

# Unfolding the Spatial and Temporal Neural Processing of Lying about Face Familiarity

Delin Sun<sup>1,2</sup>, Tatia M.C. Lee<sup>1,2,3</sup> and Chetwyn C.H. Chan<sup>4</sup>

<sup>1</sup>Laboratory of Neuropsychology, <sup>2</sup>Laboratory of Cognitive Affective Neuroscience, <sup>3</sup>The State Key Laboratory of Brain and Cognitive Sciences, The University of Hong Kong, Hong Kong, China and <sup>4</sup>Department of Rehabilitation Sciences, Applied Cognitive Neuroscience Laboratory, The Hong Kong Polytechnic University, Hong Kong

Address correspondence to Chetwyn C.H. Chan, Department of Rehabilitation Sciences, Applied Cognitive Neuroscience Laboratory, The Hong Kong Polytechnic University, Hum Hom, Kowloon, Hong Kong. Email: chetwyn.chan@polyu.edu.hk

Delin Sun and Tatia M.C. Lee contributed equally to this work

**To understand the neural processing underpinnings of deception, this study employed both neuroimaging (functional magnetic resonance imaging, fMRI) and neurophysiological (event-related potential, ERP) methodologies to examine the temporal and spatial coupling of the neural correlates and processes that occur when one lies about face familiarity. This was performed using simple directed lying tasks. According to cues provided by the researchers, the 17 participants were required to respond truthfully or with lies to a series of faces. The findings confirmed that lie and truth conditions are associated with different fMRI activations in the ventrolateral, dorsolateral, and dorsal medial-frontal cortices; premotor cortex, and inferior parietal gyrus. They are also associated with different amplitudes within the time interval between 300 and 1000 ms post face stimulus, after the initiation (270 ms) of face familiarity processing. These results support the cognitive model that suggests representations of truthful information are first aroused and then manipulated during deception. Stronger fMRI activations at the left inferior frontal gyrus and more positive-going ERP amplitudes within [1765, 1800] ms were observed in the contrast between lie and truth for familiar than for unfamiliar faces. The fMRI and ERP findings, together with ERP source reconstruction, clearly delineate the neural processing of face familiarity deception.**

**Keywords:** deception, ERP, face familiarity, fMRI, source reconstruction

## Introduction

Deception is a ubiquitous social phenomenon that has been vigorously researched in an attempt to understand its behavioral and neural underpinnings. In the past 2 decades, many studies have employed neuroimaging (e.g., functional magnetic resonance imaging, fMRI) or neurophysiological methodologies (e.g., scalp event-related potential, ERP) to understand the neural mechanisms of deception. As a result, fMRI deception studies have revealed that a neural network involving the frontal and parietal regions is related to the execution of deception (Spence et al. 2001; Langleben et al. 2002; Lee et al. 2002, 2005; Christ et al. 2009). ERP deception studies have also shown that scalp signals around or after 300 ms poststimulus onset robustly mark the difference between deceptive and truthful responses (Johnson and Rosenfeld 1992; Johnson et al. 2003, 2008). The knowledge gained from both methods has largely increased our understanding of the neural processing that occurs when one tells a lie. However, the inherent limitations of neuroimaging and neurophysiological methodologies prevent an in-depth understanding of the dynamic responses in the brain during deception. On the one hand, the low temporal resolution of the fMRI method makes findings

appear as snapshots of brain activation compressed across time. On the other hand, calculating the inner-skull sources of the scalp's electrical signals may lead to several sets of possible sources, which could hardly be differentiated mathematically by merely using neurophysiological data (Katznelson 1981; Joyce and Rossion 2005). Overall, an understanding of when and where deception-related brain activation appears has yet to be developed.

To fill in the abovementioned theoretical gap, this study employed a simple “directed lying” experimental task, during which both fMRI and ERP data were acquired. The participants were asked to view familiar or unfamiliar faces and to respond according to the cues of “truth” or “lie.” This paradigm focuses on an important component of deception: manipulating truthful information without interference of other lie-related cognitive/affective processes, such as decision-making before action (to lie or not) and feedback evaluation after the outcome (being detected or not). The “directed lie” paradigm makes it easier to identify the neural correlates (both fMRI and ERP) of cognitive manipulation during deception and, hence, contributes to successful ERP source analyses since both spatial and temporal results were clearly obtained. All participants were invited to take part in both the fMRI and ERP experiments. A cognitive model of deception consisting of 2 consecutive phases was employed for this study. The first phase activated the neural representation of truth. The other inhibited the truth and formulated the response to achieve the goal of instilling a false belief in others. This model was inspired by several previous fMRI studies (Spence et al. 2004) that consistently reported the involvement of inferior prefrontal cortex in deception. This brain region has been proved to play roles in behavioral regulation (Aron et al. 2004) and, thus, implies that the neural correlates of truthful information were evoked beforehand.

The neural processing of face stimuli has been well documented (Gobbini and Haxby 2006, 2007). Previous fMRI studies have shown that familiar stimuli elicit significant brain activation in the precuneus/posterior cingulate gyrus, posterior superior temporal sulcus/temporal-parietal junction, anterior temporal gyrus, and anterior paracingulate cortex (Gobbini and Haxby 2006, 2007). ERP studies have also shown that perceptions of familiar faces, compared with those of unfamiliar faces, were correlated with 2 enlarged ERP components. One is N250 (Schweinberger et al. 2002; Herzmann et al. 2004; Pierce et al. 2011), which peaks (negative-going) approximately 250 ms after stimulus onset at medial-frontal sites. The other is P600 (Eimer 2000a, 2000b), which is most

prominent (positive-going) between 400 and 600 ms poststimulus at posterior sensors. The scalp topography of the ERP amplitudes depends on the channels of reference employed, but the temporal information provided by the ERP methodology is more stable and important than the 2-dimensional (2D) topography.

