The "visual word form area" is involved in successful memory encoding of both words and faces

Leilei Mei [a,b], Gui Xue [c], Chuansheng Chen [b,*], Feng Xue [a], Mingxia Zhang [a], Qi Dong [a,*]

[a] State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, 100875, China
[b] Department of Psychology and Social Behavior, University of California, Irvine, CA 92697, USA
[c] Department of Psychology, University of Southern California, Los Angeles, CA 90089, USA

ABSTRACT

Previous studies have identified the critical role of the left fusiform cortex in visual word form processing, learning, and memory. However, this so-called visual word form area's (VWFA) other functions are not clear. In this study, we used fMRI and the subsequent memory paradigm to examine whether the putative VWFA was involved in the processing and successful memory encoding of faces as well as words. Twenty-two native Chinese speakers were recruited to memorize the visual forms of faces and Chinese words. Episodic memory for the studied material was tested 3 h after the scan with a recognition test. The fusiform face area (FFA) and the VWFA were functionally defined using separate localizer tasks. We found that, both within and across subjects, stronger activity in the VWFA was associated with better recognition memory of both words and faces. Furthermore, activation in the VWFA did not differ significantly during the encoding of faces and words. Our results revealed the important role of the so-called VWFA in face processing and memory and supported the view that the left mid-fusiform cortex plays a general role in the successful processing and memory of different types of visual objects (i.e., not limited to visual word forms).

© 2010 Elsevier Inc. All rights reserved.

Introduction

Previous studies have revealed the critical role of the left fusiform cortex in reading. First, functional imaging studies have observed strong activation in the left fusiform cortex when comparing words with nonwords, across both alphabetic and logographic writings (Cohen et al., 2000, 2002; Liu et al., 2008). Second, better reading skills are associated with greater involvement of the left fusiform gyrus (Brem et al., 2006; Schlaggar and McCandliss, 2007; Turkeltaub et al., 2003). Third, dyslexics on the other hand showed abnormal fusiform function compared to their normal counterparts (McCrory et al., 2005; Shaywitz et al., 2002; van der Mark et al., 2009). Fourth, evidence from lesion studies has revealed that damages to the left fusiform cortex (Gaillard et al., 2006) or its neural connections to other areas (Cohen et al., 2004; Epelbaum et al., 2008) resulted in impaired letter-by-letter reading.

Using the subsequent memory paradigm (i.e., comparing encoding-related brain activities of subsequently remembered and forgotten items) (Brewer et al., 1998; Wagner et al., 1998) and the training paradigm, recent research further showed the crucial role of the left fusiform gyrus in memory and learning of visual word forms. For instance, several studies using the subsequent memory paradigm have

revealed that strong activation in the fusiform cortex was associated with successful encoding of both familiar words (Otten et al., 2001, 2002; Otten and Rugg, 2001; Wagner et al., 1998) and novel writings (Xue et al., submitted for publication-a). In addition, evidence from artificial language training studies has suggested that the left fusiform is optimal for learning novel visual word forms (Chen et al., 2007; Dong et al., 2008; Xue et al., 2006a). Specifically, it has been found that stronger leftward laterality of the fusiform cortex when initially processing a novel writing (pre-training) was associated with better orthographic learning after two weeks' training (Xue et al., 2006a).

Although existing studies have identified the critical role of the left mid-fusiform in learning to read, it is less clear whether this brain area is specialized for visual word form processing or it performs other cognitive functions. According to the visual word form area (VWFA) perspective (Cohen and Dehaene, 2004; Cohen et al., 2000, 2002), the left fusiform region is specialized for visual word form processing by selectively responding to familiar words. However, other researchers (e.g., Price and Devlin, 2003; Xue et al., 2006b; Xue and Poldrack, 2007) have suggested that the VWFA is not specialized for the processing of familiar visual words because there is evidence that it is also involved in lexical processing (Hillis et al., 2005; Kronbichler et al., 2004), non-word visual objects such as faces, houses, and tools (see Price and Devlin (2003) for a review), and novel writings (Xue et al., 2006b; Xue and Poldrack, 2007).

Although research on the VWFA's involvement in the processing of objects other than visual words is accumulating, it is limited in two

---

* Corresponding authors. Q. Dong is to be contacted at fax: +86 10 58807615. C. Chen, fax: +1 949 824 3002.
E-mail addresses: cschen@uci.edu (C. Chen), dongqi@bnu.edu.cn (Q. Dong).

major aspects. First, these studies typically showed activation in the left mid-fusiform gyrus, but did not actually localize the activation to the VWFA. Direct comparisons between activations by familiar words and those by objects in other categories (e.g., faces) at the VWFA would provide stronger evidence. Second, only perceptual tasks were used in those studies. Thus, it is unknown whether these activations elicited by non-word objects, if they actually fall into the putative VWFA, would carry the same functional properties beyond processing into learning and memory. As mentioned above, the VWFA's activation during the processing of words (familiar or unfamiliar) usually leads to better word learning and memory. However, it is largely unknown whether activation in the same region would result in better memory of non-word objects such as faces. Although many studies have examined the neural correlates of face memory (e.g., Golarai et al., 2007; Kuskowski and Pardo, 1999; Prince et al., 2009; Xue et al., submitted for publication-b), and some have reported activation in the left mid-fusiform region (Prince et al., 2009; Xue et al., submitted for publication-b), no studies have focused on the role of VWFA in memory of faces or directly compared it with the memory of words.

Using the fMRI and the subsequent memory paradigm (Brewer et al., 1998; Wagner et al., 1998), the present study aimed to directly examine the role of the VWFA in the memory of words and faces. By using the subsequent memory paradigm, this study extended previous research by focusing on the involvement of the left fusiform region in the memory (rather than just visual processing) of words and faces. An independent localizer task was used to define the VWFA (Baker et al., 2007) and the fusiform face area (FFA) (Grill-Spector et al., 2004; Kanwisher et al., 1997; McCarthy et al., 1997). To emphasize the encoding of visual forms, an intentional encoding task was used. As shown in previous research (Bernstein et al., 2002; Otten and Rugg, 2001), perceptual and intentional encoding tasks resulted in greater engagement of the posterior regions (e.g., the fusiform cortex) in successful encoding. Subjects were explicitly instructed to memorize the visual forms. To further encourage subjects to focus on visual forms, we added homophones and the same faces from different angles to the materials to be memorized. In this study, two specific hypotheses were tested. First, we expected to replicate previous findings of the involvement of the VWFA in successful encoding of words (Otten et al., 2001, 2002; Otten and Rugg, 2001; Wagner et al., 1998). Second, we expected that the so-called VWFA would be involved in successful encoding and memory of faces. We directly compared the activation patterns and subsequent memory effects of faces and words in the VWFA.

