

available at www.sciencedirect.comwww.elsevier.com/locate/brainres**BRAIN
RESEARCH****Research Report****Effective connectivities of cortical regions for top-down face processing: A Dynamic Causal Modeling study**

Jun Li^{a,b}, Jiangang Liu^c, Jimin Liang^a, Hongchuan Zhang^e, Jizheng Zhao^{a,b}, Cory A. Rieth^g, David E. Huber^g, Wu Li^d, Guangming Shi^b, Lin Ai^f, Jie Tian^{a,d,*}, Kang Lee^{g,h,*}

^aLife Sciences Research Center, School of Life Sciences and Technology, Xidian University, Xi'an 710071, China^bSchool of Electronic Engineering, Xidian University, Xi'an 710071, China^cSchool of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China^dInstitute of Automation, Chinese Academy of Sciences, Beijing 100190, China^eDevelopmental Cognitive Neuroscience Laboratory, Northwestern University, Illinois, USA^fDepartment of Neuroimaging, Beijing Tiantan Hospital, Capital Medical University, Beijing 100050, China^gDepartment of Psychology, University of California, San Diego, La Jolla, California, USA^hInstitute of Child Study, University of Toronto, Toronto, Ontario, Canada

ARTICLE INFO

Article history:

Accepted 18 April 2010

Available online 25 April 2010

Keywords:

Face processing

Top-down processing

Bottom-up processing

Dynamic Causal Modeling (DCM)

Orbitofrontal cortex (OFC)

ABSTRACT

To study top-down face processing, the present study used an experimental paradigm in which participants detected non-existent faces in pure noise images. Conventional BOLD signal analysis identified three regions involved in this illusory face detection. These regions included the left orbitofrontal cortex (OFC) in addition to the right fusiform face area (FFA) and right occipital face area (OFA), both of which were previously known to be involved in both top-down and bottom-up processing of faces. We used Dynamic Causal Modeling (DCM) and Bayesian model selection to further analyze the data, revealing both intrinsic and modulatory effective connectivities among these three cortical regions. Specifically, our results support the claim that the orbitofrontal cortex plays a crucial role in the top-down processing of faces by regulating the activities of the occipital face area, and the occipital face area in turn detects the illusory face features in the visual stimuli and then provides this information to the fusiform face area for further analysis.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Humans are exceptionally skilled at face perception. Over a wide range of viewing conditions we effortlessly detect and recognize faces accurately and quickly. Traditionally, face processing is viewed as a feed-forward, bottom-up process in

which facial identity is processed first in the ventral visual stream and then passed along to more anterior regions of the brain such as the frontal cortex. This face processing model has been supported by many existing studies. However, in these experiments the stimuli are high quality images of actual faces, maximizing bottom-up information. This may

* Corresponding authors. J. Tian is to be contacted at Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China. Fax: +86 10 62527995. K. Lee, Institute of Child Study, University of Toronto, 45 Walmer Road, Toronto, Ontario, Canada M5R 2X2. Fax: +1 416 978 6485.

E-mail addresses: tian@ieee.org (J. Tian), kang.lee@utoronto.ca (K. Lee).

swamp any top-down influences on face perception, thus biasing interpretation of experimental results in favor of bottom-up driven processing models.

Providing evidence that supports the bidirectional interactive face processing model proposed by Haxby et al. (2000), recent studies using novel stimuli and paradigms have questioned the traditional view about the neural systems of face processing. These studies focused on top-down face processing by requiring participants to imagine faces (Ishai et al., 2000; Mechelli et al., 2004), or to interpret ambiguous face stimuli such as Mooney faces and vase/face illusion (Andrews and Schluppeck, 2004; Hasson et al., 2001), or to detect impoverished face stimuli (Gosselin and Schyns, 2001; Summerfield et al., 2006; Wild and Busey, 2004), or to even make illusory face detection of pure visual noise stimuli (Li et al., 2009; Liu et al., 2010; Zhang et al., 2008). It was found that top-down feed-backward mechanisms play an important role in face processing, which is perhaps engendered by the neural system's use of face-relevant knowledge and learned expectations that regulate the bottom-up processing of visual stimuli (Mechelli et al., 2004; Summerfield et al., 2006). Further, these studies revealed a distributed cortical network for top-down face processing (Li et al., 2009), which overlaps to a large extent with the face processing network reported in traditional bottom-up face processing studies (for reviews, see Haxby et al., 2000; Ishai et al., 2005; Ishai, 2008). Taken together, the findings from bottom-up and top-down paradigms suggest that the neural system for processing faces involves a network of neural regions distributed from occipital to frontal cortices that has both feed-forward and feed-backward connections (Fairhall and Ishai, 2007; Haxby et al., 2000; Ishai, 2008; Li et al., 2009; Mechelli et al., 2004; Summerfield et al., 2006).

Previous studies of top-down illusory face processing (Li et al., 2009; Zhang et al., 2008) used conventional analyses that cannot determine in which direction signals flow and whether connections between brain regions are modulated during the task. The term “effective connectivity” is used in referring to the connection strength between different brain regions and how these strengths vary with experimental manipulations. To understand the top-down face processing network in terms of effective connectivity, the analyses reported here used Dynamic Causal Modeling (DCM; Friston et al., 2003). This analysis not only determines the active neural connections between brain regions during the experiment, but it also determines the direction of the intrinsic and modulatory cortical pathways specifically involved in top-down face perception.

