A study on neural mechanism of face processing based on fMRI

Jiangang Liu a, Jie Tian b,c,*, Kang Lee d, Jun Li c

a School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China
b Key Laboratory of Complex Systems and Intelligence Science, Institute of Automation, Chinese Academy of Sciences, Beijing 100080, China
c Life Science Research Center, Xidian University, Xi'an 710071, China
d Institute of Child Study, University of Toronto, Canada

Received 7 June 2007; received in revised form 25 June 2007; accepted 26 June 2007

Abstract

Recently, there were debates about the specificity of lateral middle fusiform in face processing. The debates focused on whether these areas were specialized in face processing or involved in processing of visual expertise and categorization at individual level. The present study aims to investigate the neural mechanism of face processing, using Chinese characters as comparison stimuli. Chinese characters are greatly similar to faces on a variety of dimensions, among which the most significant one is that both faces and Chinese characters not only are extremely familiar to literate Chinese adults but also are processed at individual level. In the present study, faces and Chinese characters activated bilateral middle fusiform with great correlation. Greater activities were observed in the right fusiform face area (FFA) for faces than for Chinese characters. These results demonstrate that FFA is specialized in face processing per se rather than the processing of visual expertise and categorization at individual level.

© 2007 National Natural Science Foundation of China and Chinese Academy of Sciences. Published by Elsevier Limited and Science in China Press. All rights reserved.

Keywords: Face processing; Chinese characters; Fusiform gyrus; fMRI

1. Introduction

Face perception, as a crucial method by which one recognizes other members, is of great importance in social life. However, there were debates about the neural mechanism of face processing in recent studies. The debates focused on the question of whether there was specialized neural basis for processing human faces due to their social significances (domain-specificity) or general-purpose structures for processing a variety of visual information with a certain attribute, for example, the expertise with them (domain-generality) [1,2].

In recent studies, the most controversial neural structures were the bilateral middle fusiform gyrus in the extrastriate cortex [1–6]. It has been demonstrated that these areas were much more responsive to faces than to non-face stimuli [2–5,7]. The activations elicited by faces were bilateral with superiority in the right [7]. With single-unit recordings targeted to fMRI-identified face-selective regions (in temporal lobe) in two monkeys, Tsao et al. found that almost all of the visual-respective neurons in this region more strongly responded to face than to non-face objects, suggesting that there were cortical regions for face processing in macaques [8,9]. Because the activations of the fusiform gyrus have been so strongly associated with face processing, Kanwisher et al. named these regions the fusiform face area, or FFA, and have suggested that the FFA was specialized only in face processing [7].

Some studies with normal subjects using event-related potentials (ERP) and magnetoencephalography (MEG)
reported face-selective responses [10,11], especially N170 which was intimately associated with face processing [10]. In our recent studies, we also found that N170 responded selectively to face. The source accounting for the scalp topography associated with N170 was localized in the right middle fusiform [12].

However, Tarr and Gauthier challenged this domain-specific view of the fusiform gyrus [1]. They proposed that the FFA was not dedicated to face recognition alone. They argued that the FFA was flexible and its function was dependent on visual expertise and the categorization at individual level [1,13,14]. Using “Greebles”, computer-generated stimuli designed to be similar to faces in attributes, Gauthier et al. [13] found significantly increased activations in the right middle fusiform gyrus for subjects who were trained to become experts in processing Greebles. In a later study with car and bird experts, the results also suggested that expertise, rather than the nature of objects, determined the activation of the FFA [14]. Based on these evidences, Tarr and Gauthier argued that it was the interaction of individual-level categorization and visual expertise with faces that led to the activation of FFA [1].

It is puzzling that functional neural imaging studies consistently produced strong evidence to support both theories, given the fact that both theories were diametrically opposite to each other. The methodological differences between studies may account for the contradiction. For example, in the studies of Gauthier et al. [13], the Greebles were like faces and could be easily processed by the way of face perception. One could argue that the activation elicited by Greebles in FFA of Greebles experts was because that they used a face-processing method to discriminate the Greebles [2]. In contrast, in studies of Kanwisher and colleagues, the non-face objects such as houses were far more different from faces on various dimensions (e.g. within-class homogeneity, processing demand, level of processing, length and intensity of exposure, and extent of expertise). Thus, the distinctions of neural presentations between faces and those non-face stimuli were not sufficient to demonstrate the specialization of FFA in face processing.