Methods

Subjects

Twenty-two native Chinese speakers (half males; mean age = 22.8 ± 2.8 years old, with a range from 19 to 30 years) participated in this study. All subjects had normal or corrected-to-normal vision and were strongly right-handed as judged by Snyder and Harris's handedness inventory (Snyder and Harris, 1993). None of them had a previous history of neurological or psychiatric disease. Informed written consent was obtained from the subjects before the experiment. This study was approved by the IRB of the National Key Laboratory of Cognitive Neuroscience and Learning at Beijing Normal University.

Materials

Four types of stimuli, including faces, Chinese words, common objects, and scrambled images of objects, were used in the localizer tasks. Each type contained 40 items. Faces and objects were taken by the same digital camera. The subsequent memory task consisted of

132 Chinese words and 132 famous faces that were neutral in emotion expressions. Famous faces were used so their familiarity to the subjects would be similar to that of familiar words. Each type of materials was further divided into two matched groups, one for the encoding task and the other as foils in the subsequent memory task. All stimuli were presented in gray-scale and 227 × 283 pixels in size.

All Chinese words were medium- to high-frequency words (higher than 25 per million according to the Chinese word frequency dictionary) (Wang and Chang, 1985), with 4–12 strokes, and 2–3 units according to the definition by Chen et al. (1996). Visual complexity (i.e., number of strokes and units) and word frequency was strictly matched across the study words, the foils, and words used in the localizer task.

The famous faces were obtained from the internet and normalized to the same resolution, brightness, and size. These stimuli were evaluated by 11 research assistants in the laboratory before experiment to ensure they were highly familiar to Chinese subjects (i.e., no items scored less than 5 on a 6-point scale with 1 representing "never seen it before" and 6 representing "very familiar"). Familiarity level and gender of the faces were matched across the study faces and the foils.

fMRI task

The fMRI task began with a localizer scan while the subject was passively viewing the four types of stimuli (faces, Chinese words, common objects, and scrambled images of objects). The 40 images of each type of materials were repeated once in the scan. The whole scan consisted of 16 consecutive 20 s epochs (4 for each type of materials), which were separated by 14 s fixation periods. Each image was presented for 750 ms, followed by a 250 ms blank interval. To ensure that subjects were awake and attentive, they were instructed to press a key whenever they noticed an image with white frame. This happened twice per epoch. The localizer scan lasted for 9 m 42 s.

After the localizer scan, participants were scanned while being asked to intentionally encode faces and words. A mixed design was used for the encoding scan, in which 6 blocks of faces interleaved with 6 blocks of words. The order of the blocks was counterbalanced across subjects. Each block included 11 stimuli and 2 successively presented fillers (homophones in the word block and different angles of the faces in the face block). During scanning, subjects were told about the fillers and were explicitly instructed to memorize the visual forms of faces or words. Subjects were further told that homophones and faces of different angles would be added in the subsequent memory test to encourage them to focus on the visual forms. In the actual test, however, no fillers were added to simplify the design. For each trial, the stimulus was presented for 2 s, followed by a blank that randomly varied from 1 to 5 s (mean = 2 s) to improve design efficiency. To avoid the primacy and recency effects, two other fillers were separately placed at the beginning and the end of the sequence. In total, the scan included 160 trials and lasted for 10 min 34 s.

Post-scan behavioral test

Three hours after scanning, a recognition test was administered to assess subjects' memory performance. Fillers in fMRI scan were excluded in this test. Consequently, a total of 132 faces and 132 words were used. For both types of the stimuli, half of them were those used in the fMRI encoding task, whereas the other half had not been seen by the subjects during the fMRI scan. All stimuli were randomly intermixed. For each stimulus, the subjects had to decide whether they had seen it during the scan on a 6-point confidence scale, ranging from 1 (definitely new) to 6 (definitely old). Each stimulus would stay on the screen until the subjects responded. The next item would appear after a 1 s blank.

MRI data acquisition

Data were acquired with a 3.0 T Siemens MRI scanner at the MRI Center of Beijing Normal University. Single-shot T2*-weighted gradient-echo, EPI sequence was used for functional imaging acquisition with the following parameters: $TR/TE/\theta = 2000$ ms/30 ms/90°, $FOV = 200 \times 200$ mm, matrix $= 64 \times 64$, and slice thickness $= 4$ mm. Thirty contiguous axial slices parallel to AC–PC plane were obtained to cover the whole cerebrum and partial cerebellum. Anatomical MRI was acquired using a T1-weighted, three-dimensional, gradient-echo pulse-sequence. Parameters for this sequence were: $TR/TE/\theta = 2530$ ms/3.39 ms/7°, $FOV = 256 \times 256$ mm, matrix $= 192 \times 256$, and slice thickness $= 1.33$ mm. One hundred and twenty-eight sagittal slices were acquired to provide a high-resolution structural image of the whole brain.

Image preprocessing and statistical analysis

Initial analysis was carried out using tools from the FMRIB's software library (www.fmrib.ox.ac.uk/fsl) version 4.1.2. The first 3 volumes in each time series were automatically discarded by the scanner to allow for T1 equilibrium effects. The remaining images were then realigned to compensate for small head movements (Jenkinson and Smith, 2001). Translational movement parameters never exceeded 1 voxel in any direction for any subject or session. Data were spatially smoothed using a 5-mm full-width-half-maximum Gaussian kernel. The smoothed data were then filtered in the temporal domain using a nonlinear high-pass filter with a 100-s cutoff. A 2-step registration procedure was used whereby EPI images were first registered to the MPRAGE structural image, and then into standard (Montreal Neurological Institute [MNI]) space, using affine transformations with FLIRT (Jenkinson and Smith, 2001) to the avg152 T1 MNI template.