Following the recognition of face familiarity, cognitive manipulation of the truthful information occurs. Consistent findings across previous fMRI deception studies have shown that a frontal-parietal brain network, including the ventrolateral, dorsolateral, and dorsal medial-frontal cortices; premotor cortex, and inferior parietal gyrus, was activated during deception (Lee et al. 2002, 2005; Phan et al. 2005; Christ et al. 2009). In a recent fMRI study by Bhatt et al. (2009), researchers investigated the neural correlates of concealing or revealing the identities of individuals seen in study sets. Results showed more activation for deception than for telling truth conditions in the right middle frontal gyrus, red nucleus, inferior frontal gyrus, supramarginal gyrus, superior frontal gyrus, anterior cingulate gyrus, dorsolateral prefrontal cortex, and bilateral precuneus. The findings were largely consistent with previous neuroimaging results on deception and suggested that deception involving face identity also requires the suppression of truthful information. At the same time, many ERP studies have indicated that the late positive complex, which is a positive-going deflection appearing between 300 and 1000 ms after the onset of stimuli, was smaller for deceptive than for truthful conditions (Johnson et al. 2003, 2004) at posterior sites. We hypothesized that lie and truth conditions were associated with different fMRI activation in the ventrolateral, dorsolateral, and dorsal medial-frontal cortices; premotor cortex, and inferior parietal gyrus and were also associated with different amplitudes within the time interval between 300 ms and 1000 ms post face stimulus after the initiation of face familiarity processing.

## Materials and Methods

### Subjects

Seventeen Chinese males (age: 25–40 years; education level: 12–22 years), recruited from a local community, participated in this study. They were invited to participate in both the fMRI and ERP experiments. Among them, 12 completed both experiments. Two and 3 of them only volunteered for the fMRI and the ERP experiments, respectively. Only male participants were recruited in order to exclude possible gender differences in terms of facial recognition (Proverbio et al. 2006; Hoffmann et al. 2010) and/or deceptive behaviors (Farrow et al. 2003). All the participants were right-handed, as assessed by the Edinburgh Handedness Inventory (Oldfield 1971), and had normal or corrected-to-normal vision. None of them reported having a history of neurological or mental disorders. All the participants provided written informed consent. The study was conducted under the approval of the Institutional Review Board of the University of Hong Kong and the Hong Kong West Cluster Hospital Authority.

### Materials

There were 2 sets of photos: (1) 50 faces of the participants' acquaintances in the local community; and (2) 50 faces of strangers. Both sets of stimuli were male Chinese faces with neutral facial expressions and were matched for age. All the stimuli were transformed into grayscale. The luminance, contrast, and resolution of the photos were adjusted to approach equivalence using the Adobe Photoshop software (San Jose, CA, USA).

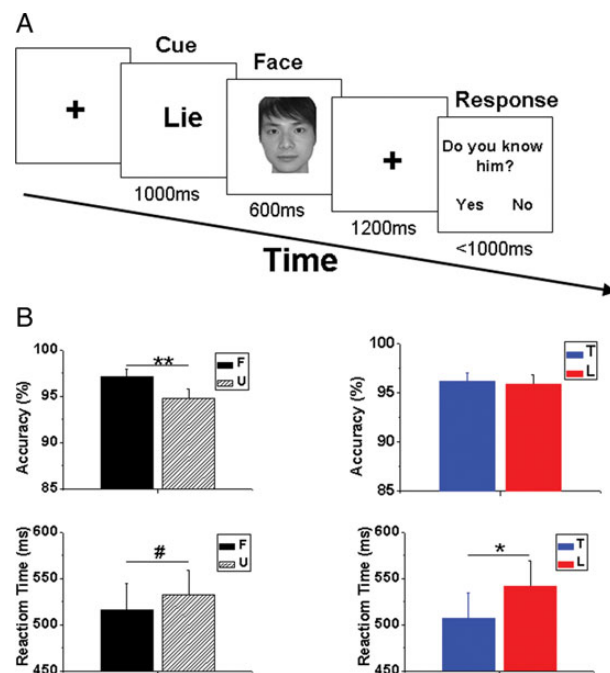
Prior to the experimental task, each participant was asked to give the name of the person in the photo sets and to rate the familiarity

(1 = totally unfamiliar, 9 = very familiar) and valence (1 = very negative, 9 = very positive) of the face on a 9-point scale for both sets of photos. The familiar face stimuli recruited in the formal task were those named correctly with a familiarity score over 5 and a valence score between 4 and 6. The unfamiliar face stimuli were chosen if the participant could not name the person, provided a zero familiarity score, and gave a valence score between 4 and 6. Finally, 30 familiar and unfamiliar faces were chosen for each participant in the formal task.

### Experimental Task

The task used a directed lie paradigm, in which the participant was instructed to respond to the question, "Do you know him?" according to the Chinese text cue (truth or lie) presented beforehand. If the face was of an acquaintance, the correct answer was "yes" for the cue truth and "no" for the cue lie. If the face was of a stranger, the correct response was no for the cue truth and yes for the cue lie. Therefore, there were a total of 4 conditions: truth and familiar face (TF), truth and unfamiliar face (TU), lie and familiar face (LF), as well as lie and unfamiliar face (LU). The participant was asked to respond as accurately and quickly as possible. If the cue was lie, the participant had to lie as convincingly as possible in order to deceive the computer system, which was monitoring his neural activities. Before the formal task, each participant practiced repeatedly until he reached 90% accuracy in performance. During practice, the familiar and unfamiliar faces were replaced by photos of 10 famous public figures and 10 strangers not included in the formal task, respectively.

