for characters in the VWFA ($t(10) = 4.92$, $p < 0.001$) were significantly above the chance level (33.33%). For the L-L classification accuracies, the FFA showed above-chance accuracy for its non-preferred stimuli as well, i.e., characters ($t(10) = 5.82$, $p < 0.001$) and textures ($t(10) = 8.09$, $p < 0.001$), while the VWFA showed above-chance accuracy for textures ($t(10) = 3.49$, $p < 0.006$) but not for faces ($t(10) = 1.06$, $p > 0.30$). The results for the textures are shown in Fig. 4 without further discussion as we are not interested in the representations of textures in either the FFA or VWFA.

The cross-condition classification analyses showed that the accuracies for both faces and characters were significantly above chance in both the FFA and VWFA (FFA: face, $t(10) = 14.76$, $p < 0.001$; character, $t(10) = 11.70$, $p < 0.001$; VWFA: face, $t(10) = 6.18$, $p < 0.001$; character, $t(10) = 6.55$, $p < 0.001$). The H-L and L-H classification accuracy for faces was significantly higher than that for characters in the FFA ($t(10) = 4.45$, $p < 0.001$), while in the VWFA, the accuracy for characters was significantly higher than for faces ($t(10) = 1.94$, $p < 0.04$; see Fig. 4).

The above results indicate the spatial patterns representing faces and characters may be similar under different attention conditions in both the FFA and the VWFA. Testing this conclusion, the next analysis checked whether CNR level could explain the difference of classification performance under the two attention conditions (high vs. low). Fig. 5 shows the L-H (CNR) classification accuracies as a function of the CNR amplifying factor within the FFA_H and the VWFA_H. In the FFA, the L-H (CNR) accuracies for both faces and characters increased with the level of CNR. The initial L-H (CNR) classification accuracy for faces was significantly lower than in the H-H classification ($t(10) = -3.81$, $p < 0.003$, the first bar and the solid horizontal line in Fig. 5A). However, when the CNR amplifying factor reached 1.4, the L-H (CNR) classification accuracy for faces was increased to be no different from the H-H classification ($t(10) = 1.13$, $p > 0.30$). Furthermore, when the amplifying factor was bigger than 2.2, the L-H (CNR) accuracies for faces were even higher than in the H-H condition ($t(10) = 2.65$, $p < 0.03$).

Similarly, for characters, the initial L-H (CNR) accuracy was significantly lower than in the H-H classification ($t(10) = -3.85$, $p < 0.003$, the second bar and the dotted horizontal line in Fig. 5A), but it was significantly increased as the amplifying factor was increased. When the factor reached 1.4, the L-H (CNR) accuracy for characters was similar as for the H-H classification ($t(10) = 1.42$,

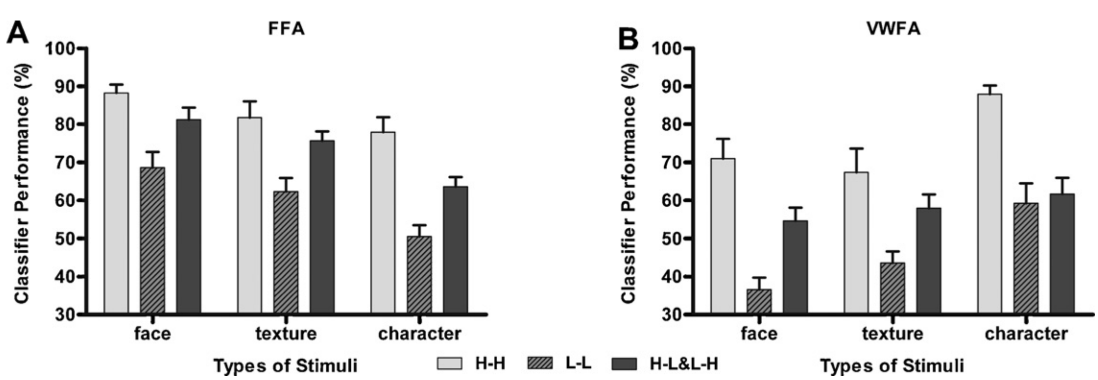

Fig. 4. Accuracies for the within-condition classifications (H-H, L-L) and the cross-condition classifications (H-L and L-H). In both the FFA (A) and VWFA (B), the accuracies of H-H classification for faces and characters were significantly higher than that of L-L classification ($p < 0.001$). The L-L classification accuracies for faces in the FFA ($p < 0.001$) and for characters ($p < 0.001$) in the VWFA were significantly above chance level (33.33%). The cross-condition classification showed that the accuracies for both faces and characters were significantly above chance level in both the FFA and VWFA ($p < 0.001$). The category-preference of both regions was retained in the cross-condition classifications with higher accuracy in for faces than for characters in FFA ($p < 0.001$), and for characters than for faces in VWFA ($p < 0.05$). Error bars indicate across-subject standard deviations.