Faces are truly a unique class of visual stimuli in our environment. Firstly, faces are omni-present. Humans have become experts at processing faces at an extremely high level of proficiency due to daily exposure to them since early childhood. Secondly, faces have a canonical upright orientation. They are seen far more frequently upright than inverted. Thirdly, unlike many other visual stimuli, humans process faces not only at the categorical level (e.g. its race), but also at the individual level, with the latter as the primary focus. Finally, it is well known that the fusiform gyrus is highly responsive to faces (for review, see Ref. [15]). It is extremely difficult to find an ideal comparison stimulus in our environment. However, Chinese written symbols (i.e. characters) are perhaps the only class of visual stimuli that closely resemble the attributes of face stimuli. Chinese characters are omni-present in the Chinese societies and Chinese people are exposed to them from early childhood. The characters have a canonical upright orientation. Literate Chinese adults are experts at processing thousands of individual Chinese characters. Also, such processing is performed at the individual Chinese character level because each Chinese character carries specific meanings. Furthermore, the fusiform gyrus is also responsive to Chinese character processing [16–18]. However, unlike faces that are products of evolution, Chinese characters are cultural products.

The present study aims to investigate the neural mechanism of face processing. The Chinese characters were used as comparison stimuli because of their extreme similarity to faces. It is hypothesized that if FFA is more responsive to face than to Chinese characters, it may be specialized in face processing, or at least, this specificity of FFA is not fully attributed to the interaction of visual expertise and individual categorization. Alternatively, if FFA responds equally to face and Chinese characters, then the activation of FFA may result from visual expertise and individual categorization.

2. Materials and methods

2.1. Subjects

Eleven healthy, right-handed Chinese undergraduates (5 males, mean age: 21.3, SD: 1.76) without recent substance abuse or past history of neurological or psychiatric illness participated in the present study. All participants provided written informed consent to participation in the study, which was approved by the ethics committee of Sichuan University Huaxi Hospital.

2.2. Stimuli

Two types of stimuli (faces and Chinese characters) were used in the present study. The face stimuli included 60 face pictures created from photographs of Asian undergraduates of University California at San Diego with counterbalance of gender. Chinese characters stimuli contained 60 high-frequency phonograms. Faces and Chinese characters stimuli were converted into two-tone images by Photoshop 8.0.

2.3. Procedure

Experimental scans consisted of four block-designed sessions, two sessions for face task and the other two sessions for Chinese character task. The two face-task sessions of each subject were randomly assigned into two different groups (face group I and face group II), and similarly, the two character-task sessions of each subject were randomly assigned into two different groups (character group I and character group II). Each face-
task session included three 30 s face stimulus epochs interleaved by three 30 s scrambled-picture epochs as baseline. Each face epoch contained six face trials. Each face trial began with a 500 ms fixation followed by the first face presentation for 500 ms, then 1500 ms fixation followed by the second face presentation for 1000 ms, finally 1500 ms fixation during which the participants were asked to judge whether the two faces in this trial were identical or not. During the scrambled-picture epochs, participants only passively viewed the pictures. Each character-task session was similar to face-task session, except including Chinese character pictures instead of face pictures. The four sessions were scanned in an interleaved order.

2.4. fMRI data acquisition

During each session, 60 whole brain T2*-weighted axial images were acquired by a 3.0 T MRI scanner (GE Signa Excite System, America) with a quadrature RF coil using standard EPI sequence (36 contiguous axial slices, slice thickness 4 mm, TR = 3000 ms, TE = 30 ms, FOV = 240 mm, flip angle = 90°, matrix size 64 × 64).

2.5. fMRI data analysis

Image data analysis included preprocessing and statistical analysis, which were performed using statistical parametric mapping (SPM2, Wellcome Department of Imaging Neuroscience, London, UK; http://www.fil.ion.ucl.ac.uk/spm) [19] and implemented in Matlab 7.1 (the Mathworks Inc., USA). During preprocessing, after spatial realignment to the first volume, all scans of each participant were normalized to a standard EPI template (MNI152 provided by the Montreal Neurological Institute) and resampled to 2 × 2 × 2 mm voxels for statistic analysis, and then spatially smoothed with an isotropic 6 mm full-width-half-maximal (FWHM) Gaussian kernel to decrease spatial high frequency noise and ensure the validity of inferences based on parametric tests. The time series of each session was high-pass filtered (high-pass filter = 128 s) to remove scanner drifts [19].

For each participant, the image data of each session were analyzed using general linear model (GLM), where the regressor was created by convolving a canonical hemodynamic response function (HRF) with a delta function corresponding to the presentation sequence of each stimulus type. Movement parameters were used in GLM as regressors to account for residual effects related to movement.