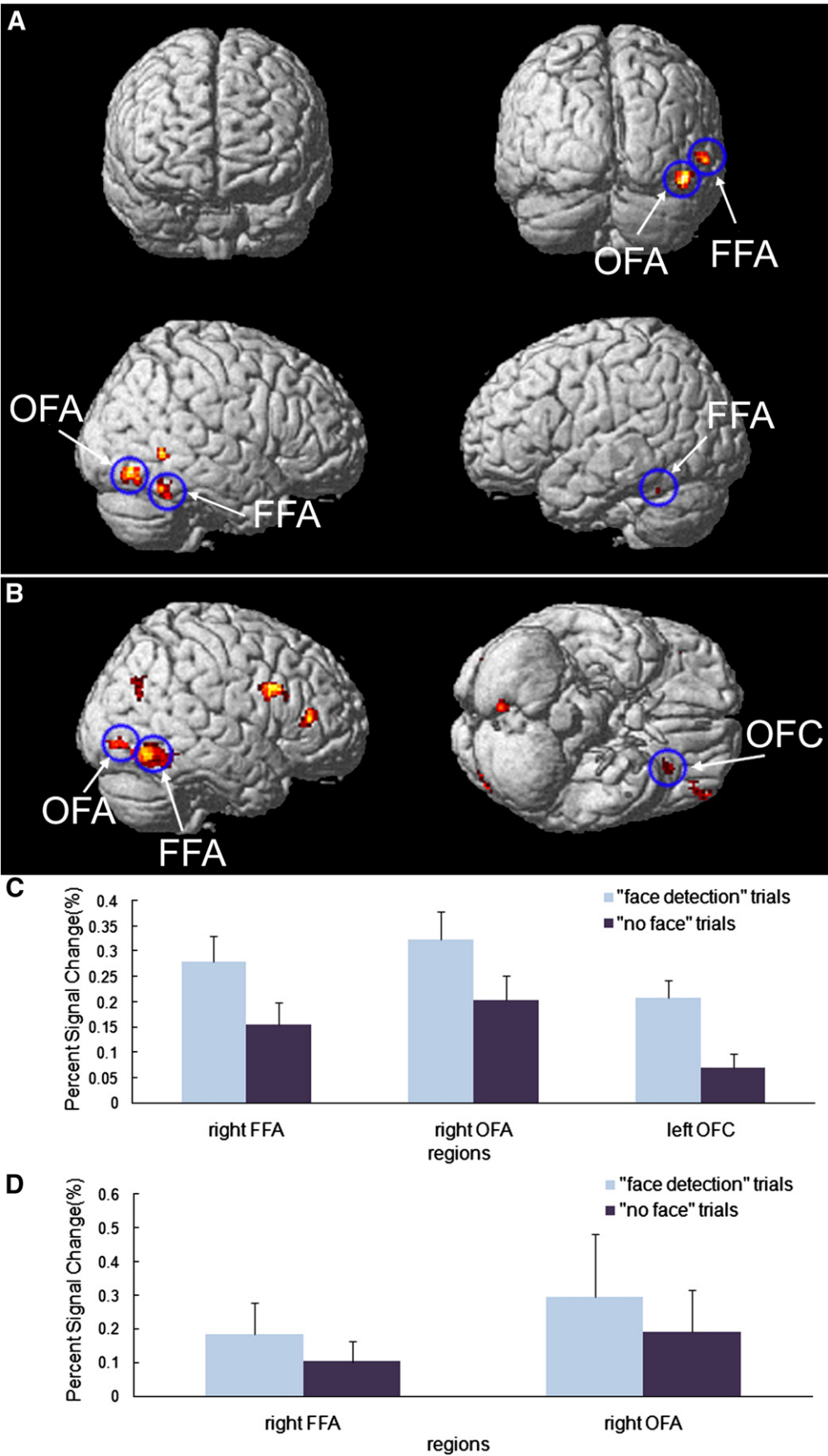
The present study focused on three cortical regions identified by traditional analyses: the fusiform face area (FFA), occipital face areas (OFA), and orbitofrontal cortex (OFC). It is now well established that the FFA and OFA play an important role in face processing and are part of the bottom-up and top-down face processing networks (Fairhall and Ishai, 2007; Liu et al., 2009; Mechelli et al., 2004; Summerfield et al., 2006; Zhang et al., 2009). Thus, these are “core regions” of face processing. However, the function of the OFC and its relation to the core face processing regions is less clear.

A number of functional neuroimaging studies have identified OFC activation both during face processing and during

processing of non-face objects. It has been proposed that the OFC is involved in encoding novel information (Frey and Petrides, 2000; Frey et al., 2004) as well as in mediating the perception of attractive and sexually relevant faces (Ishai, 2007; Kranz and Ishai, 2006; O'Doherty et al., 2001). More importantly, recent studies of object recognition revealed that the OFC plays a key role in top-down object processing (Bar et al., 2006; Bar, 2009; Johnson, 2005; Kveraga et al., 2007a,b). Specifically, it has been proposed that the OFC uses low spatial frequency visual information to form a coarse prediction of the most likely candidate object, which is used to prime the corresponding object processing areas in the ventral occipital-temporal cortex in a top-down manner. This hypothesis predicts that during top-down face detection, the OFC should have functional connections to the OFA or the FFA. The present study tested this prediction.

Based on our recent investigation on top-down face processing (Li et al., 2009; Zhang et al., 2008), we adopted a novel paradigm that promotes illusory face detection in response to images that only contain noise. Participants were told that half of the images in the experiment contained faces and the other half did not. Their task was to detect which images contained faces. An initial training stage of the experiment did indeed contain faces on 50% of the trials. During this training, the faces became more and more difficult to detect by mixing higher degrees of noise with the faces. Eventually, participants were only shown pure noise images, although they were instructed that there were still faces on 50% of the trials and that face detection would be very difficult. The noise images were a mixture of Gaussian blobs of different spatial frequencies placed randomly throughout the image. These complex noise images lend themselves to a large number of interpretations and participants readily continued to detect faces. Thus, we were able to study top-down influences using false detections of faces to images containing only noise, which avoids contamination from strong bottom-up face information. Furthermore, an independent localizer task was performed to validate the ventral occipito-temporal face-sensitive areas identified by the illusory face detection task. Our recent study using this method revealed a complex distributed cortical network for top-down face processing (Li et al., 2009). However, this finding was obtained by using simple correlational analyses (Psychophysiological Interaction, or PPI) with the right FFA as the seed region. Not only is this method unable to measure the directional effective connectivity between different brain regions involved in top-down face processing, but this method also has potential methodological problems such as “double dipping” in which the same data is used for more than one analysis (see Kriegeskorte et al., 2009).

To determine the directional top-down effective connectivities involved in the top-down face processing as well as to avoid the methodological problems associated with PPI, we used DCM (Friston et al., 2003) in combination with Bayesian model selection (Penny et al., 2004) to analyze the data obtained in Li et al. (2009). The use of DCM has several advantages in addition to avoiding double dipping (Stephan et al., 2010). First, this analytic method provides information not only about the intrinsic effective connectivities among various brain regions (i.e., connections strengths that are constant throughout the



experiment), but it also provides information about enhanced connectivities due to a specific processing demand (i.e., illusory face detection). Rather than revealing simple correlational relationships, DCM extracts directional relationships, providing information about how different brain regions are functionally connected during object processing (e.g. in a feed-forward or feed-backward manner). This information is particularly important for the present study considering that we were interested in the interplay between the FFA, OFA and OFC in top-down face processing.

2. Results

2.1. Behavioral results

The average proportion of trials on which subjects responded “face” was 34%, with a standard deviation (SD) of 14%, across the 480 pure noise images. The mean reaction times of the “face” and “no face” responses were 723 ms (SD=126 ms) and 698 ms (SD=119 ms) respectively, which are not significantly different from each other ($t(10)=1.6$, $p=0.169$).

2.2. Conventional fMRI analysis

In the localizer task, in the right hemisphere, all twelve subjects showed activation in the right middle fusiform gyrus and the right lateral occipital cortex in response to passive viewing of faces, as compared with other objects. In contrast, in the left hemisphere, only seven subjects exhibited face-specific responses in the left middle fusiform gyrus and eight subjects did so in the left lateral occipital cortex. This right hemisphere dominance is highly consistent with the existing findings (Kanwisher et al., 1997; Kim et al., 1999; O’Craven and Kanwisher, 2000). For the group level results of the localizer task (uncorrected $p<0.0001$, extent threshold: 15 voxels), the peak coordinates of the middle fusiform gyrus (Talairach coordinate, right: 42, –51, –14; left: –38, –49, –18) and the lateral occipital cortex (right: 48, –72, –6) are consistent with the loci of the FFA and the OFA reported in previous studies of face processing (Grill-Spector et al., 2004; Kanwisher et al., 1997; Rossion et al., 2003) (Fig. 1A). We also found significant activation in the group map for the right posterior superior temporal sulcus (57, –52, 3). However, only six subjects exhibited face-specific responses in this area at the individual level. Consequently, this area was not included in the following DCM analysis.

To identify brain regions specifically contributing to top-down face processing, blood oxygen level dependent (BOLD) responses

Table 1 – Average cluster activation for the main effects of “face” trials compared to “non-face” trials with conventional analysis.