At the first level, the data were modeled by two general linear models within the FILM module of FSL. The first model was used to compute the subsequent memory effect, in which four events were modeled for each type of materials (words and faces): Definitely remembered (confidence rating of 6), Possibly remembered (confidence ratings of 4, 5), Forgotten (confidence ratings of 1, 2, 3), and nuisance events (fillers). The second model was used to compute overall activations of faces and words, in which three events were modeled: faces, words, and nuisance events (fillers). This was separately modeled because of the uneven distribution of events across the three different memory conditions. In both models, events were modeled at the time of the stimulus presentation. These event onsets and their durations were convolved with canonical hemodynamic response function (double-gamma) to generate the regressors used in the general linear models. Temporal derivatives and the 6 motion parameters were included as covariates of no interest to improve statistical sensitivity. Null events were not explicitly modeled, and therefore constituted an implicit baseline. For each subject, 4 contrast images were computed in the first model for each type of materials. The four contrasts were (a) Definitely remembered minus null events, (b) Possibly remembered minus null events, (c) Forgotten events minus null events, and (d) Definitely remembered minus Forgotten events. In addition, we examined the memory-success-by-material interaction by using the contrast $[1 -1 -1 1]$ of Definitely remembered words, Forgotten words, Definitely remembered faces, and Forgotten faces. For the second model, 2 contrast images (faces minus null events; words minus null events) were computed.

Corresponding to the first-level analysis, two second-level models were separately constructed. The first model was used to average the subsequent memory effect across subjects for both words and faces. In the second model, individuals' discriminability indices (d') were added as covariates to examine the relationship between encoding-related brain activities and individuals' subsequent memory performance for both words and faces. The d' values were computed for all responses regardless of confidence using the following formula: $d' = Z_{(hit\ rate)} - Z_{(false\ alarm)}$. In both second-level models, group activations were computed using mixed-effects models (treating subjects as a random effect) with FLAME stage 1 with automatic outlier detection (Beckmann et al., 2003; Woolrich, 2008; Woolrich et al., 2004). Unless otherwise indicated, group images were thresholded with a height threshold of $z > 2.3$ and a cluster probability, $P < 0.05$, corrected for whole-brain multiple comparisons using the Gaussian random field theory.

Regions of interest analysis

Regions of interest (ROIs) were functionally defined using data from the localizer scan. The VWFA was defined as the contiguous voxels in the left occipitotemporal cortex surviving the words > objects contrast. The FFA was defined as the contiguous voxels in the middle fusiform gyrus surviving the faces > objects contrast. For both ROIs, a stringent threshold ($Z > 3.7$, cluster probability < 0.05, corrected for whole-brain multiple comparisons using the Gaussian random field theory) was used to select highly consistent voxels across subjects. The VWFA contained 166 voxels that centered at $x = -46$, $y = -62$, $z = -16$ (MNI, $Z = 5.20$). The FFA contained 184 voxels that were located only in the right fusiform (MNI: 44, -50, -26, $Z = 5.10$). These two locations were close to those reported in previous studies: the VWFA (Bolger et al., 2005; Cohen and Dehaene, 2004; Cohen et al., 2002; Vigneau et al., 2005) and the FFA (Gauthier et al., 1999; Grill-Spector et al., 2004; Kanwisher et al., 1997; McCarthy et al., 1997). For the ROI analyses, the mean effect size (i.e. contrast of parameter estimate, COPE) was extracted for each subject and each contrast and then used for further statistical analysis.

Results

Behavioral results

For the localizer tasks, subjects correctly responded to $97 \pm 4\%$ of the 32 items with white frame. This indicates that subjects were attentive to the stimuli during the localizer scan.

For the memory results, we first divided subjects' responses into four types: Definitely remembered (confidence rating of 6), Possibly remembered (confidence ratings of 4, 5), Possibly forgotten (confidence ratings of 2, 3), and Definitely forgotten (confidence rating of 1). The ratios of the four types of responses for both types of materials (words and faces) are shown in Table 1. The discrimination index (Pr) (i.e., probability of hits minus probability of false alarms) was used to compute the accuracy of recognition (Otten et al., 2001; Snodgrass and Corwin, 1988). For the Definitely remembered responses, the Prs were significantly greater than chance (i.e., zero) for both words (mean: 0.26, $t(21) = 6.79$, $p < 0.001$) and faces (mean: 0.61, $t(21) = 15.56$, $p < 0.001$), although recognition performance was better for faces ($t(21) = 8.39$, $p < 0.001$). For the Possibly remembered responses, however, the Prs were not reliably greater than zero for both types of materials (-0.03 for words and 0.01 for faces).

Table 1

Recognition memory performance (% correct) for old (trained) and new (foils) items by type of materials (words and faces).

Materials		Definitely old	Maybe old	Maybe new	Definitely new
Words	Old	0.44 (0.21)	0.29 (0.18)	0.20 (0.12)	0.07 (0.07)
	New	0.18 (0.15)	0.32 (0.16)	0.34 (0.16)	0.16 (0.17)
Faces	Old	0.69 (0.15)	0.06 (0.06)	0.10 (0.11)	0.15 (0.10)
	New	0.08 (0.13)	0.05 (0.05)	0.15 (0.18)	0.72 (0.25)

Note. Numbers in parentheses are standard deviations.

Table 2
Brain regions demonstrating the subsequent memory effect.