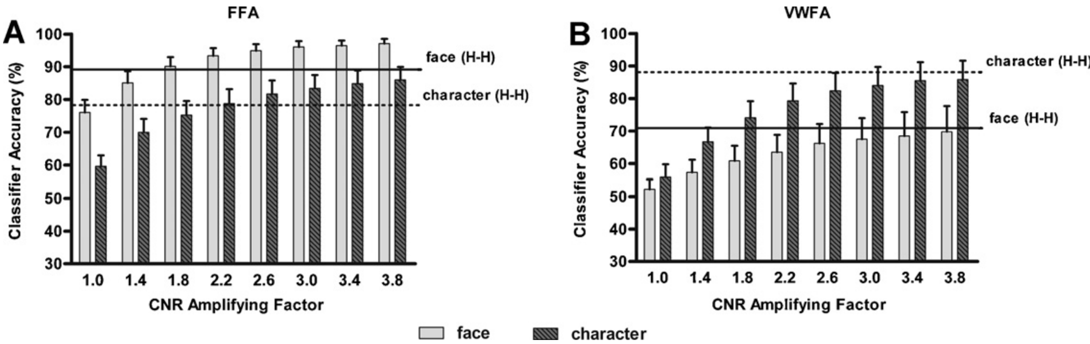

Fig. 5. Cross-condition classification accuracies for faces and characters increased with the CNR amplifying factor in both FFA and VWFA. The solid and dotted horizontal lines indicate the H-H classification accuracies for faces and characters in the two ROIs.

$p > 0.19$). In the VWFA, the initial L-H (CNR) accuracy for characters was significantly lower than in the H-H classification ($t(10) = -11.04$, $p < 0.001$, the second bar and the dotted horizontal line in Fig. 5B), but when the factor reached 2.2, the difference was no longer present ($t(10) = 1.64$, $p > 0.13$). For faces, the initial L-H (CNR) accuracy was significantly lower than in the H-H classification ($t(10) = -3.50$, $p < 0.006$, the first bar and the solid horizontal line in Fig. 5B), but the difference was no longer there when the factor reached 1.8 ($p > 0.10$). In summary, the L-H (CNR) accuracies for faces and characters increased monotonically with the CNR level in both ROIs, and were comparable to the H-H accuracies when the CNR level was high enough.

4. Discussion

Visual processing in the human brain involves highly specialized coding mechanisms to handle the vast amounts of visual input. Category selectivity is one manifestation of such general principles. Attention provides a different type of mechanism to enhance information processing efficiency by selecting what is relevant and ignoring what is irrelevant. To examine the interaction between category selectivity and attention, the present study used the multi-voxel pattern analysis to examine how attention modulates neural coding in two well-known brain regions showing category selectivity, i.e., the FFA and the VWFA. Specifically, the issue addressed was whether different levels of attention alter the spatial pattern of the neural representations for faces and Chinese characters in the FFA and VWFA, respectively. A related issue examined was whether the Contrast Noise Ratio affects MVPA performance under different attention conditions.

With the same set of critical stimuli, we manipulated the task instructions to create two levels of attention to the test stimuli, a High Attention condition and a Low Attention condition. Our results showed that, compared with the high attention condition, both the BOLD responses and the within-condition classification performance were significantly reduced in the low attention condition. The results are consistent with previous findings (Murray & Wojciulik, 2004; Reddy & Kanwisher, 2007; Sterzer et al., 2008; Wojciulik et al., 1998) and confirm the validity of the attention manipulation.

The cross-condition classification results showed that the spatial pattern of neural activation in response to a certain category of stimulus under the high attention condition can be used to decode the neural responses to the same stimulus category under the low attention condition, and vice versa, with significantly above chance accuracy. This result was found for all three categories of stimuli in the two ROIs examined, regardless of category preference of the ROIs. In addition, classification accuracy for the preferred category of an ROI was significantly higher than that for the non-preferred categories. Confirming the functional selectivity for the FFA and VWFA, the results suggest that similar neural populations were involved in coding a specific category of stimulus under different attention conditions; otherwise, one would not expect such high levels of cross-condition classification accuracies.