In each trial, a text cue was presented for 1000 ms, followed by the appearance of a face for 600 ms, which was then replaced by a fixation for 1200 ms. Finally, a text cue (Do you know him?) appeared to prompt a response within 1000 ms (see Fig. 1). During the fMRI experiment, the inter-trial-interval (ITI) varied randomly from 1000 to 3000 ms to allow for signal differentiation of the rapid event-related trials. In the ERP experiment, the ITI was shortened (between 800 and 1200 ms) to reduce the length of the data recordings, because long ERP experiments cause fatigue and inattention. The formal task was comprised of 160 trials in the fMRI experiment and 480 in the ERP experiment. Each condition had 40 (fMRI) or 120 (ERP) trials. More



**Figure 1.** (A) Schematic paradigm of the directed lie task. (B) Behavioral performance. F: familiar face; U: unfamiliar face; T: response of truth; L: response of lie. #  $0.05 < P < 0.06$ , \* $P < 0.05$ , \*\* $P < 0.001$ . Error bar denotes standard error of the mean (SEM).

trials were recruited in the ERP experiment to compensate for the relatively low signal-to-noise ratio in ERP data recordings. The order of trials was pseudorandomized for each participant and was counterbalanced across participants. Response keys for yes and no were counterbalanced across participants as well. All the fMRI experiments were performed before the ERP experiments because the time slots were reserved beforehand. To reduce the order effect of the 2 experiments, participants were asked to take part in the ERP experiments long after (10–12 months) the fMRI experiments.

### **fMRI Image Acquisition and Analysis**

All the fMRI data were acquired by a 3-T Philips Achieva scanner with a standard head coil. A  $T_1$ -weighted, spin-echo pulse sequence (time repetition, TR = 7 ms; time echo, TE = 3.2 ms; slice thickness = 1 mm) was employed to obtain anatomical images in sagittal planes.  $T_2^*$ -weighted, gradient-echo, echo-planar imaging pulse sequence (TR = 1800 ms, TE = 30 ms, matrix size =  $64 \times 64$ , field of view =  $230 \text{ mm} \times 230 \text{ mm}$ , flip angle =  $90^\circ$ , slice thickness = 4 mm) parallel to the anterior commissure–posterior commissure plane was used to acquire the functional images.

The fMRI data preprocessing was done by the SPM8 software (Wellcome Department of Cognitive Neurology, London, UK). The initial 6 volumes of each run were discarded to ensure the magnetization reached equilibrium. The remaining functional scans of each participant were first realigned and resliced to form a mean image. Temporal differences in image acquisition were then corrected by slice timing. Through coregistration, the functional images were matched with the participant's own anatomical image, which was segmented into gray/white matter. Then, the functional images were normalized to a standard anatomical template of East Asian brains (Talairach et al. 1988). The normalized functional images were smoothed by an 8-mm full-width half-maximum Gaussian filter (FWHM). The hemodynamic response function was used to model the fMRI signal, and a high-pass filter was set at 128 s to reduce low frequency noise.

Four contrasts (only for the trials with correct response) for each participant were entered into a flexible factorial model with 2 within-group factors [i.e., Type of Response (lie vs. truth) and Face (familiar vs. unfamiliar)]. Results were voxel-level height thresholded at  $P < 0.001$  and survived after peak- or cluster-level family wise error (FWE) correction,  $P < 0.05$ .

### **ERP Data Recording and Analysis**

Electroencephalogram (EEG) signals were recorded by a fabric cap (Neuroscan Company) embedded with 128 Ag–AgCl electrodes. All channel recordings were referenced to the electrode at the left mastoid. Channel impedances were kept  $< 5 \text{ k}\Omega$ . The scalp signals were preprocessed by the Scan 4.3 software (Neuroscan Company). All EEG channels were re-referenced to the average of all EEG channels. The signals were band-pass filtered between 0.1 and 30 Hz using a zero phase-shift digital filter. Ocular artifacts were mathematically corrected (Gratton et al. 1983). Continuous recordings were cut into epochs from  $-200$  to  $1847 \text{ ms}$  after the presentation of a face. Baseline correction was then performed by recruiting the data from  $-200$  to  $0 \text{ ms}$  as baseline. Artifact rejection was performed, so that signals exceeding  $\pm 100 \mu\text{V}$  in any given epoch were automatically discarded. Trials (only for those with correct responses) of the same condition were averaged, so that each participant had 4 ERP datasets (TF, TU, LF, and LU).

The 4 ERP datasets per participant were then transformed into SPM8 file format and down-sampled to 200 Hz to increase processing speed. The ERP amplitude at each channel per time point was used to fit the intensity of a 3D image by linear interpolation (Kiebel and Friston 2004a, 2004b). The cross-section (i.e., the  $x$ – $y$  plane, where the  $x$  dimension reflects “left–right” and the  $y$  dimension denotes “anterior–posterior” in the scalp) of this 3D image represents the horizontal projection of a standard 128-channel Neuroscan EEG cap, and the  $z$  dimension represents the timeline. Unlike in the fMRI method, a voxel was defined as  $4.25 \text{ mm}$  ( $x$  dimension in the scalp)  $\times$   $5.38 \text{ mm}$  ( $y$  dimension in the scalp)  $\times$   $5 \text{ ms}$  ( $z$  dimension, i.e., timeline). The 3D image data were smoothed by an FWHM of [9, 9 mm, 20 ms] (Henson et al. 2008; Boly et al. 2011).