Brain regions	Words				Faces			
	x	y	z	Z	x	y	z	Z
Left hippocampus					−34	−22	−16	3.20
Left amygdala					−20	−8	−16	3.37
Right parahippocampal gyrus					22	−18	−18	4.05
Right amygdala					26	−2	−18	3.33
Left fusiform gyrus	−48	−64	−14	3.74	−44	−54	−14	3.56
					−42	−72	−12	3.33
Right fusiform gyrus					40	−60	−16	3.75
					38	−78	−10	3.23
Left inferior occipital gyrus	−36	−88	−10	2.95	−28	−94	−16	3.45
	−26	−68	32	4.09	−32	−96	2	3.50
Right inferior occipital gyrus					26	−92	−8	3.73
Left superior occipital gyrus	−24	−70	44	3.91	−34	−88	20	3.07
Right superior occipital gyrus	26	−74	42	3.52	36	−84	16	3.66

Based on these results and previous findings that the subsequent memory effect was mainly found for the "definitely remembered" items (Brewer et al., 1998; Otten et al., 2001; Wagner et al., 1998), we contrasted Definitely remembered items with Forgotten items to maximize the signal-to-noise ratio in the following analysis of the subsequent memory effect.

The left fusiform supported successful encoding of words

To examine neural correlates of successful encoding of words, we first examined the subsequent memory effect by comparing brain activations of the Definitely remembered and the Forgotten items. The brain regions demonstrating the subsequent memory effect are listed in Table 2. As expected, the Definitely remembered items showed more activation than the Forgotten ones in the left fusiform gyrus (MNI: −48, −64, −14, $Z = 3.74$). More activation was also found in the left inferior occipital cortex and bilateral superior occipital cortex (see Fig. 1A).
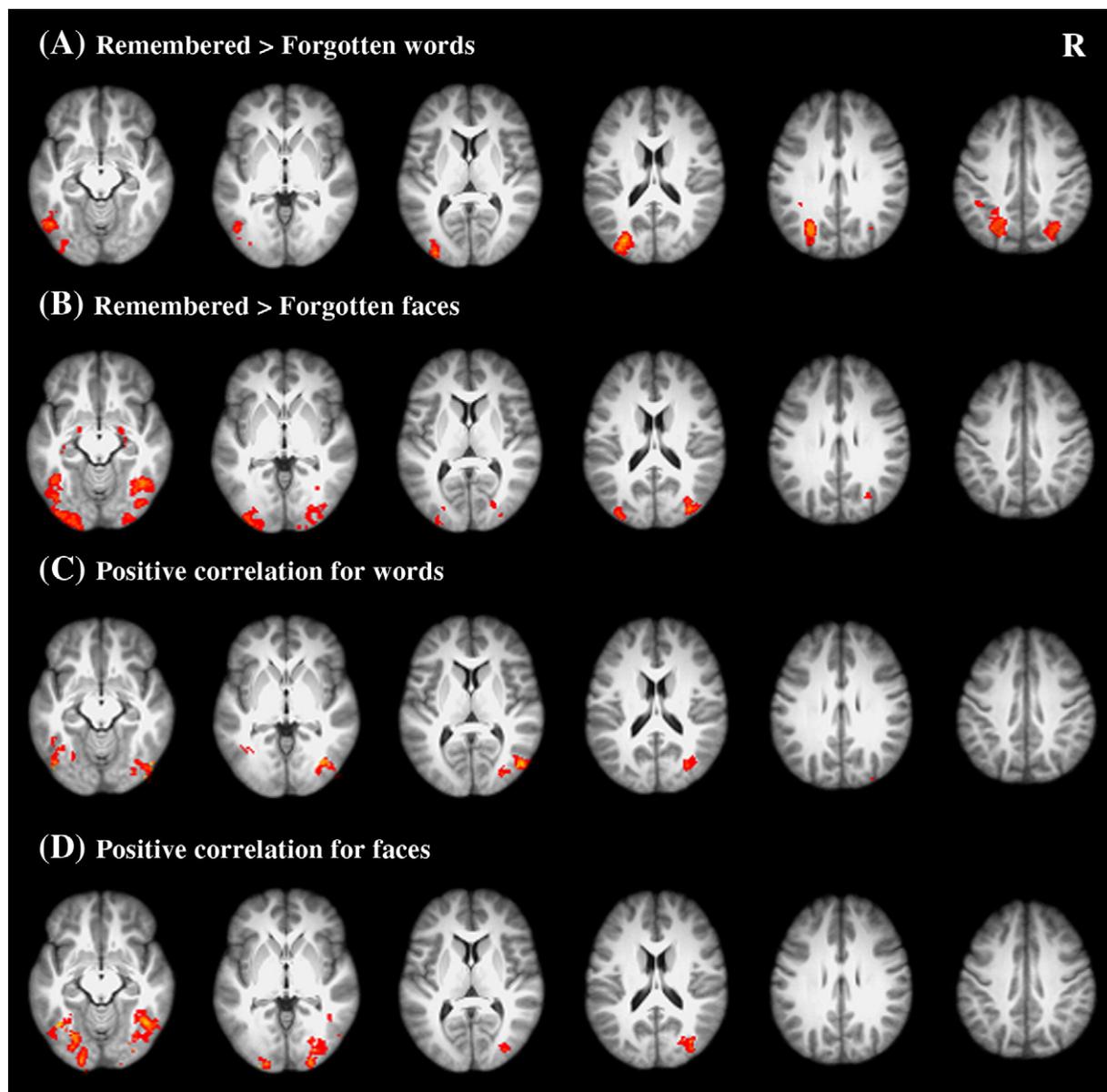

Fig. 1. Neural correlates of successful encoding of both words and faces. (A–B) Brain regions demonstrating the subsequent memory effect for words and faces, respectively. (C–D) Brain regions showing significant positive correlations with recognition memory of words and faces, respectively. R = right.

Table 3
Brain regions showing positive correlations with recognition performance (d').

Brain regions	Word				Face			
	x	y	z	Z	x	y	z	Z
Left fusiform gyrus	−42	−60	−10	3.46	−40	−60	−20	4.94
	−32	−64	−16	3.79	−30	−72	−18	5.20
Right fusiform gyrus					40	−56	−14	4.44
	32	−78	−14	4.35	44	−72	−16	3.92
Left inferior occipital gyrus	−48	−74	−14	4.70	−24	−98	−4	5.65
Right inferior occipital gyrus	50	−72	10	4.23				
	36	−72	−2	4.36	26	−92	−2	3.73
Right superior occipital gyrus					34	−80	22	4.37

We also performed whole-brain correlational analysis, in which brain activity during encoding was correlated with individuals' recognition performance (i.e., d' values) (see Methods). Consistent with the results of the subsequent memory effect analysis, positive correlations were found in the left fusiform (extending from the medial portion MNI: −32, −60, −16, $Z = 3.79$ to the lateral portion MNI: −42, −60, −10, $Z = 3.46$, see Table 3). It is worth noting that the two regions identified by different analyses were located closely to each other. Other regions showing positive correlations included bilaterial inferior occipital cortex and the posterior portion of the right fusiform cortex (see Fig. 1C).