Parameter estimates of regressors were obtained at each voxel of the whole brain for each participant, and contrast images between stimuli were generated by combining the parameters of corresponding regressors with appropriate contrasts. The contrast images of individual participant underwent a second-level random effects analysis with t test, regarding participants as a random variable [20].

Across all subjects, one-sample t test was performed on face task and Chinese task, and a paired t test analysis was performed between those two tasks. For each subject, the correlations of response patterns were calculated between two face-task sessions or two character-task sessions (within-category correlation), and between face-task session and character-task sessions (between-category correlation). All group analyses and correlation analyses were limited to bilateral fusiforms.

3. Result

3.1. Behavior results

Participants were highly accurate in face and Chinese character conditions (96.3% and 96.6%). Mean correct reaction times for the face and Chinese character conditions were 698.5 ms and 657.2 ms, respectively. Paired t tests were performed on mean accuracy and mean correct reaction time of processing between faces and Chinese characters. No significant differences in mean accuracy were found between two stimuli (p > 0.05). In contrast, the significant difference in mean accurate reaction time between such two stimuli revealed that subjects took more time to discriminate faces than Chinese characters, given approximately equal reaction accuracy.

3.2. fMRI data results

Fig. 1 shows the activation elicited by faces (Fig. 1a) and by Chinese characters (Fig. 1b) in fusiform. It can be seen that faces and Chinese characters both activated bilateral fusiform with superiority in the right (p < 0.001 corrected). The regions activated by these two stimuli were greatly similar. For each subject, this similarity was quantitatively measured by correlation of within-category and between-category response patterns in fusiform. Fig. 2 shows the mean within-category and between-category correlations across all subjects. We can see that both within-category and between-category correlations were significantly larger than that at chance (p < 0.001). The within-category correlation of face (Fig. 2 left green line in Fig. 2) was significantly greater than between-category correlation (Fig. 2 blue lines in Fig. 2) (p < 0.05) and within-category correlation of Chinese characters (Fig. 2 right green line in Fig. 2) (p < 0.01). Fig. 3 gives the comparison of activation between faces and Chinese characters. As revealed by Fig. 3, greater activities elicited by faces relative to Chinese characters were observed in right middle fusiform that was consistent with FFA (p < 0.001 uncorrected), whereas greater activities elicited by Chinese characters relative to faces were not observed in the fusiform even at the significant level of p < 0.05. Table 1 summarizes the activation elicited by faces or Chinese characters relative to base, and the distinction of activation patterns between faces and Chinese characters.
Fig. 1. Activated regions elicited by faces relative to base (a) and Chinese characters relative to base (b) in fusiform across all subjects (Z is defined by MNI atlas).

Fig. 2. The mean correlated coefficients of within-category sessions (green) and between-category sessions (blue) across all subjects. (a) and (b) are the activations elicited by faces for face group I and face group II, respectively. (c) and (d) are the activations elicited by Chinese characters for character group I and character group II, respectively. The correlated coefficient between each pair of sessions is labeled between them. In each sub-figure, the coordinates of transverse slices are Z = −20 (left) and Z = −16 (right). Z is defined by MNI atlas.

Fig. 3. Comparison of activation between faces and Chinese characters. The hot color (positive value) indicates greater activities elicited by faces relative to Chinese characters. No activation of Chinese characters relative to faces is observed in fusiform even in the level of p < 0.05. Z is defined by MNI atlas.
4. Discussion

The present study aims to investigate the neutral mechanism of face processing. The Chinese characters were used as comparison stimuli because of their extreme similarity to faces. Different from the linear-arranged alphabetical words, Chinese characters, the primary units of Chinese, have nonlinear square configuration, and hence contained more configural information [21]. The face pictures were converted into the two-tone image, so that they were matched with Chinese characters in terms of luminance, color and gray intensity level.