Regions	BA	Voxel	Z	Talairach		
				x	y	z
R fusiform gyrus	37	343	4.22	46	–51	–8
L fusiform gyrus	37	425	5.04	–40	–57	–11
R inferior occipital gyrus	19	343	4.39	44	–78	0
L inferior occipital gyrus	18	425	4.45	–36	–84	–5
L inferior parietal lobule	40	665	5.41	–42	–40	44
L orbitofrontal cortex	47	65	4.05	–28	21	–4
R inferior frontal gyrus	46	95	5.01	44	33	6
L inferior frontal gyrus	9	169	4.40	50	11	25
L middle frontal gyrus	9	50	4.19	–53	7	29
L middle frontal gyrus	46	79	5.11	–40	30	11
L middle frontal gyrus	10	51	4.33	–46	51	3
R precuneus	7	274	5.1	24	–52	43
Medial frontal gyrus	8	57	4.61	–2	24	49
Declive		74	4.87	12	–71	–17

Coordinates of the peak voxel are shown for each cluster. All activations are significant at $p<0.0001$ (uncorrected); k cluster ≥ 50 voxels; R, right hemisphere; L, left hemisphere; BA, Brodmann’s area; voxel size is $2 \times 2 \times 2$ mm³.

during the “face detection” trials and “no face” trials were contrasted. Table 1 and Fig. 1B present the patterns of activation at the group level. The maxima of clusters (uncorrected $p<0.0001$, extent threshold: 50 voxels) were used as the reference for choosing the individual regions of interests (ROIs) (Table 1 and Fig. 1B). The loci were in agreement with the results of the localizer task and the findings of the existing studies for the right FFA (46, –51, –8) (Kanwisher et al., 1997), right OFA (44, –78, 0) (Rossion et al., 2003), and left OFC (–28, 21, –4) (Bar et al., 2006; Kveraga et al., 2007a).

As shown in Fig. 1C, a greater intensity of the mean Percent Signal Changes (PSCs, calculated using Marsbar 0.41, <http://marsbar.sourceforge.net/>) across 11 subjects (one subject was excluded due to the lack of identifiable OFC ROI) were found in all three ROIs defined by the illusory face detection task for the “face detection” trials compared to the “no face” trials. Paired t tests were conducted between the two conditions (“face detection” trials and “no face” trials) on the three ROIs across all 11 subjects, which revealed significant differences in all ROIs ($t(10)=12.2$, $p<0.001$ in right FFA, $t(10)=10.4$, $p<0.001$ in right OFA, and $t(10)=8.4$, $p<0.001$ in left OFC).

We also calculated the mean PSCs between the “face detection” trials and “no face” trials in the ROIs for the FFA and the OFA defined by the localizer task for each subject. Similar

Fig. 1 – (A) Group level statistical analysis maps of the localizer task defined by contrasting fully visible faces with non-face stimuli. The threshold was set at $T>5.45$ ($p<0.0001$, uncorrected) and cluster >15 voxels. (B) Group level statistical analysis maps of the “face detection” trials compared to the “no face” trials which was used to select ROIs for the individual level analysis. The threshold was set at $T>5.45$ ($p<0.0001$, uncorrected) and minimum cluster volume = 50 voxels. The areas in blue circles indicate the ROIs used in the DCM models, which includes the right FFA, right OFA and left OFC. (C) Mean percent signal changes for “face detection” trials and “no face” trials across 11 subjects in the three ROIs of the right FFA, right OFA and left OFC as identified by the illusory face detection task. (D) Mean percent signal changes for “face detection” trials and “no face” trials across 11 of the same subjects in the two ROIs of the right FFA and right OFA as identified by the localizer task. Error bars denote standard errors. FFA, Fusiform Face Area; OFA, Occipital Face Area; OFC, Orbitofrontal Cortex.

Table 2 – Model comparison between the optimal model of Model 12a and other models with DCM: Group Bayes factors (GBF) and binomial p values.

Model 1a	Model 1b	Model 2a	Model 2b	Model 3a	Model 3b	Model 4a	Model 4b
GBF=9.9e+116 $p=0$	GBF=8.6e+101 $p=0$	GBF=9.1e+38 $p=0^{**}$	GBF=1.0e+47 $p=0^{**}$	GBF=2.6e+76 $p=0$	GBF=3.9e+75 $p=0$	GBF=9.4e+35 $p=0.011^*$	GBF=2.5e+44 $p=0^{**}$
Model 5a	Model 5b	Model 6a	Model 6b	Model 7a	Model 7b	Model 8a	Model 8b
GBF=1.4e+65 $p=0$	GBF=4.2e+67 $p=0$	GBF=1.3e+67 $p=0.006$	GBF=3.9e+72 $p=0$	GBF=1.2e+69 $p=0^*$	GBF=9.4e+71 $p=0$	GBF=9.9e+116 $p=0$	GBF=1.9e+61 $p=0$
Model 9a	Model 9b	Model 10a	Model 10b	Model 11a	Model 11b	Model 12a	Model 12b
GBF=2.7e+163 $p=0$	GBF=3.4e+95 $p=0$	GBF=8.1e+30 $p=0$	GBF=4.7e+23 $p=0^*$	GBF=1.7e+42 $p=0$	GBF=1.3e+37 $p=0$	–	GBF=348 $p=0^{***}$
Model 13a	Model 13b	Model 14a	Model 14b	Model 15a	Model 15b	Model 16a	Model 16b
GBF=1.1e+12 $p=0$	GBF=5.1e+11 $p=0.006$	GBF=2.5e+13 $p=0^*$	GBF=4.8e+15 $p=0$	GBF=1.3e+25 $p=0^*$	GBF=2.2e+24 $p=0^{**}$	GBF=9.7e+34 $p=0^{**}$	GBF=6.8e+33 $p=0^{**}$

Three subjects (***), two subjects (**) and one subject (*) not included because AIC and BIC did not favor the same model.

results were obtained (Fig. 1D): significant differences of PSCs between the “face detection” trials and “no face” trials were found in both the localizer defined FFA ($t(10)=5.07$, $p<0.001$) and OFA ($t(10)=3.25$, $p=0.004$) across the same 11 subjects. These results indicated that the ROIs defined by the contrasts between the “face detection” trials with the “no face” trials in the illusory face detection task can be defined as face-selective ROIs in the subsequent DCM analysis.

2.3. DCM results

2.3.1. Model comparison

Table 2 shows the group Bayes factors (GBF) and the binomial p values for all model comparisons between thirty-two possible models of directional connections among the FFA, OFA, and OFC based on the existing studies (Fig. 4 for all models; also see

the method below). In all comparisons, strong evidence ($\text{GBF}>348$, $p<0.015$) was found in favor of Model 12a. Therefore, this model was assumed to underlie the top-down face processing task and the resultant parameters for this model were considered next.

2.3.2. Effective connectivity analysis

The group averaged parameters of Model 12a were calculated using the DCM averaging routing of the SPM5 package.