Similar to the statistical method used for the fMRI analysis, 4 contrast images, that is, TF, TU, LF, and LU, from each participant were entered into a flexible factorial model with 2 within-group factors, that is, Type of Response and Face. Results were voxel-level height thresholded at  $P < 0.001$  and survived after peak- or cluster-level FWE correction,  $P < 0.05$ .

### **ERP Source Reconstruction with fMRI Spatial Priors**

The ERP source reconstruction analyses were carried out using the group inversion (imaging method) module in the SPM8 software. This approach considered the scalp EEG data as a reflection of the activities of a large number of dipolar sources distributed over the cortical sheet with fixed locations and orientations. It restricted the activated sources to being the same in all participants, with only the degree of activation varying (Litvak and Friston 2008). The first step of this analysis was to match an SPM standard template head model to each individual's EEG electrode positions. The boundary element model was then used for the forward model calculation. Datasets related to the contrast of interest (i.e., conditions F and U for the contrast F – U; conditions T and L for the contrast L – T; conditions TF + LU and TU + LF for the contrast F (L – T) – U(L – T)) from each participant were inverted by the multiple sparse priors approach (Friston et al. 2008) for each time window of interest according to the scalp ERP group statistical results. The MNI coordinates of peak significance in the fMRI group analysis were recruited to serve as spatial priors that could further restrict the source locations (Henson et al. 2010). For each participant, the intensity of source activity for each task condition within the time window of interest was converted into the brightness of a 3D image that could be overlapped onto a standard MNI brain template. Last, after spatial smoothing with an FWHM of 8 mm, the images were entered into paired  $t$ -tests. Because we already knew the location of the sources from fMRI results, we recruited a lenient threshold (voxel-level height thresholded at  $P < 0.05$ , uncorrected, with an extent threshold of  $> 25$  voxels) to check whether the brain regions of interest were activated within the time windows of interest.

## **Results**

### **Behavioral Performance**

To investigate how the behavioral performance was influenced by the Type of Response (lie vs. truth) and Face (familiar vs. unfamiliar) factors and whether the performance varied between fMRI and ERP experiments, we analyzed the behavioral data of the 12 volunteers who participated in both experiments. The reaction times and response accuracies of each condition recorded in the 2 studies were entered into a 2 (fMRI vs. ERP) by 2 (lie vs. truth) by 2 (familiar vs. unfamiliar) repeated-measures analysis of variance using SPSS 13 (SPSS, Inc.).

Results showed that participants responded more slowly ( $F_{1, 11} = 6.690$ ,  $P = 0.025$ ) when the cue was lie ( $539.8 \pm 29.1 \text{ ms}$ ) than when it was truth ( $506.7 \pm 28.2 \text{ ms}$ ). Moreover, there was a trend ( $F_{1, 11} = 4.784$ ,  $P = 0.051$ ) for quicker responses to familiar ( $515.0 \pm 29.1 \text{ ms}$ ) than to unfamiliar ( $531.4 \pm 27.3 \text{ ms}$ ) faces. No significant interaction effect was found between Type of Response and Face ( $F_{1, 11} = 1.371$ ,  $P = 0.266$ ) for the reaction time. Response accuracies were not significantly different ( $F_{1, 11} = 0.130$ ,  $P = 0.725$ ) between truth ( $96.2 \pm 0.9\%$ ) and lie ( $95.9 \pm 1.0\%$ ) responses, but were higher ( $F_{1, 11} = 14.023$ ,  $P = 0.003$ ) for familiar ( $97.3 \pm 0.9\%$ ) than for unfamiliar ( $94.8 \pm 1.1\%$ ) faces. No significant interaction effect was found between the Type of Response and Face ( $F_{1, 11} = 1.534$ ,  $P = 0.241$ ) for the accuracy (see Fig. 1B).

These data also illustrated whether behavioral performance varied between the 2 experiments. Participants responded quicker ( $F_{1, 11} = 186.718$ ,  $P < 0.001$ ) in the ERP (mean  $\pm$  standard

error =  $287.4 \pm 18.1$  ms) than those in the fMRI experiment ( $759.1 \pm 42.8$  ms). This might be attributed to the stronger training effect in the ERP than in the fMRI experiment, because more trials were conducted in the ERP experiment. More importantly, the difference in reaction time between studies did not interact with either Type of Response ( $F_{1, 11} = 3.611$ ,  $P = 0.084$ ) or Face ( $F_{1, 11} = 2.677$ ,  $P = 0.130$ ). This suggests that the possible training effect did not influence the strategy associated with face familiarity and deception processing. To further support this idea, the reaction time for lies was significantly correlated between the 2 experiments for both familiar (Pearson correlation coefficient = 0.582,  $P = 0.047$ ) and unfamiliar (Pearson correlation coefficient = 0.678,  $P = 0.015$ ) faces. Participants' response accuracies were not found to be different between experiments ( $F_{1, 11} = 1.688$ ,  $P = 0.220$ ). The difference in accuracy between experiments did not interact with the Type of Response ( $F_{1, 11} = 2.712$ ,  $P = 0.128$ ) or Face ( $F_{1, 11} = 1.389$ ,  $P = 0.263$ ).

### fMRI Data

Fourteen participants (12 of whom participated in both experiments) were involved in this analysis. The results are listed in Table 1. Compared with unfamiliar faces, familiar face stimuli were associated with stronger fMRI activations at the left precuneus, medial-frontal gyrus, middle temporal gyrus, and right inferior and superior temporal gyrus. Unfamiliar faces were associated with stronger activation at the right inferior frontal gyrus (see Fig. 2A).