Bilateral fusiform supported successful encoding of faces

We also performed the same two analyses for faces. Both analyses showed that successful encoding of faces was associated with more activation in bilateral fusiform cortex, extending to the posterior portions as well as the inferior and superior occipital cortex (Figs. 1B and D). It should be noted that regions in the left fusiform cortex associated with successful encoding of both words and faces were close to the VWFA localized in this study. Meanwhile, regions in the right fusiform cortex associated with successful encoding of faces were close to the FFA (see Tables 2 and 3 for specific coordinates).

Other brain regions, including bilateral medial temporal lobe (MTL) that extended to the amygdala, were also found to show the subsequent memory effect. These regions might be recruited because
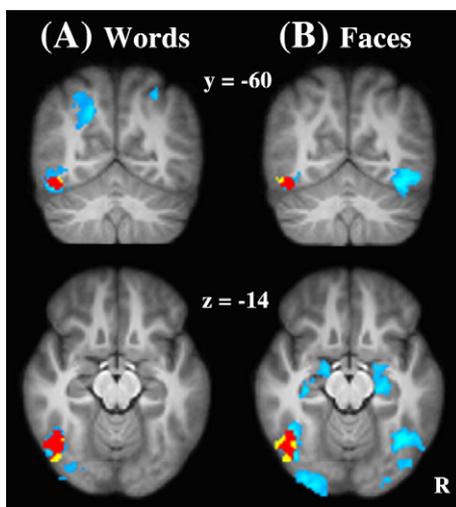

Fig. 2. The overlaps between the VWFA and the left fusiform activation. The VWFA (yellow), brain regions showing the subsequent memory effect (light blue), and their overlaps (red) are overlaid on the coronal (top) and sagittal (bottom) slices of the group mean structural image for words (A) and faces (B). R = right.
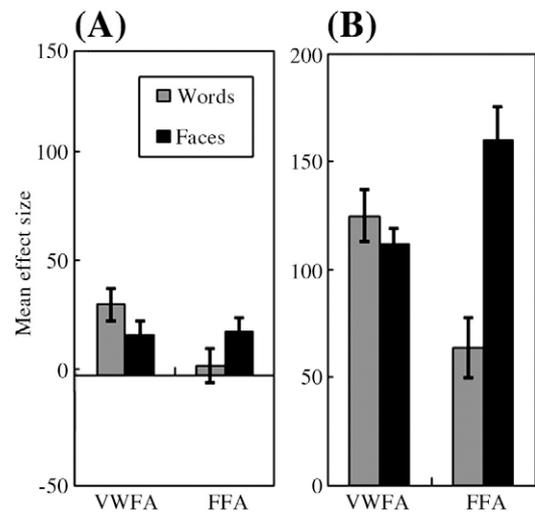

Fig. 3. The mean effect size in the VWFA and FFA for words and faces. (A) The subsequent memory effect; (B) overall activation during encoding. Error bars represent standard error of the mean.

of subjects' likely emotional reactions to these familiar people (LaBar and Cabeza, 2006).

The so-called VWFA was responsible for successful encoding of both words and faces

Having revealed the important association between the left fusiform cortex and successful encoding of both words and faces, we then examined whether this association occurred in the VWFA by performing the following four analyses. First, we overlaid group maps of Definitely remembered versus Forgotten words and faces on the VWFA mask (see Fig. 2). It was found that regions in the left fusiform that showed the subsequent memory effect for both words and faces largely overlapped with the VWFA.

Second, we extracted the mean effect size of the subsequent memory effect for words and faces from the VWFA and the FFA. As shown in Fig. 3A, the VWFA demonstrated the subsequent memory effect for both words and faces, whereas the FFA showed the subsequent memory effect only for faces (region-by-material interaction: $F(1,21) = 24.11$, $p < 0.001$). Although the subsequent memory effect was slightly greater for words than faces in the VWFA and for faces than words in the FFA, they were not statistically significant (VWFA: $F(1,21) = 2.84$, n.s.; FFA: $F(1,21) = 2.77$, n.s.). We further compared the overall activation of words with that of faces and observed no significant difference in the VWFA as well ($F(1,21) = 1.63$, n.s.) (see Fig. 3B).

Third, we examined the memory success (Definitely remembered vs. Forgotten) × material (words vs. faces) interaction throughout the whole brain using the VWFA as the prethreshold mask. No voxels within the VWFA demonstrated a significant effect even with a liberal threshold ($Z = 2.3$, uncorrected).

Finally, we correlated brain activity in the VWFA and the FFA during encoding with recognition of words and faces (see Fig. 4). Correlational analyses showed that activity in the VWFA was positively correlated with memory performance for both words and faces (words: $r = 0.52$, $p < 0.05$; faces: $r = 0.54$, $p < 0.01$), whereas that of the FFA was not significantly correlated with memory of either type of materials (words: $r = 0.12$, n.s.; faces: $r = 0.31$, n.s.).