Both face and Chinese character stimuli activated bilateral fusiform with great similarity, especially in the right hemisphere (Fig. 1), consistent with previous studies of processing of face [7,14] and Chinese characters [16–18]. Recent studies have provided converging evidence on the fact that the left fusiform selectively responded to alphabetical words [22–25], for example, “the visual word form area” (VWFA), which was suggested to be specialized for visual word processing [26,28]. In contrast, the representation of Chinese characters was usually observed in bilateral fusiform [16–18]. Unlike alphabetical words, Chinese characters contain more configural information and share many common characteristics with faces. The most important similarity between faces and Chinese characters is that literate Chinese adults have much more expertise with them than with other objects, which automatizes the processing of faces and Chinese characters at individual level [1,2,27]. It has been suggested that the right hemisphere may be involved in more configural processing, whereas the left hemisphere may be more associated with the processing of feature [29,30]. Based on this evidence, Koutstaal et al. [31] demonstrated that the right fusiform presented more sensitivity to alterations of configural form, whereas the left fusiform was more sensitive to the change of feature information. This right lateralization of configural processing may be a potential interpretation of the bilateral representation of Chinese characters in the present and previous studies. Within ventral occipitotemporal junction, the cells with similar selectivity cluster to column, and the columns selectively responding to different but related attributes were mixed and overlapping [32]. The objects are represented in a continuous and distributed network [32–34]. The similarity between patterns of response to faces and Chinese characters in fusiform, in the present study, may be attributed to the similarities of attributes between them.

To quantitatively measure such similarity of response patterns, the correlation of within-category (between two face-task sessions or between two character-task sessions) and between-category (between face-task sessions and character-task sessions) response patterns in fusiform was calculated. As shown in Fig. 2, it is of interest that strong between-category correlations were observed (blue lines in Fig. 2). Haxby et al. [35] used similar method to investigate the distinctive response pattern to face and house. Different from the present study, it is reported that the mean within-category correlations (for faces 0.81 and for house 0.87) were greatly larger than the mean between-category correlations (−0.40 and −0.47). Such difference from the present study may reflect that Chinese characters might be more “like” faces than other non-face objects (e.g. house, tools). Therefore, it is reasonable to suggest that the correlation between response pattern to faces and to non-face objects may be modulated by the similarity of their attributes that include not only the visual form of objects but also the other characteristics, for example, the expertise and processing mode. However, how to accurately decide on such attributes and quantitatively measure their similarity between objects in different categories on need to be investigated.

The distinction in neural representation between faces and Chinese characters was revealed by the comparison between activations elicited by these two stimuli (Fig. 3). Within the fusiform, the activation associated with faces focused on the middle lateral region (42, −49, −14) which was consistent with the loci of FFA (40 ± 6, −55 ± 8, −10 ± 5) [7] and (39 ± 3, −40 ± 7, −16 ± 5) [3]. These results were also in agreement with other studies, in which increased activities of faces relative
to houses and chairs were observed in the right fusiform (40, −52, −19) [33]. In contrast, no significantly greater activities of Chinese characters relative to faces were observed in fusiform. Such distinction between neural representations of faces and Chinese characters may result from the difference in some attributes between them such as producing mode (evolutional vs. cultural), structural information (oval vs. square) and feature information (strokes vs. facial features). As mentioned above, Chinese characters are extremely similar to faces in terms of expertise and processing level. So, the distinction in neural representation between faces and Chinese characters may suggest that the activation of FFA elicited by face relative to other objects is not attributed, or at least not absolutely attributed, to the visual expertise with them and categorization at individual level. It provided strong evidence supporting the role of FFA specialized in face processing. Such specificity of neural response to face reflects the unique attributes of face that is learned by people with implicit feedback through long time [27]. However, how to accurately describe these attributes is unclear. This finding may also provide a potential explanation for the results of some neuropsychological studies that some prosopagnosic patients preserved normal ability to discriminate between objects with which they had expertise, even in the same category [36–38], and other patients with deficits in discrimination between familiar objects could identify faces successfully [39].

In the present study, Chinese characters were used as comparison stimuli because of their extreme similarity to faces on a variety of dimensions. The patterns of response to faces and to Chinese characters were highly correlated. However, despite this high correlation, the distinction of neural representation between faces and Chinese characters was observed, whose locus was in agreement with that of FFA. Such distinction demonstrated that FFA was specialized in the processing of face per se rather than the processing of visual expertise and categorization at individual level.

Acknowledgements

This work was supported by Chair Professors of Cheung Kong Scholars Programme, PCSIRT (Grant No. IRT0645) and CAS Hundred Talents Program, the National Program on Key Basic Research Projects (Grant No. 2006CB705700), the National High-Tech R&D Program of China (Grant No. 2006AA04Z216), the Joint Research Fund for Overseas Chinese Young Scholars (Grant No. 30528027), National Natural Science Foundation of China (Grant Nos. 30600151, 30500131, 60532050), and Beijing Natural Science Fund (Grant Nos. 4051002, 4071003). The authors thank Prof. Qiyong Gong and other researchers in MRI Research Center of Huaxi Hospital of Sichuan University for their help in fMRI data acquisition.

References