Deceptive responses elicited stronger fMRI activation at the left inferior, middle, and superior frontal gyrus; left inferior parietal gyrus, and right inferior frontal gyrus (see Fig. 3A). Note that the effect of Type of Response survived very strict thresholds.

**Table 1**  
fMRI results (voxel-level height threshold  $P < 0.001$ , peak- or cluster-level  $P < 0.05$ , FWE correction)

Brain area	Cluster	Z	MNI coordinates		
			x	y	z
<b>Main effect of Face</b>					
F > U					
L precuneus (BA31)	2079	6.953	-8	-56	20
L medial-frontal gyrus (BA10)	819	6.197	-2	60	-4
R inferior temporal gyrus (BA21)	184	5.370	54	-6	-24
L middle temporal gyrus (BA39)	425	5.168	-44	-68	20
R superior temporal gyrus (BA39)	201	3.945	50	-58	12
F < U					
R inferior frontal gyrus (BA44/45)	204	4.521	46	14	2
<b>Main effect of Type of Response<sup>a</sup></b>					
L > T					
L inferior frontal gyrus (BA47)	536	6.786	-48	18	0
L inferior parietal gyrus (BA40)	294	6.105	-48	-54	52
L middle frontal gyrus (BA6)	88	5.859	-42	10	50
L superior frontal gyrus (BA6/8)	306	5.519	8	16	62
R inferior frontal gyrus (BA47)	42	5.333	50	16	0
L < T					
None					
<b>Interaction effect between Type of Response and Face</b>					
F(L - T) > U(L - T)					
L inferior frontal gyrus (BA47)	202	4.300	-46	18	0
F(L - T) < U(L - T)					
None					

Cluster represents the number of voxels within the significant cluster. Z indicates the z value. F: familiar face; U: unfamiliar face; T: truthful response; L: deceptive response.

<sup>a</sup>Results of the main effect of Type of Response were all voxel-level thresholded at  $P < 0.05$  (FWE correction) with cluster size of  $> 25$  voxels.

An interaction effect between Face and Type of Response was also found at the left inferior frontal gyrus (see Fig. 4A). Further analysis of this region showed that brain activations (percent signal changes) for the contrast L - T were larger for familiar than for unfamiliar faces ( $t_{(13)} = 5.603$ ,  $P < 0.001$ ). The brain activation for the contrast L - T was positively correlated with the reaction time for the same contrast when the faces were unfamiliar ( $R = 0.579$ ,  $P = 0.030$ ). This correlation did not apply when the faces were familiar ( $R = 0.258$ ,  $P = 0.373$ ).

### ERP Data

Fifteen participants (12 of whom participated in both experiments) were included in this analysis. The results are listed in Table 2. Within the time interval of [270, 670] ms post face onset, familiar faces were associated with more positive-going amplitudes at frontal-central sites and more negative-going amplitudes at bilateral temporal sites (see Fig. 2B).

Compared with truthful responses, lies were associated with less negative-going amplitudes at left occipital-parietal sites within 2 separate time windows post face onset, that is, [425, 510] and [790, 930] ms post face onset (see Fig. 3B).

A significant interaction effect between Face and Type of Response was found at left temporal sites within a short interval, that is [1765, 1800] ms, just before the text cue response (see Fig. 4B). Further analysis of this region showed that the amplitudes extracted from channel 5 (which was nearest to the peak significance) at left temporal scalp areas averaged [1765, 1800] ms for the contrast L - T and were larger ( $t_{(14)} = 5.191$ ,  $P < 0.001$ ) for familiar ( $0.753 \pm 0.227$   $\mu$ V) than for unfamiliar faces ( $-0.379 \pm 0.218$   $\mu$ V). The pattern of the interaction effect was similar in the ERP and fMRI results. The ERP amplitudes for the condition LF were positively correlated with the fMRI percent signal changes extracted from the ROI at the left inferior frontal gyrus for the condition LF ( $R = 0.745$ ,  $P = 0.005$ ).