Discussion

This study was designed to examine the role of the so-called visual word form area (VWFA) (Cohen and Dehaene, 2004; Cohen et al., 2002) in face processing and memory. We measured brain activity
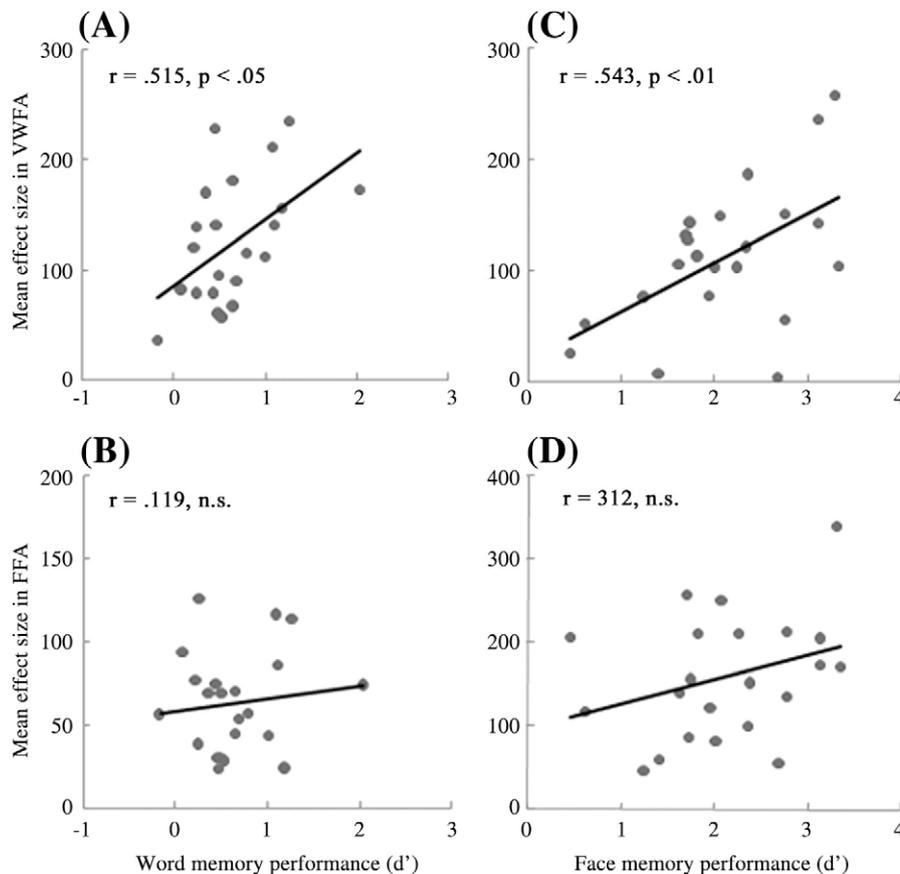
Fig. 4. Activity in the VWFA during encoding predicted recognition memory (d'). The four scatter plots show correlations for VWFA activity and word memory (A), FFA activity and word memory (B), VWFA activity and face memory (C), and FFA activity and face memory (D).

with fMRI during an intentional encoding task and collected data on memory with a post-scan recognition test. Analyses based on both the within-subject subsequent memory effect and cross-subject correlations revealed that activity in the putative VWFA during encoding predicted recognition memory for both words and faces. More importantly, the subsequent memory effect did not differ significantly by types of materials (i.e., words and faces). These results not only argue against the VWFA perspective, which considers the left mid-fusiform gyrus as specialized for the processing of visual word forms, but also show the important role of this region for face processing and especially face memory (it was at least as important as the right fusiform for the latter).

The critical role of the VWFA in successful encoding of both words and faces

Previous studies have identified the critical involvement of the left fusiform cortex in visual word form processing (e.g., Cohen and Dehaene, 2004; Cohen et al., 2002), encoding (Otten et al., 2001, 2002; Otten and Rugg, 2001; Wagner et al., 1998), and learning (Chen et al., 2007; Dong et al., 2008; Xue et al., 2006a). Moreover, a recent study successfully increased orthographic memory by reducing neural repetition suppression in the left fusiform cortex with prolonged repetition lags (Xue et al., submitted for publication-a). That study provided a causal link between the left fusiform's activity and orthographic learning. Consistent with these results, the present study further confirmed the critical involvement of the putative VWFA in successful encoding of visual word forms using both within-subject (i.e., the subsequent memory effect) and cross-subject (i.e., brain–behavior correlation) analyses. Taken together, it is clear that

the VWFA plays a fundamental role in visual word form processing, encoding, and learning.

Although the VWFA is critically involved in reading familiar words, its function is not as exclusive or specialized as the VWFA hypothesis would claim (Cohen and Dehaene, 2004; Cohen et al., 2000, 2002). As mentioned earlier in the introduction, there is already accumulating evidence that the VWFA is involved in processing and encoding non-word visual objects such as faces, scenes, and tools (Bernstein et al., 2002; Garoff et al., 2005; Kirchhoff et al., 2000; Prince et al., 2009; Xue et al., submitted for publication-b; see Price and Devlin, 2003, for a review). Our study provided more evidence against the VWFA perspective by directly comparing the role of the so-called VWFA (precisely localized via passive viewing tasks) in face vis-à-vis word processing and memory. Specifically, we found that strong activation in the VWFA was associated with successful encoding of faces as well as words. More importantly, the present study found that activation in the VWFA did not differ significantly between the successful encoding of words and that of faces. In sum, these results suggest that the VWFA was not specialized exclusively for the processing of visual word forms.

Implications on the left fusiform's function

Similar activation in the left fusiform cortex for successful encoding of words and faces was probably because of our explicit instruction on feature/part processing strategies. It has been proposed that the left and right hemispheres might be specialized for processing high- versus low-spatial-frequency information (Kitterle and Selig, 1991), part versus whole (Robertson and Lamb, 1991), and feature versus holistic information (Grill-Spector, 2001), respectively (see Hellige et al., forthcoming for a review). Therefore, reliance

on the left fusiform cortex might be caused by great involvement of feature/part information processing for face encoding in this study.

Consistent with this view, previous studies have found that inversion eliminated or reduced the right hemisphere advantage for faces because of the whole representation of faces was disrupted (Hillger and Koenig, 1991; Leehey et al., 1978). More interestingly, tachistoscopic studies have observed left-hemisphere advantage for feature processing of faces induced by task manipulation (Hillger and Koenig, 1991). Finally, one neuroimaging study (Rossion et al., 2000) revealed a right-fusiform advantage for the processing of faces as a whole and a left-fusiform superiority for the processing of face features.