2.3.2.1. *Intrinsic connectivities and direct input effects.* Intrinsic connectivities represent the default state of interregional coupling and are measured through interactions among brain regions that are independent of the task (Friston et al., 2003). Fig. 2A summarizes the significant intrinsic connectivities among the three brain regions. All of the three

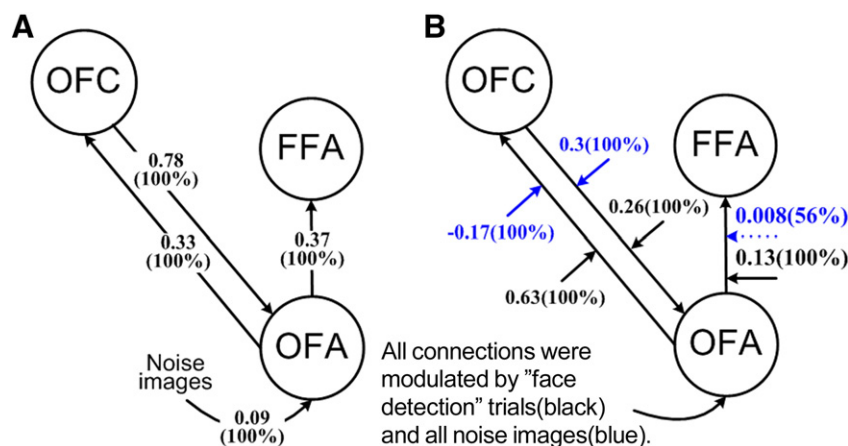


Fig. 2 – Connectivity parameters for the optimal Model 12a. Strengths averaged across individuals are presented. (A) Intrinsic connections and direct input effects. (B) Modulatory effects for the “face detection” trials (black) and those for presentation of a noise image regardless of face detection (blue). Solid lines indicate significant intrinsic (A) or modulatory (B) effects ($p>0.95$) and dotted lines indicate non-significant effects.

intrinsic connectivities were highly reliable, as revealed by posterior probabilities of 1.0. The strength of influence of the OFA on the FFA was 0.37, the OFA on the OFC was 0.33, and the OFC on the OFA was 0.78. The direct input effect of visual stimuli (presentation of a noise image) on the OFA was 0.09 and significant (posterior $p=1.0$) (Fig. 2A).

2.3.2.2. Modulatory effects. Modulations of the connectivities between the three regions for the “face detection” trials as well as those for all noise images are depicted in Fig. 2B. Regarding the modulatory effects of the “face detection” trials, it was found that the reciprocal connectivities between the OFC and the OFA were enhanced significantly (into the OFC: strength=0.63, posterior $p=1.0$, from the OFC: strength=0.26, posterior $p=1.0$), and the OFA had a significant modulatory effect on the FFA (strength=0.13, posterior $p=1.0$).

In contrast, with regard to the modulatory effects of all noise images (regardless of face detection), only the OFC was found to have significant positive modulatory effects on the OFA (strength=0.3, posterior $p=1.0$). Moreover, the OFA had a significantly strong negative modulatory effect on the OFC (strength=-0.17, posterior $p=1.0$) but very weak positive modulatory effect on the FFA (strength=0.008, posterior $p=0.56$).

3. Discussion

The present study examined illusory face detection to pure noise images to investigate the neural networks involved in top-down face processing. Consistent with previous findings using PPI (Li et al., 2009), conventional BOLD signal analysis identified three core brain regions, namely the FFA, OFA, and OFC, that were highly responsive to trials on which participants detected a face. Focusing on these three regions, application of Bayesian model selection to Dynamic Causal Modeling determined the model that best accounted for both intrinsic and modulatory effective connections among the three areas.

3.1. The optimal intrinsic effective connectivity model

Out of 32 plausible intrinsic (i.e., response independent) effective connectivity models between the OFA, FFA, and OFC, the optimal model was as follows. First, there were reciprocal connections between the OFC and the OFA, although the feed-backward connection of the OFC to the OFA was stronger (0.78). This is consistent with a recent proposal that the posterior OFC facilitates visual recognition in a top-down manner based on low spatial frequency information; within the dorsal magnocellular pathway, this low spatial frequency information is rapidly available through projections from early visual areas to the prefrontal cortex (Bar et al., 2006; Kveraga et al., 2007a,b).

Second, there were significant feed-forward connections from the OFA to the FFA. This is consistent with the proposal that the OFA is involved in the initial processing of face-like features before providing this information to the FFA for further analysis (Fairhall and Ishai, 2007; Kveraga et al., 2007a; Mechelli et al., 2004).

Third, there was a failure to find reciprocal connections from the OFC to the FFA. This finding is inconsistent with a recent study that found top-down facilitation of object recognition by the OFC (Kveraga et al., 2007a). In contrast to our results, this study found feed-backward connectivity from the OFC to the FFA but not to the OFA. However, in the previous study, the tested models included only feed-forward connectivity from the OFA to the OFC and a reciprocal connectivity between the OFC and the FFA. Thus, it is possible that the feed-backward connectivity from the OFC to the OFA was misattributed to the FFA due to the lack of the feed-backward connection between the OFC and the OFA. In contrast, we tested models that allowed reciprocal connectivities between the OFC and the OFA and between the OFC and the FFA.

3.2. The modulatory effects of the optimal intrinsic effective connectivity model

Our analysis further revealed that feed-backward connectivity from the OFC to the OFA was enhanced significantly above its intrinsic values both for “face detection” trials and also in response to the pure noise image, regardless of face detection. This result is consistent with recent studies reporting that the OFC is involved in the top-down visual processing of objects within posterior areas (Bar et al., 2006; Kveraga et al., 2007a). It has been suggested that the OFC provides a top-down analysis of low-frequency information through feed-backward connections to visual areas. This feed-backward connectivity may aid the OFA's analysis of feature information when searching for a face in a pure noise image.

We found that the feed-forward connectivity from the OFA to the FFA was enhanced significantly (strength=0.13, $p=1.0$) for “face detection” trials, but it was not significant when we only considered the modulatory effects of presenting a noise image without regard to face detection. This result is consistent with connectivity patterns of the OFA and FFA in face processing found in previous studies (Fairhall and Ishai, 2007; Mechelli et al., 2004). This result also fits well with recent reports that the right FFA represents the identity of faces based on face features initially identified by the OFA (Rotshtein et al., 2005).

Furthermore, we found that the positive feed-forward connectivity from the OFA to the OFC was enhanced significantly (strength=0.63, $p=1.0$) in the “face detection” trials. In contrast, when the modulatory effect of the pure noise images regardless of face detection was considered, this connectivity was significantly negative. These results are consistent with the suggestion that the OFA plays a critical early role in the analysis of faces by detecting face features to construct an initial representation of a face image (Calder and Young, 2005; Haxby et al., 2000; Pitcher et al., 2007). In our paradigm, when participants formed an expectation that the pure noise images contained face-like features, the OFC may have sent signals via the feed-backward connectivity from the OFC to the OFA to assist face detection and extraction of face-like features in the noise image. On this account, if face-like features were detected, the enhanced feed-forward connectivity from the OFA to the OFC enhanced the expectation that the noise image contained a face, which in turn enhanced the likelihood that the OFA detected and extracted face-like features from the noise image (and so on, recursively). However, when the OFA did not detect