We also manipulated the CNR factor to examine its effect on the cross-condition classification. The fitted signal changes are usually assumed to reflect the neural responses to a stimulus. The regression residual reflects the noise consisting of stochastic competing neural activities and fMRI measurement noise (Worsley et al., 1996) that remains at a relatively consistent level across different conditions. Under the "signal + noise" model of the measured neural activity (Friston et al., 1995), the CNR adjustment should have mainly enhanced the strength of the "signal" while keeping the "noise" unchanged, that is, the adjustment changes the relative strength of the signal but not the spatial pattern of the signal. Therefore, the result from the CNR adjustment is consistent with the second model described in the introduction, that is, similar neuronal populations are involved in stimulus coding under different attention conditions differing only in the level of signal strength. In such a situation, one would expect to see reasonably high accuracies for cross-condition classification, which can be further increased when the signal is boosted in the low attention condition, as we indeed observed here. This finding suggests that a certain category of stimulus is coded by a specific neural substrate and different levels of attention merely change the response level of such a substrate quantitatively, as opposed to engaging a different neural substrate.

In comparison, if the BOLD responses to a certain category of stimulus under different attention conditions originated from distinct neural representations, that is, one set of neurons was used for stimulus coding for the low attention condition and a different set of neurons was used for stimulus coding for the high attention condition (per the first computational model described in Section 1), one would not expect high cross-condition classification accuracies, nor to expect such accuracies to improve with increasing CNR, and in particular, to improve to a level matching the within-condition classification. Taken together, our results suggest that the spatial patterns of neural coding of category selectivity are specialized in the FFA and VWFA, where similar populations are used for certain stimulus categories (i.e., faces and words, respectively). Although some studies (e.g., Fischer & Whitney, 2009; Murray & Wojciulik, 2004) appear to imply that neural spatial patterns differed across attention conditions in other visual regions (i.e., primary visual areas and lateral occipital cortex), our current findings do not necessarily contradict with these early observations. It is likely that in previous studies attention did not completely change, but rather "sharpen", neuronal activation patterns. Such relatively weak changes are technically difficult to detect with the current resolution of MVPA. Another possible reason to account for the seemingly discrepancy is the difference of paradigms used in the current study and studies of Fischer and Whitney (2009) and Murray and Wojciulik (2004). Our study compared the spatial patterns between different categories, while these two studies focused on visually very similar stimuli.

Using a paradigm requiring participants to attend to one of two simultaneously presented pictures, Reddy and Kanwisher (2007) demonstrate that information about the preferred visual categories in the FFA and PPA was somehow preserved even when the stimuli were unattended. Taking a different approach of cross-condition classification, the present study supports and extends their finding by showing that the pattern of neural responses in the FFA and the VWFA remains unchanged when attention is altered.

Note that previous studies measuring the overall magnitude of BOLD signals reached different conclusions regarding the functional selectivity of the VWFA for visual words (Baker et al., 2007; Hasson et al., 2002; Indefrey et al., 1997; Reinke et al., 2008; Tagamets et al., 2000). The present results obtained from the MVPA method suggest that the spatial pattern of activation potentially constitutes a better indicator of category selectivity (Reddy & Kanwisher, 2007), supporting the functional selectivity of the VWFA for visual words from a different perspective.

The present results showed that manipulation of the CNR of the data affected MVPA performance. Therefore, when claiming a difference on MVPA performance, the effect of the CNR should be considered, particularly when comparing and interpreting MVPA performance across different groups of subjects or different ROIs (Coutanche et al., 2011; Diana, Yonelinas, & Ranganath, 2008; Kamitani & Tong, 2006). This issue has been rarely investigated. For example, Smith et al. (2011) manipulated the amplitude of neural activities by varying the contrast of the stimulus, and found that as the response amplitude was increased with contrast, the orientation classification performance increased approximately linearly with the logarithm of stimulus contrast.

In conclusion, using cross-condition classifications, the present study investigated whether attention alters the spatial coding of category selectivity in the FFA and VWFA. Our findings provide compelling evidence that preferred stimulus categories in the FFA and the VWFA (i.e., faces for FFA and words for VWFA) evoke similar spatial representations under different attention conditions, and attention primarily modulates the CNR of the signals, a dominant factor of classification accuracy. These findings add to our understandings to the mechanisms of attentional modulation on the functional specificity in the FFA and VWFA.

Acknowledgments

We thank Sheng He, George Northoff, Wen Zhou, and John X. Zhang for valuable comments on the manuscript. This study was supported by NSFC (30900366 to ZY; 31070905 to XW; 31070903 to YJ) and Youth fund of IPCAS (O9CX012001 to ZY).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neuropsychologia.2012.01.026.

References

Baker, C. I., Liu, J., Wald, L. L., Kwong, K. K., Benner, T., & Kanwisher, N. (2007). Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(21), 9087–9092.

Cohen, L., & Dehaene, S. (2004). Specialization within the ventral stream: The case for the visual word form area. *NeuroImage*, *22*(1), 466–476.

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M. A., et al. (2000). The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain*, *123*(2), 291–307.

Cohen, L., Lehéricy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the visual word form area. *Brain*, *125*(5), 1054–1069.

Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, *29*(3), 162–173.