Future research should investigate the hemispheric specialization perspective (Hellige et al., forthcoming) by comparing fusiform activation in tasks that encourage feature/part processing with tasks that encourage holistic/whole processing using both left-hemisphere-superiority (e.g., words) (Cohen and Dehaene, 2004) and right-hemisphere-superiority (e.g., faces) materials (Willems et al., 2009). According to the hemispheric specialization perspective, the left fusiform gyrus would be involved in processing many different visual objects and its level of involvement would vary depending on processing strategies (e.g., emphasizing features or parts versus holistic information or wholes) regardless of material.

Future research can also examine the hemispheric specialization perspective using a language learning paradigm. Previous studies have revealed critical associations between cerebral asymmetry in the fusiform cortex and orthographic learning (Chen et al., 2007; Xue et al., 2006a). Specifically, individuals with greater leftward fusiform activation learned novel visual word forms better after orthographic learning. From the hemispheric specialization perspective, this association might reflect the fact that good learners enhanced the relative involvement of the left fusiform cortex by using more feature/part processing strategies when learning the novel visual forms, and consequently improved learning outcomes. It should be interesting to investigate this possibility by comparing the visual word form learning that encourages feature/part processing strategies with learning that emphasizes holistic/whole processing strategies.

Acknowledgments

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. PNAS 104, 9087–9092.

Beckmann, C.F., Jenkinson, M., Smith, S.M., 2003. General multilevel linear modeling for group analysis in FMRI. NeuroImage 20, 1052–1063.

Bernstein, L.J., Beig, S., Siegenthaler, A.L., Grady, C.L., 2002. The effect of encoding strategy on the neural correlates of memory for faces. Neuropsychologia 40, 86–98.

Bolger, D.J., Perfetti, C.A., Schneider, W., 2005. Cross-cultural effect on the brain revisited: universal structures plus writing system variation. Hum. Brain Mapp. 25, 92–104.

Brem, S., Bucher, K., Halder, P., Summers, P., Dietrich, T., Martin, E., Brandeis, D., 2006. Evidence for developmental changes in the visual word processing network beyond adolescence. NeuroImage 29, 822–837.

Brewer, J.B., Zhao, Z., Desmond, J.E., Glover, G.H., Gabrieli, J.D., 1998. Making memories: brain activity that predicts how well visual experience will be remembered. Science 281, 1185–1187.

Chen, Y.P., Allport, D.A., Marshall, J.C., 1996. What are the functional orthographic units in Chinese word recognition: the stroke or the stroke pattern? Q. J. Exp. Psychol. 49, 1024–1043.

Chen, C., Xue, G., Dong, Q., Jin, Z., Li, T., Xue, F., Zhao, L., Guo, Y., 2007. Sex determines the neurofunctional predictors of visual word learning. Neuropsychologia 45, 741–747.

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

Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M.-A., Michel, F., 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, 291–307.

Cohen, L., Lehericy, 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, 1054–1069.

Cohen, L., Henry, C., Dehaene, S., Martinaud, O., Lehericy, S., Lemer, C., Ferrieux, S., 2004. The pathophysiology of letter-by-letter reading. Neuropsychologia 42, 1768–1780.

Dong, Q., Mei, L., Xue, G., Chen, C., Li, T., Xue, F., Huang, S., 2008. Sex-dependent neurofunctional predictors of long-term maintenance of visual word learning. Neurosci. Lett. 430, 87–91.

Epelbaum, S., Pinel, P., Gaillard, R., Delmaire, C., Perrin, M., Dupont, S., Dehaene, S., Cohen, L., 2008. Pure alexia as a disconnection syndrome: new diffusion imaging evidence for an old concept. Cortex 44, 962–974.

Gaillard, R., Naccache, L., Pinel, P., Clemenceau, S., Volle, E., Hasboun, D., Dupont, S., Baulac, M., Dehaene, S., Adam, C., Cohen, L., 2006. Direct intracranial, FMRI, and lesion evidence for the causal role of left inferotemporal cortex in reading. Neuron 50, 191–204.

Garoff, R.J., Slotnick, S.D., Schacter, D.L., 2005. The neural origins of specific and general memory: the role of the fusiform cortex. Neuropsychologia 43, 847–859.

Gauthier, I., Tarr, M.J., Anderson, A.W., Skudlarski, P., Gore, J.C., 1999. Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. Nat. Neurosci. 2, 568–573.

Golarai, G., Ghahremani, D.G., Whitfield-Gabrieli, S., Reiss, A., Eberhardt, J.L., Gabrieli, J.D.E., Grill-Spector, K., 2007. Differential development of high-level visual cortex correlates with category-specific recognition memory. Nat. Neurosci. 10, 512–522.

Grill-Spector, K., 2001. Semantic versus perceptual priming in fusiform cortex. Trends Cogn. Sci. 5, 227–228.

Grill-Spector, K., Knouf, N., Kanwisher, N., 2004. The fusiform face area subserves face perception, not generic within-category identification. Nat. Neurosci. 7, 555–562.

Hellige, J.B., Laeng, B., Michimata, C., forthcoming. Processing Asymmetries in the Visual System. In: Hugdahl, K., Westerhausen, R. (Eds.) The two halves of the brain: information processing in the cerebral hemispheres. MA:MIT Press, Cambridge.

Hillger, L.A., Koenig, O., 1991. Separable mechanisms in face processing: evidence from hemispheric specialization. J. Cogn. Neurosci. 3, 42–58.

Hillis, A.E., Newhart, M., Heidler, J., Barker, P., Herskovits, E., Degaonkar, M., 2005. The roles of the "visual word form area" in reading. NeuroImage 24, 548–559.

Jenkinson, M., Smith, S., 2001. A global optimisation method for robust affine registration of brain images. Med. Image Anal. 5, 143–156.

Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J. Neurosci. 17, 4302–4311.

Kirchhoff, B.A., Wagner, A.D., Maril, A., Stern, C.E., 2000. Prefrontal–temporal circuitry for episodic encoding and subsequent memory. J. Neurosci. 20, 6173–6180.

Kitterle, F.L., Selig, L.M., 1991. Visual field effects in the discrimination of sine-wave gratings. Percept. Psychophys. 50, 15–18.

Kronbichler, M., Hutzler, F., Wimmer, H., Mair, A., Staffen, W., Ladurner, G., 2004. The visual word form area and the frequency with which words are encountered: evidence from a parametric fMRI study. NeuroImage 21, 946–953.