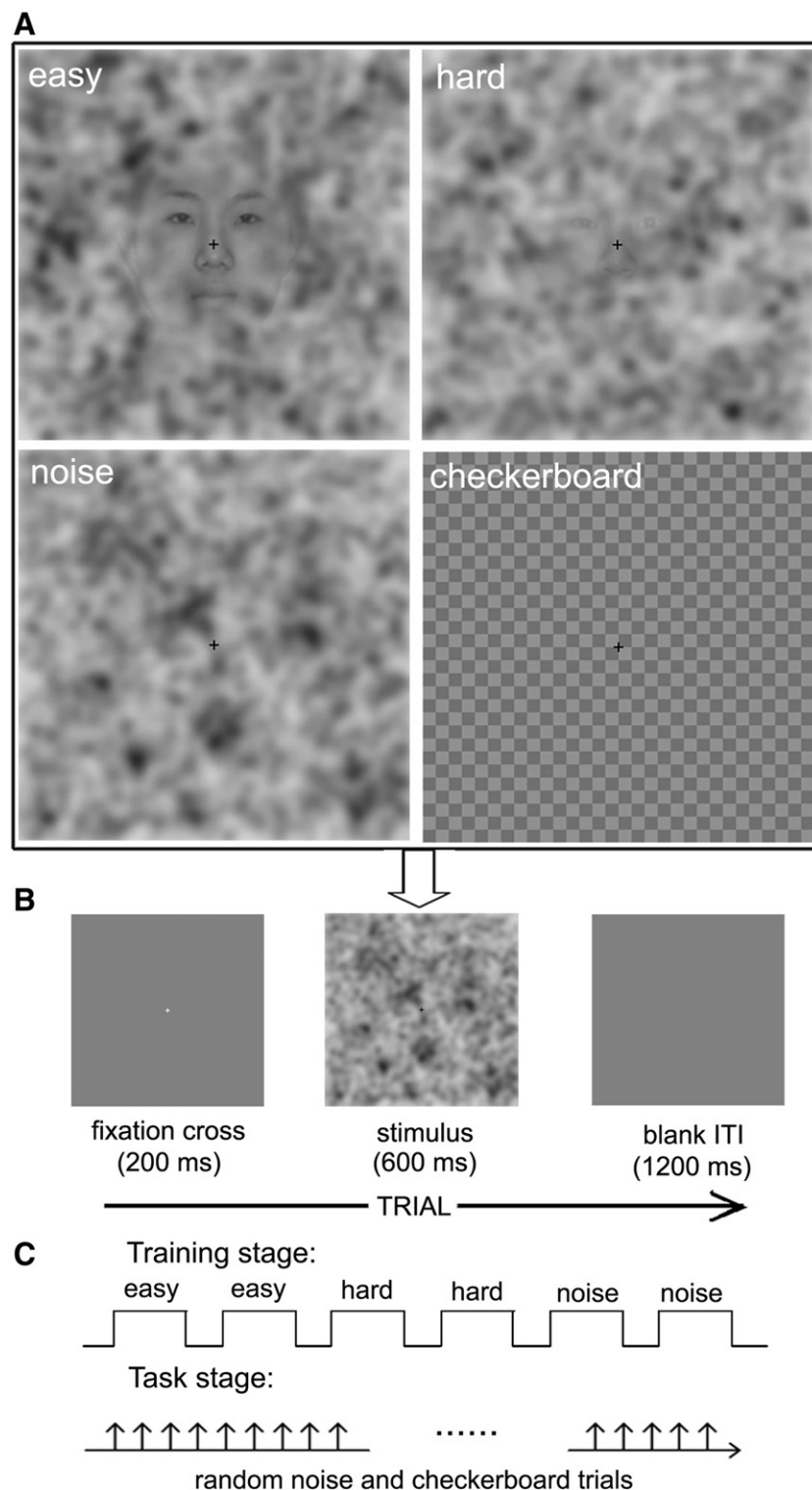


Fig. 3 – Example stimuli and illustration of the experimental design used. (A) Sample stimuli used in the experiment for easy-to-detect-face trials, hard-to-detect-face trials, pure noise trials and checkerboard trials (from top left to bottom right). (B) The sequence of displays in a trial from both the training stage and the illusory face detection task stage. (C) The types of experimental design utilized in the training stage and the illusory face detection task stage respectively. A block design was employed in the training stage and an event-related design was employed in the illusory face detection task stage.

any face-like features, the connectivity from the OFA to the OFC was de-activated, resulting in a negative modulatory effect (see Friston et al., 2003; Plailly et al., 2008).

It should be noted that in both the intrinsic and modulatory networks, the OFA played a central role. One might expect the FFA to play a larger role because it typically reveals a high level

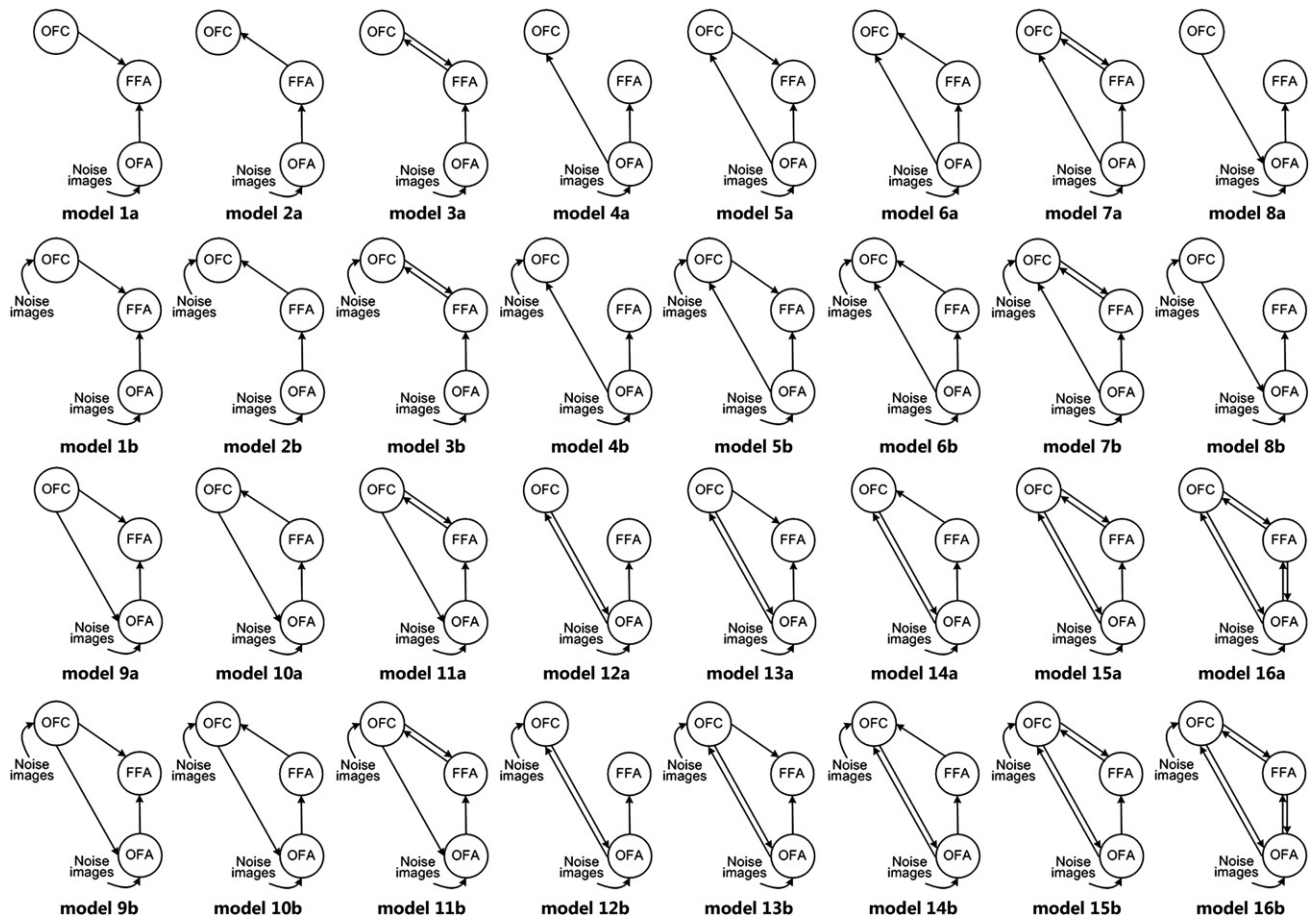


Fig. 4 – Model networks of interregional connections and experimental inputs. Both the “face detection” trials and presentation of noise images regardless of face detection modulated all couplings among the FFA, OFA and OFC.

of activations to faces with the use of conventional BOLD signal analyses. Furthermore, FFA activation is high regardless of whether bottom-up face information is strong (Gauthier et al., 2000; Kanwisher et al., 1997; Kanwisher and Yovel, 2006) or weak (Li et al., 2009; Zhang et al., 2008). However, it is important to keep separate the notions of connection strength versus resultant activation. More specifically, although the OFA may serve as a key generator of illusory face detection, the largest activation may still be in the FFA considering that the FFA receives input from the OFA (Fairhall and Ishai, 2007; Haxby et al., 2000). Furthermore, the pure noise images in the present study forced participants to search for local information that resembled facial features. Had we instructed or primed the participants to search for configural facial information, the role of the FFA might have been enhanced (Schiltz and Rossion, 2006). This possibility needs to be tested in future studies.