### ERP Source Reconstruction

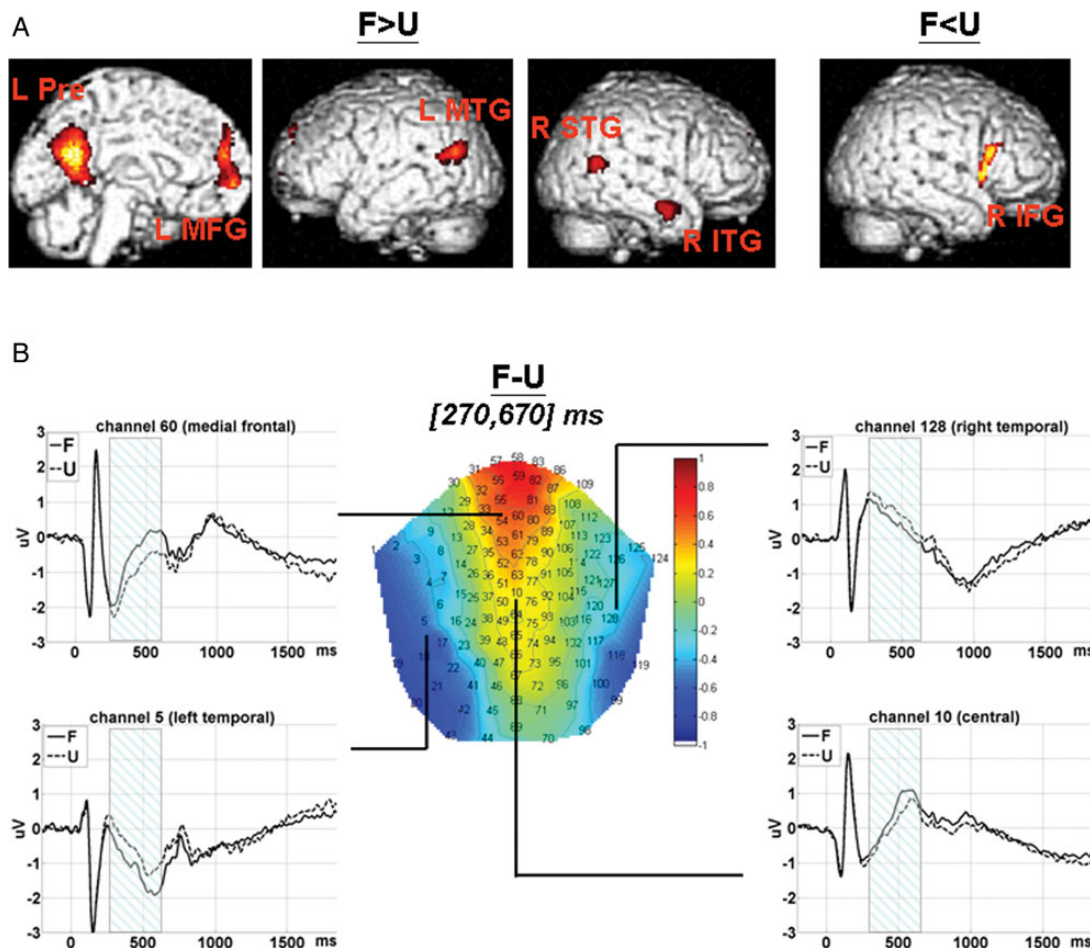
The statistical results for the source intensities were shown in Figure 5. Since the ERP data showed that the early significant result of contrast L - T ([425, 510] ms) appeared within the interval for the significant result of contrast F - U ([270, 670] ms), we calculated the source intensities for F - U within 3 consecutive time windows (i.e., [270, 425], [425, 510], and [510, 670] ms) to see the brain activation differentiating familiar and unfamiliar faces. Results showed that familiar faces were associated with stronger intensities within varying time windows: [270, 425] ms for the left middle temporal gyrus, medial-frontal gyrus, and right superior temporal gyrus; [425, 510] ms for the left precuneus; and [510, 670] ms for the left middle temporal gyrus and precuneus.

Within the time window of [425, 510] ms, deceptive responses were found to be associated with weaker intensities at the left inferior frontal gyrus. On the other hand, within the interval of [790, 930] ms, stronger intensities were found for deceptive responses at the left inferior parietal gyrus.

Within the interval of [1765, 1800] ms, familiar faces were associated with stronger intensities at the left inferior frontal gyrus for the contrast L - T.

### Discussion

Consistent with our a priori hypothesis, the results of this study have clearly shown that, on one hand, lies elicited stronger fMRI



**Figure 2.** Main effect of Face (F vs. U). (A) fMRI SPM T-map. Voxels within the significant clusters were height threshold at  $P < 0.001$  and survived peak- or cluster-level  $P < 0.05$  FWE correction. L Pre: left precuneus; L MFG: left medial-frontal gyrus; L MTG: left middle temporal gyrus; R ITG: left inferior temporal gyrus; R STG: right superior temporal gyrus; R IFG: right inferior frontal gyrus. (B) ERP 2D topography for the contrast (F – U) within [270, 670] ms and ERP waveforms (time-locked to face onset) extracted from channels at medial-frontal (60), central (10), and left (5)/right (128) temporal scalp areas are shown, respectively. The shadowed intervals indicate the time windows during which amplitudes for F were significantly different from those for U. F: familiar face; U: unfamiliar face.

activation at the bilateral inferior frontal gyrus (i.e., ventrolateral prefrontal cortex), left superior (i.e., dorsal medial), middle frontal gyrus (i.e., premotor cortex), and inferior parietal gyrus. On the other hand, lies were also associated with less negative-going amplitudes at left occipital-parietal sites in 2 time windows (i.e., [425, 510] and [790, 930] ms). Both intervals were behind the initiation (270 ms) of face familiarity processing. The findings support the a priori cognitive model for deception, meaning the representation of truthful information was first aroused and then manipulated during deception.