Kuskowski, M.A., Pardo, J.V., 1999. The role of the fusiform gyrus in successful encoding of face stimuli. NeuroImage 9, 599–610.

LaBar, K.S., Cabeza, R., 2006. Cognitive neuroscience of emotional memory. Nat. Rev. Neurosci. 7, 54–64.

Leehey, S., Carey, S., Diamond, R., Cahn, A., 1978. Upright and inverted faces: the right hemisphere knows the difference. Cortex 14, 411–419.

Liu, C., Zhang, W.-T., Tang, Y.-Y., Mai, X.-Q., Chen, H.-C., Tardif, T., Luo, Y.-J., 2008. The visual word form area: evidence from an fMRI study of implicit processing of Chinese characters. NeuroImage 40, 1350–1361.

McCarthy, G., Puce, A., Gore, J.C., Allison, T., 1997. Face-specific processing in the human fusiform gyrus. J. Cogn. Neurosci. 9, 605–610.

McCrory, E.J., Mechelli, A., Frith, U., Price, C.J., 2005. More than words: a common neural basis for reading and naming deficits in developmental dyslexia? Brain 128, 261–267.

Otten, L.J., Rugg, M.D., 2001. Task-dependency of the neural correlates of episodic encoding as measured by fMRI. Cereb. Cortex 11, 1150–1160.

Otten, L.J., Henson, R.N.A., Rugg, M.D., 2001. Depth of processing effects on neural correlates of memory encoding: relationship between findings from across- and within-task comparisons. Brain 124, 399–412.

Otten, L.J., Henson, R.N.A., Rugg, M.D., 2002. State-related and item-related neural correlates of successful memory encoding. Nat. Neurosci. 5, 1339–1344.

Price, C.J., Devlin, J.T., 2003. The myth of the visual word form area. NeuroImage 19, 473–481.

Prince, S.E., Dennis, N.A., Cabeza, R., 2009. Encoding and retrieving faces and places: distinguishing process- and stimulus-specific differences in brain activity. Neuropsychologia 47, 2282–2289.

Robertson, L.C., Lamb, M.R., 1991. Neuropsychological contributions to theories of part/whole organization. Cogn. Psychol. 23, 299–330.

Rossion, B., Dricot, L., Devolder, A., Bodart, J.-M., Crommelinck, M., Gelder, B.d., Zoontjes, R., 2000. Hemispheric asymmetries for whole-based and part-based face processing in the human fusiform gyrus. J. Cogn. Neurosci. 12, 793–802.

Schlaggar, B.L., McCandliss, B.D., 2007. Development of neural systems for reading. Annu. Rev. Neurosci. 30, 475–503.

Shaywitz, B.A., Shaywitz, S.E., Pugh, K.R., Mencl, W.E., Fulbright, R.K., Skudlarski, P., Constable, R.T., Marchione, K.E., Fletcher, J.M., Lyon, G.R., Gore, J.C., 2002. Disruption of posterior brain systems for reading in children with developmental dyslexia. Biol. Psychiatry 52, 101–110.

Snodgrass, J.G., Corwin, J., 1988. Pragmatics of measuring recognition memory: applications to dementia and amnesia. J. Exp. Psychol. 117, 34–50.

Snyder, P.J., Harris, L.J., 1993. Handedness, sex, and familial sinistrality effects on spatial tasks. Cortex 29, 115–134.

Turkeltaub, P.E., Gareau, L., Flowers, D.L., Zeffiro, T.A., Eden, G.F., 2003. Development of neural mechanisms for reading. Nat. Neurosci. 6, 767–773.

van der Mark, S., Bucher, K., Maurer, U., Schulz, E., Brem, S., Buckelmuller, J., Kronbichler, M., Loenneker, T., Klaver, P., Martin, E., Brandeis, D., 2009. Children with dyslexia lack multiple specializations along the visual word-form (VWF) system. NeuroImage 47, 1940–1949.

Vigneau, M., Jobard, G., Mazoyer, B., Tzourio-Mazoyer, N., 2005. Word and non-word reading: what role for the Visual Word Form Area? NeuroImage 27, 694–705.

Wagner, A.D., Schacter, D.L., Rotte, M., Koutstaal, W., Maril, A., Dale, A.M., Rosen, B.R., Buckner, R.L., 1998. Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. Science 281, 1188–1191.

Wang, H., Chang, R.B., 1985. Modern Chinese Frequency Dictionary. Beijing Language University Press, Beijing.

Willems, R.M., Peelen, M.V., Hagoort, P., 2009. Cerebral lateralization of face-selective and body-selective visual areas depends on handedness. Cereb. Cortex bhp234.

Woolrich, M., 2008. Robust group analysis using outlier inference. NeuroImage 41, 286–301.

Woolrich, M.W., Behrens, T.E.J., Beckmann, C.F., Jenkinson, M., Smith, S.M., 2004. Multilevel linear modelling for FMRI group analysis using Bayesian inference. NeuroImage 21, 1732–1747.

Xue, G., Poldrack, R.A., 2007. The neural substrates of visual perceptual learning of words: implications for the visual word form area hypothesis. J. Cogn. Neurosci. 19, 1643–1655.

Xue, G., Chen, C., Jin, Z., Dong, Q., 2006a. Cerebral asymmetry in the fusiform areas predicted the efficiency of learning a new writing system. J. Cogn. Neurosci. 18, 923–931.

Xue, G., Chen, C., Jin, Z., Dong, Q., 2006b. Language experience shapes fusiform activation when processing a logographic artificial language: an fMRI training study. NeuroImage 31, 1315–1326.

Xue, G., Mei, L., Chen, C., Lu, Z., Dong, Q., Poldrack, R., submitted for publication-a. Facilitating orthographic learning by reducing the neural repetition suppression in the left fusiform cortex. Hum. Brain Mapp.

Xue, G., Mei, L., Chen, C., Lu, Z., Dong, Q., Poldrack, R., submitted for publication-b. Spaced learning enhances subsequent recognition memory by reducing neural repetition suppression. J. Cogn. Neurosci.