In conclusion, the present study used an illusory face detection paradigm to study top-down face processing in response to pure noise images. Conventional BOLD signal analysis revealed three regions specifically involved in illusory face detection. These regions also included the orbitofrontal cortex in addition to the fusiform and occipital areas that were previously known to be involved in both top-down and bottom-up processing of faces. Analysis using Dynamic

Causal Modeling and Bayesian model selection revealed both intrinsic and modulatory effective connectivities among these three cortical regions. These results suggest that the OFC plays a crucial role in top-down face processing by regulating OFA activity, and the OFA in turn detects elements of the noise images that resemble face features and then provides this information to the FFA for further analysis.

4. Experimental procedures

4.1. Subjects

Twelve normal, right-handed subjects (seven males, age 23.8 ± 1.4 years), with normal or corrected-to-normal vision, participated in this study. All subjects gave written informed consent for the procedure in accordance with protocols approved by the Human Research Protection Program of Tiantan Hospital, Beijing, China.

4.2. Design and procedure

The experiment included two stages: an initial training stage and an illusory face detection task stage (Fig. 3C). Four types of stimuli were used: face images overlaid with

50% noise which are easy for participants to detect faces, face images overlaid with 75% noise which are hard for face detection, pure noise images which did not include faces, and checkerboard images that were used as the baseline (Fig. 3A). During the initial training stage, a block design was employed, and subjects completed one session which consisted of six blocks of progressively more difficult face detection (Fig. 3C). The first two blocks used an even mix of pure noise images (noise trials) and face images overlaid with 50% noise (easy-to-detect-face trials), the second two blocks used an even mix of pure noise images (noise trials) and face images overlaid with 75% noise (hard-to-detect-face trials), and the final two training blocks consisted of pure noise images (noise trials) only. Participants were instructed that in each block, half of the 20 noise images contained a face. They were told that the task would become more difficult over time and to indicate whether or not they saw a face in the noise image with their left or right button press (counterbalanced across subjects). For each trial, the image was presented for 600 ms after a 200 ms fixation cross, followed by a blank screen for 1200 ms (Fig. 3B). The training stage taught participants the nature of the experiment and gradually brought them to a point that promoted illusory face detection.

After training, subjects completed four sessions of the face detection task which used an event-related fMRI procedure. Forty checkerboard images (checkerboard trials) were used as control trials and 120 pure noise images (noise trials) were presented randomly in each of session. Subjects were instructed that the task was the same as the third phase of training and that of 50% of the noise trials would include faces. During the checkerboard trials, no responses were required. The responses from each participant were divided into “face detection” trials (when the subject reported detecting face) and “no face” trials (when the subject reported not-detecting face). Participants were scanned in the same session for both the training task and illusory face detection task. However, only those data recorded during the task stage were analyzed. Behavioral data were obtained by recording the participants’ responses during performance of the tasks while in the MRI scanner.

After the illusory face detection task stage, a classical block fMRI design used as a localizer task to identify the brain areas activated while viewing fully visible faces. Each subject completed two sessions separated into eight blocks of three object types (faces, randomly selected objects, and scrambled pictures). Each block consisted of 20 trials in which a stimulus was presented for 600 ms after a 400 ms fixation cross. Of these 20 stimuli, two randomly chosen stimuli contained a white border, which was used as catch trial. Subjects were instructed to press the right button on a response device whenever a white border appeared around a picture. Following each block, there was a fourteen second fixation baseline condition.

4.3. Data acquisition

Functional and structural images were acquired using a 3.0 T-whole body scanner (Siemens Trio a Tim, Erlangen, Germany) at Tianan Hospital. T1-weighted high-resolution ($1 \times 1 \times 1$ mm)

structural images were obtained using a magnetization prepared rapid acquisition gradient echo (MPRAGE) sequence (FOV = 256). Functional images were acquired by using a multislice echo planar imaging (EPI) sequence covering the whole cerebrum (32 axial slices acquired in interleaved sequence, 4 mm slice thickness, TR = 2 s, echo time (TE) = 20 ms, flip angle 90° , matrix = 64×64) with a resolution of 3.75×3.75 mm. Each session consisted of 166 functional volumes. The first three volumes of each functional scan were discarded to compensate for scanner equilibration.

4.4. Conventional image analysis

Imaging data were analyzed using statistics parametric mapping software (SPM5, Wellcome Department of Cognitive Neurology, London, UK, <http://www.fil.ion.ucl.ac.uk/spm>). EPI volumes were spatially realigned to correct for movement artifacts, transformed to the MNI (Montreal Neurological Institute) standard space, and smoothed using a 6-mm Gaussian kernel. Statistical analysis at the first level relied on a general linear model (GLM), with the two types of trials (“face” versus “no face” trials) as the conditions of interest and a regressor encoding all noise image trials for illusory face detection task and three types of conditions for localizer task. The trials from the four separate illusory face detection task sessions were concatenated to form a single session for each individual. Session effects were accounted for by adding four session regressors into the GLM (Bitan et al., 2005). To correct for low-frequency components, a high-pass filter with a cutoff period of 128 s was used (Friston et al., 2002). Task-specific activations were obtained by contrasting hemodynamic responses during the face and no face trials for the illusory face detection task, and during the viewing of fully visible faces and non-face stimulus for the localizer task. Statistical evaluation of group data was based on second-level random effects analysis with a height threshold of $p < 0.0001$ (uncorrected) and extent threshold of $k > 50$ voxels for the illusory face detection task (see Table 1 and Fig. 1B) or of $k > 15$ voxels for the localizer task (see Fig. 1A). The results of the localizer task were used to validate the ventral occipito-temporal face-sensitive areas identified by the illusory face detection task. The group results of the illusory face detection task were then used for choosing the regions of interest (ROIs) for the effective connectivity analysis as required by DCM.

4.5. Selection of volumes of interest

The general goal of DCM is to make inferences about the possible connectivity among brain regions and the influence of one region on another in a given experimental context. The three brain regions whose role in the top-down face processing was investigated were the right FFA, the right OFA, and the left OFC (Table 1 and Fig. 1B). For the FFA and OFA, there is overwhelming evidence from electrophysiological and imaging studies that indicates a right hemisphere advantage for face processing (Allison et al., 1999; Kanwisher et al., 1997; Kim et al., 1999; O’Craven and Kanwisher, 2000), which is supported by the results of our localizer task. For the OFC, only the left hemisphere showed greater activity when subjects detected a face, consistent with the previous studies

on top-down object processing (Bar et al., 2006). Thus, to ensure sufficient power, we only used the right FFA, the right OFA and the left OFC as the ROIs for the present DCM analyses.