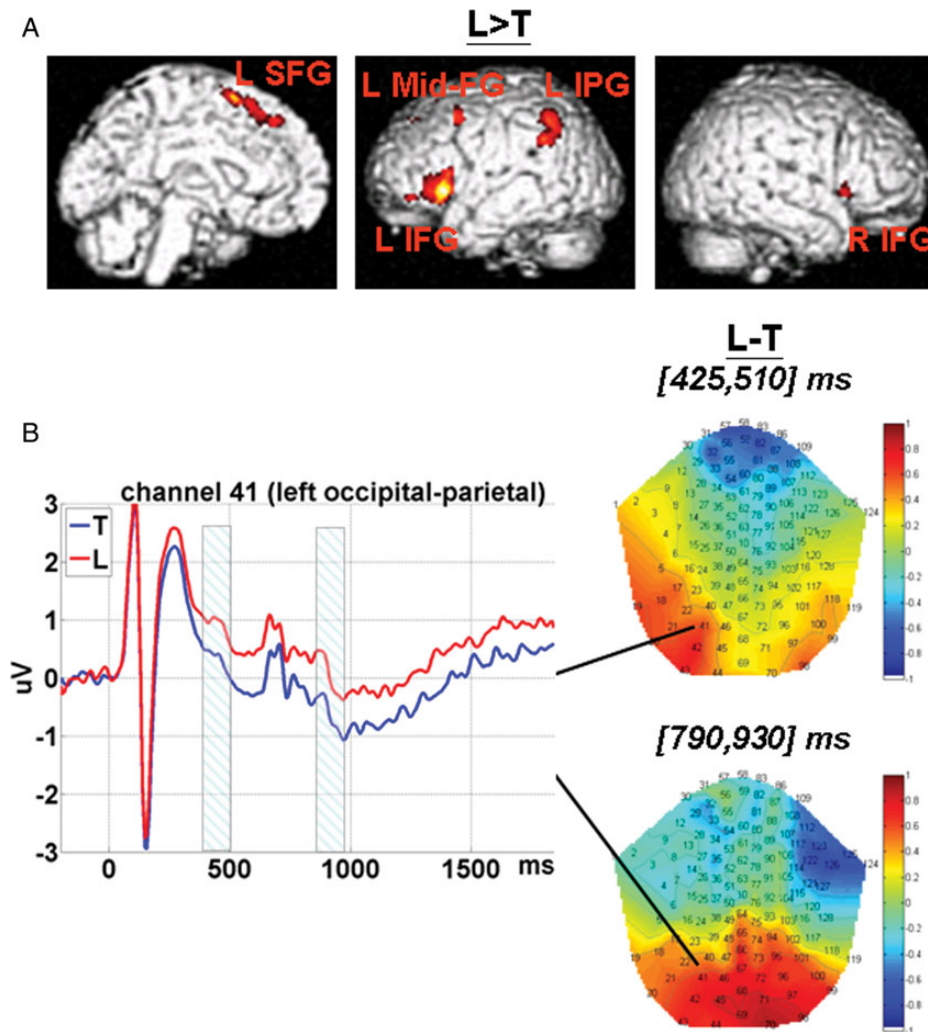
To the best of our knowledge, this is the first project on deception employing both neuroimaging and neurophysiological methodologies using the same experimental task on one sample. The findings help unfold the neural processing of deception into consecutive but distinct subprocesses. The findings provide insights into both spatial and temporal information on the neural processing of deception.

Viewing familiar faces was found to elicit stronger fMRI signals at the precuneus, medial-frontal gyrus (i.e., anterior paracingulate cortex), and superior/middle/inferior temporal gyrus (i.e., posterior superior temporal sulcus and anterior temporal gyrus). The findings were consistent with results from previous fMRI studies on face familiarity (Gobbini and Haxby 2006,

2007). Familiar faces were associated with more positive-going amplitudes at frontal-central sites and more negative-going amplitudes at bilateral temporal sites during the time course of [270, 670] ms, which covered the time windows of N250 and P600 in the literature (Eimer 2000a, 2000b). These results showed that our task paradigm successfully elicited neural correlates of face familiarity recognition and, hence, supported the validity of the paradigm employed in this study.

The brain regions associated with face identity deception have been investigated by Bhatt et al. (2009), who employed a “line-up” task in which 3 faces were presented in a line. The 3 faces were either all unfamiliar or 2 unfamiliar and 1 familiar. Bhatt et al. argued that their task paradigm approximated real-world conditions, such as line-up identifications in police stations. Our task paradigm only showed one face in each trial. The purpose was to make the cognitive processing of interest (i.e., cognitive manipulation) more prominent by simplifying the task, allowing us to better capture the neural signals.

The different task designs reflected the balance between approximating reality and getting good neural signals. Interestingly, the neural correlates of deception revealed in the present study were largely consistent with those detected by Bhatt et al. (2009), suggesting that manipulating face familiarity



**Figure 3.** Main effect of the Type of Response (L vs. T). (A) fMRI SPM T-map. Voxels within the significant clusters survived peak-level  $P < 0.05$  FWE correction, cluster extent threshold of  $>25$  voxels. L SFG: left superior frontal gyrus; L Mid-FG: left middle frontal gyrus; L IPG: left inferior parietal gyrus; L IFG: left inferior frontal gyrus; R IFG: right inferior frontal gyrus. (B) ERP 2D topographies for the contrast (L – T) were shown within 2 time windows, that is, [425, 510] and [790, 930] ms, respectively. ERP waveforms (time-locked to face onset) extracted from the channel at left occipital-parietal (41) area were also shown. The shadowed intervals indicate the time windows in which amplitudes for L were significantly different from those for T. L: response of lie; T: response of truth.

recruits similar neural correlates across task designs. The same activated brain regions were found in many other fMRI studies on deception (Lee et al. 2002, 2005; Phan et al. 2005; Christ et al. 2009). Therefore, manipulating information, which is one of the most important parts of deception, may employ similar neural circuits across different types of deception.