Coutanche, M. N., Thompson-Schill, S. L., & Schultz, R. T. (2011). Multi-voxel pattern analysis of fMRI data predicts clinical symptom severity. *NeuroImage*, *57*(1), 113–123.

Detre, G. J., Polyn, S. M., Moore, C. D., Natu, V. S., Singer, B. D., Cohen, J. D., et al. (2006). The Multi-Voxel Pattern Analysis (MVPA) toolbox. In *Poster presented at the annual meeting of the organization for human brain mapping*. Florence, Italy,

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2008). High-resolution Multi-Voxel Pattern Analysis of category selectivity in the medial temporal lobes. *Hippocampus*, *18*(6), 536–541.

Fischer, J., & Whitney, D. (2009). Attention narrows position tuning of population responses in V1. *Current Biology*, *19*(16), 1356–1361.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D., & Frackowiak, R. S. J. (1995). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, *2*(4), 189–210.

Hasson, U., Levy, I., Behrmann, M., Hendler, T., & Malach, R. (2002). Eccentricity bias as an organizing principle for human high-order object areas. *Neuron*, *34*(3), 479–490.

Haynes, J. D., & Rees, G. (2005). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nature Neruoscience*, *8*(5), 686–691.

Indefrey, P., Kleinschmidt, A., Merboldt, K. D., Kruger, G., Brown, C., Hagoort, P., et al. (1997). Equivalent responses to lexical and nonlexical visual stimuli in occipital cortex: A functional magnetic resonance imaging study. *NeuroImage*, *5*(1), 78–81.

Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neruoscience*, *8*(5), 679–685.

Kamitani, Y., & Tong, F. (2006). Decoding seen and attended motion directions from activity in the human visual cortex. *Current Biology*, *16*(11), 1096–1102.

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S. F., & Baker, C. I. (2009). Circular analysis in systems neuroscience: The dangers of double dipping. *Nature Neruoscience*, *12*(5), 535–540.

Li, X. (1983). The distribution of left and right handedness in Chinese people. *Acta Psychologica Sinica*, *15*, 268–275 (in Chinese).

MacEvoy, S., & Epstein, R. (2009). Decoding the representation of multiple simultaneous objects in human occipitotemporal cortex. *Current Biology*, *19*(11), 943–947.

McAdams, C. J., & Maunsell, J. H. R. (1999). Effects of attention on orientation-tuning functions of single neurons in Macaque cortical area V4. *Journal of Neuroscience*, *19*(1), 431–441.

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, *7*(7), 293–299.

Murray, S. O., & Wojciulik, E. (2004). Attention increases neural selectivity in the human lateral occipital complex. *Nature Neruoscience*, *7*(1), 70–74.

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, *10*(9), 424–430.

Reddy, L., & Kanwisher, N. (2007). Category selectivity in the ventral visual pathway confers and diverted attention. *Current Biology*, *17*(23), 2067–2072.

Reddy, L., Kanwisher, N. G., & VanRullen, R. (2009). Attention and biased competition in multi-voxel object representations. *Proceeding National Academy of Sciences of the United States of America*, *106*(50), 21447–21452.

Reinke, K., Fernandes, M., Schwindt, G., O'Craven, K., & Grady, C. L. (2008). Functional specificity of the visual word form area: General activation for words and symbols but specific network activation for words. *Brain and Language*, *104*(2), 180–189.

Smith, A. T., Kosillo, P., & Williams, A. L. (2011). The confounding effect of response amplitude on MVPA performance measures. *NeuroImage*, *56*(2), 525–530.

Sterzer, P., Haynes, J. D., & Rees, G. (2008). Fine-scale activity patterns in high-level visual areas encode the category of invisible objects. *Journal of Vision*, *8*(15), 1–12.

Tagamets, M. A., Novick, J. M., Chalmers, M. L., & Friedman, R. B. (2000). A parametric approach to orthographic processing in the brain: An fMRI study. *Journal of Cognitive Neuroscience*, *12*(2), 281–297.

Treue, S., & Martinez Trujillo, J. C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*, *399*(6736), 575–579.

Treue, S., & Maunsell, J. H. R. (1999). Effects of attention on the processing of motion in macaque middle temporal and medial superior temporal visual cortical areas. *Journal of Neuroscience*, *19*(17), 7591–7602.

Vul, E., Harris, C., Winkielman, P., & Pashler, H. (2009). Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspectives on Psychological Science*, *4*(3), 274–290.

Wojciulik, E., Kanwisher, N., & Driver, J. (1998). Covert visual attention modulates face-specific activity in the human fusiform gyrus: fMRI study. *Journal of Neurophysiology*, *79*(3), 1574–1578.

Worsley, K. J., Marrett, S., Neelin, P., Vandal, A. C., Friston, K. J., & Evans, A. C. (1996). A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping*, *4*(1), 58–73.