For simplicity, the three regions of interest (ROIs) were specified for each individual based on the coordinates of the peak activation obtained in the group analysis of the illusory face detection task. The center of each ROI (defined as a sphere of 6-mm radius) was located at the most significant voxel that was nearest to the peak coordinates in the group analysis for each individual. Subject-specific local maxima were constrained to lie within 12 mm (twice the width of the Gaussian smoothing kernel) of the group maximum in the appropriate SPM. In terms of individual ROI selection, of the twelve participants, we could not identify a ROI for the OFC in one participant. Thus, this participant's data were excluded from the DCM analysis.

4.6. Dynamic Causal Modeling (DCM)

DCM treats the brain as a dynamic input-state-output system. It is a nonlinear system identification procedure and uses Bayesian parameter estimation to draw inferences about the effective connectivity between different regions of the brain and the manner in which experimental conditions affect this connectivity (Friston et al., 2003). In DCM, three different sets of parameters are estimated: (i) the direct influence of an external stimulus on a given region, (ii) the intrinsic or latent connection between regions representing the interregional influence in the absence of any experimental manipulation, and (iii) the modulatory effects representing changes in intrinsic connection strength induced by the external experimental input (Friston et al., 2003). This modulatory effect can be used to identify neural networks that are involved when subjects detected faces in the pure noise images. The reported analysis adopted a two-stage procedure that is formally identical to the summary statistic approach used in a fixed effects analysis of neuroimaging data. The parameters of the first level (subject-specific) DCM models were taken to a second level (between-subjects) using the fixed effects approach (Acs and Greenlee, 2008; Booth et al., 2008; Stephan et al., 2010).

4.7. Choice of models

The neural model analyzed in DCM is experiment dependent and requires specific hypotheses. In other words, the user of DCM must specify the brain regions included in the model, which brain regions receive direct input from the presenting stimuli, the anatomical connectivity between regions, and which experimental conditions modulate connectivity (Penny et al., 2004). To this end, we proposed thirty-two plausible models. Each model contained three brain regions: the FFA, OFA, and OFC. These regions were included because previous studies found them to be involved in face or object processing in general, and top-down processing in particular.

As shown in Fig. 4, all models assumed that the FFA receives input from the OFA. It has been well established using DCM analysis that the right OFA has a direct feed-forward influence on the right FFA (Fairhall and Ishai, 2007; Summerfield et al., 2006). Thus, in our models, we assumed feed-forward connectivity from the OFA to FFA. As discussed in the Introduction, the

OFC has been found to be involved when encoding novel information (Frey and Petrides, 2000; Frey et al., 2004) and recognizing objects in a top-down manner (Bar et al., 2006; Kveraga et al., 2007a,b). Consequently, we tested the role of OFC in face top-down processing by assuming all possible connectivity architectures between the OFA, FFA and OFC. In all models, “face detection” trials were allowed to modulate all the connections among the three regions to examine the effect of top-down face detection. In addition, the presentation of a noise image, regardless of face detection, was also allowed to modulate all the connections among the three regions to test the bottom-up effects of the pure noise images.

The difference between Model a versus Model b in Fig. 4 was whether only the OFA received stimulus input, or whether both the OFA and OFC received stimulus input, respectively. This difference is motivated by the dual pathway face processing model proposed by Johnson (2005). He suggested that the face processing network may receive visual input from two separate pathways. One is a cortical pathway that receives visual input from the primary visual cortex to be processed in face-responsive areas of the fusiform gyrus and inferior occipital gyrus. The other is a subcortical pathway that receives visual input from regions such as the superior colliculus, pulvinar, and amygdala. This subcortical pathway appears to provide input to a number of cortical regions that are known to be involved in both top-down and bottom-up face processing. Thus, the present study compared these thirty-two models in regard to top-down face processing (see Fig. 4).

4.8. Selection of the optimal model

An optimal model is one that fits the data well, but does so with a minimum of free parameters (i.e., minimal model complexity). Therefore, to determine which of the 32 competing models was optimal, Bayesian model selection (Penny et al., 2004) was implemented using SPM5. This procedure identified the connectivity model showing the highest positive evidence at the individual subject level in the applied Bayesian framework (Raftery, 1995). Model evidence is calculated through the balance of model accuracy and model complexity (Penny et al., 2004). At the group level, group Bayes factors (GBFs) and binomial p values were used to determine the winning model based on the results of subject-specific Bayes factors for model comparison. Individual Bayes factors were calculated both with the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). If these differed in their outcome, then that subject's Bayes factor was not included in group results. GBFs were computed by multiplying the individual Bayes factors of the same model comparison across subjects (Stephan et al., 2007). Additionally, a conservative test for the reliability of the GBF procedure was calculated by finding the binomial p values for the probability of obtaining j or more Bayes factors > 1 in n subjects under an assumption of chance (Ethofer et al., 2006).

4.9. The averaged model

The three sets of connectivity parameters of the winning model from the individual subject models were then entered

into the DCM averaging routine provided by SPM5 to obtain a representative averaged model at the group level (Acs and Greenlee, 2008; Garrido et al., 2007). This allowed us to summarize the results from the different subject-specific DCMs.

Acknowledgments

This paper is supported by the Joint Research Fund for Overseas Chinese Young Scholars under Grant No. 30528027, the National Natural Science Foundation of China under Grant Nos. 60910006, 30873462, 30970769, 30970771, 30873462, 30870685, 60621001, 30970774, 60901064, 60902083, the Chair Professors of Cheung Kong Scholars Program of Ministry of Education of China, CAS Hundred Talents Program, Changjiang Scholars and Innovative Research Team in University (PCSIRT) under Grant No. IRT0645, the Project for the National Key Basic Research and Development Program (973) under Grant No. 2006CB705700, the Knowledge Innovation Program of the Chinese Academy of Sciences under Grant No. KG CX2-YW-129, KSCX2-YW-R-262, 863 program under Grant No. 2008AA01Z411, the Fundamental Research Funds for the Central Universities, NSF BCS-0843773, and NIH R01 HD046526.

REFERENCES

- Acs, F., Greenlee, M.W., 2008. Connectivity modulation of early visual processing areas during covert and overt tracking tasks. *Neuroimage* 41, 380–388.
- Allison, T., Puce, A., Spencer, D.D., McCarthy, G., 1999. Electrophysiological studies of human face perception. I: Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cereb. Cortex* 9, 415–430.
- Andrews, T.J., Schluppeck, D., 2004. Neural responses to Mooney images reveal a modular representation of faces in human visual cortex. *Neuroimage* 21, 91–98.
- Bar, M., 2009. The proactive brain: memory for predictions. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1235–1243.
- Bar, M., Kassam, K.S., Ghuman, A.S., Boshyan, J., Schmid, A.M., Dale, A.M., Hamalainen, M.S., Marinkovic, K., Schacter, D.L., Rosen, B.R., Halgren, E., 2006. Top-down facilitation of visual recognition. *Proc. Natl. Acad. Sci. U. S. A.* 103, 449–454.
- Bitan, T., Booth, J.R., Choy, J., Burman, D.D., Gitelman, D.R., Mesulam, M.M., 2005. Shifts of effective connectivity within a language network during rhyming and spelling. *J. Neurosci.* 25, 5397–5403.
- Booth, J.R., Mehdiratta, N., Burman, D.D., Bitan, T., 2008. Developmental increases in effective connectivity to brain regions involved in phonological processing during tasks with orthographic demands. *Brain Res.* 1189, 78–89.
- Calder, A.J., Young, A.W., 2005. Understanding the recognition of facial identity and facial expression. *Nat. Rev. Neurosci.* 6, 641–651.
- Ethofer, T., Anders, S., Erb, M., Herbert, C., Wiethoff, S., Kissler, J., Grodd, W., Wildgruber, D., 2006. Cerebral pathways in processing of affective prosody: a dynamic causal modeling study. *Neuroimage* 30, 580–587.
- Fairhall, S.L., Ishai, A., 2007. Effective connectivity within the distributed cortical network for face perception. *Cereb. Cortex* 17, 2400–2406.
- Frey, S., Petrides, M., 2000. Orbitofrontal cortex: a key prefrontal region for encoding information. *Proc. Natl. Acad. Sci. U. S. A.* 97, 8723–8727.
- Frey, S., Kostopoulos, P., Petrides, M., 2004. Orbitofrontal contribution to auditory encoding. *Neuroimage* 22, 1384–1389.
- Friston, K.J., Glaser, D.E., Henson, R.N., Kiebel, S., Phillips, C., Ashburner, J., 2002. Classical and Bayesian inference in neuroimaging: applications. *Neuroimage* 16, 484–512.
- Friston, K.J., Harrison, L., Penny, W., 2003. Dynamic causal modelling. *Neuroimage* 19, 1273–1302.
- Garrido, M.I., Kilner, J.M., Kiebel, S.J., Stephan, K.E., Friston, K.J., 2007. Dynamic causal modelling of evoked potentials: a reproducibility study. *Neuroimage* 36, 571–580.
- Gauthier, I., Tarr, M.J., Moylan, J., Skudlarski, P., Gore, J.C., Anderson, A.W., 2000. The fusiform “face area” is part of a network that processes faces at the individual level. *J. Cogn. Neurosci.* 12, 495–504.
- Gosselin, F., Schyns, P.G., 2001. Bubbles: a technique to reveal the use of information in recognition tasks. *Vision Res.* 41, 2261–2271.
- 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.
- Hasson, U., Hendler, T., Ben Bashat, D., Malach, R., 2001. Vase or face? A neural correlate of shape-selective grouping processes in the human brain. *J. Cogn. Neurosci.* 13, 744–753.
- Haxby, J.V., Hoffman, E.A., Gobbini, M.I., 2000. The distributed human neural system for face perception. *Trends Cogn. Sci.* 4, 223–233.
- Ishai, A., 2007. Sex, beauty and the orbitofrontal cortex. *Int. J. Psychophysiol.* 63, 181–185.
- Ishai, A., 2008. Let's face it: it's a cortical network. *Neuroimage* 40, 415–419.
- Ishai, A., Ungerleider, L.G., Haxby, J.V., 2000. Distributed neural systems for the generation of visual images. *Neuron* 28, 979–990.
- Ishai, A., Schmidt, C.F., Boesiger, P., 2005. Face perception is mediated by a distributed cortical network. *Brain Res. Bull.* 67, 87–93.
- Johnson, M.H., 2005. Subcortical face processing. *Nat. Rev. Neurosci.* 6, 766–774.
- Kanwisher, N., Yovel, G., 2006. The fusiform face area: a cortical region specialized for the perception of faces. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 361, 2109–2128.
- 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.
- Kim, J.J., Andreasen, N.C., O'Leary, D.S., Wiser, A.K., Ponto, L.L., Watkins, G.L., Hichwa, R.D., 1999. Direct comparison of the neural substrates of recognition memory for words and faces. *Brain* 122 (Pt 6), 1069–1083.
- Kranz, F., Ishai, A., 2006. Face perception is modulated by sexual preference. *Curr. Biol.* 16, 63–68.
- Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S., Baker, C.I., 2009. Circular analysis in systems neuroscience: the dangers of double dipping. *Nat. Neurosci.* 12, 535–540.
- Kveraga, K., Boshyan, J., Bar, M., 2007a. Magnocellular projections as the trigger of top-down facilitation in recognition. *J. Neurosci.* 27, 13232–13240.
- Kveraga, K., Ghuman, A.S., Bar, M., 2007b. Top-down predictions in the cognitive brain. *Brain Cogn.* 65, 145–168.
- Li, J., Liu, J., Liang, J., Zhang, H., Zhao, J., Huber, D.E., Rieth, C.A., Lee, K., Tian, J., Shi, G., 2009. A distributed neural system for top-down face processing. *Neurosci. Lett.* 451, 6–10.
- Liu, J., Tian, J., Li, J., Gong, Q., Lee, K., 2009. Similarities in neural activations of face and Chinese character discrimination. *Neuroreport* 20, 273–277.
- Liu, J., Li, J., Zhang, H., Rieth, C.A., Huber, D.E., Li, W., Lee, K., Tian, J., 2010. Neural correlates of top-down letter processing. *Neuropsychologia* 48, 636–641.

- Mechelli, A., Price, C.J., Friston, K.J., Ishai, A., 2004. Where bottom-up meets top-down: neuronal interactions during perception and imagery. *Cereb. Cortex* 14, 1256–1265.
- O'Craven, K.M., Kanwisher, N., 2000. Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *J. Cogn. Neurosci.* 12, 1013–1023.
- O'Doherty, J., Kringelbach, M.L., Rolls, E.T., Hornak, J., Andrews, C., 2001. Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat. Neurosci.* 4, 95–102.
- Penny, W.D., Stephan, K.E., Mechelli, A., Friston, K.J., 2004. Comparing dynamic causal models. *Neuroimage* 22, 1157–1172.
- Pitcher, D., Walsh, V., Yovel, G., Duchaine, B., 2007. TMS evidence for the involvement of the right occipital face area in early face processing. *Curr. Biol.* 17, 1568–1573.
- Plailly, J., Howard, J.D., Gitelman, D.R., Gottfried, J.A., 2008. Attention to odor modulates thalamocortical connectivity in the human brain. *J. Neurosci.* 28, 5257–5267.
- Raftery, A.E., 1995. Bayesian model selection in social research. *Sociol. Methodol.* 25, 111–163.
- Rossion, B., Caldara, R., Seghier, M., Schuller, A.M., Lazeyras, F., Mayer, E., 2003. A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain* 126, 2381–2395.
- Rotshtein, P., Henson, R.N., Treves, A., Driver, J., Dolan, R.J., 2005. Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nat. Neurosci.* 8, 107–113.
- Schiltz, C., Rossion, B., 2006. Faces are represented holistically in the human occipito-temporal cortex. *Neuroimage* 32, 1385–1394.
- Stephan, K.E., Weiskopf, N., Drysdale, P.M., Robinson, P.A., Friston, K.J., 2007. Comparing hemodynamic models with DCM. *Neuroimage* 38, 387–401.
- Stephan, K.E., Penny, W.D., Moran, R.J., den Ouden, H.E., Daunizeau, J., Friston, K.J., 2010. Ten simple rules for dynamic causal modeling. *Neuroimage* 49, 3099–3109.
- Summerfield, C., Egner, T., Greene, M., Koechlin, E., Mangels, J., Hirsch, J., 2006. Predictive codes for forthcoming perception in the frontal cortex. *Science* 314, 1311–1314.
- Wild, H.A., Busey, T.A., 2004. Seeing faces in the noise: stochastic activity in perceptual regions of the brain may influence the perception of ambiguous stimuli. *Psychon. Bull. Rev.* 11, 475–481.
- Zhang, H., Liu, J., Huber, D.E., Rieth, C.A., Tian, J., Lee, K., 2008. Detecting faces in pure noise images: a functional MRI study on top-down perception. *Neuroreport* 19, 229–233.
- Zhang, H., Tian, J., Liu, J., Li, J., Lee, K., 2009. Intrinsically organized network for face perception during the resting state. *Neurosci. Lett.* 454, 1–5.