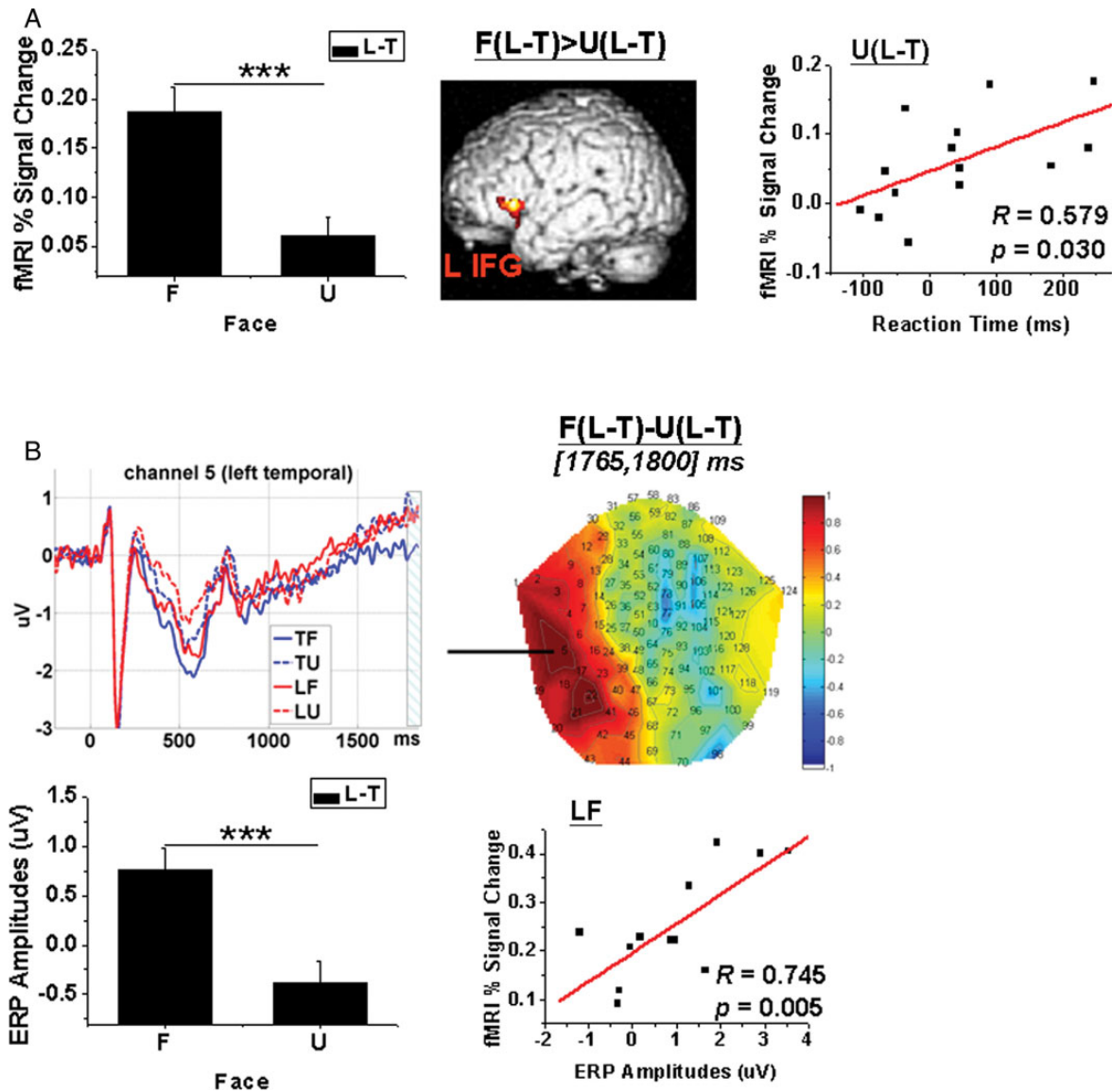
#### Early Retrieval of Truth

An interesting result further improved our understanding of the neural processing of deception. Deceptive responses were found to be associated with less negative-going amplitudes at left occipital-parietal sites within a time interval of [425, 510] ms, which was just in the middle of the time window for recognizing face familiarity, that is, [270, 670] ms. This suggests that the earliest neural processing of deception appears after the early component but before the late component of face familiarity recognition. In other words, the neural processing of deception begins earlier than the accomplishment of face recognition. Gobbini and Haxby (2007) have proposed a neural model for face perception in which the anterior paracingulate,

posterior superior temporal sulcus, and temporal-parietal junction serve in the retrieval of personal traits, intentions, attitudes, and mental states of familiar individuals. The anterior temporal areas play roles in the representation of semantic and biographical information. The precuneus participates in the retrieval of episodic memories. This model does not mention the temporal sequence of the activations of these brain regions. Our source reconstruction analyses showed that, within the time interval (i.e., [270, 425] ms) just before the early processing of deception (i.e., [425, 510] ms), greater source intensities at the left middle temporal gyrus, medial-frontal gyrus, and right superior temporal gyrus were associated with familiar faces, suggesting that the early processing of deception is elicited immediately after the retrieval of personal traits of familiar individuals.

#### Early Processing of Deception

Within the time window of [425, 510] ms, deceptive responses were found to elicit weaker intensities at the left inferior



**Figure 4.** Interaction effect between Face and Type of Response. (A) fMRI results. Within the SPM T-map, voxels in the significant clusters were height threshold  $P < 0.001$  and survived peak- or cluster-level  $P < 0.05$  FWE correction. The percent signal changes extracted from L IFG for the contrast U(L – T) were positively correlated with the reaction time for the same contrast ( $N = 14$ ). L IFG: left inferior frontal gyrus. (B) ERP results. The 2D topography [F(L – T) – U(L – T)] was averaged across [1765, 1800] ms (time-locked to face onset). The waveform was extracted from channel 5 at left temporal scalp areas. The shadowed interval covers the time windows in which amplitudes for F(L – T) were significantly different from those for U(L – T). The amplitudes for condition LF extracted from channel 5 averaged across [1765, 1800] ms were positively correlated with percent signal changes for LF extracted from L IFG ( $N = 12$ ). Notice that the patterns of difference between F and U for the contrast (L – T) were similar between (A, left panel) percent signal changes extracted from the L IFG and (B, left panel) amplitudes extracted from channel 5 at left temporal scalp areas. F: familiar face; U: unfamiliar face; L: response of lie; T: response of truth. \*\*\* $P < 0.001$ . Error bar denotes SEM.

frontal gyrus, known as Broca’s Area, dominating language production (Costafreda et al. 2006). Therefore, the decreased activation at the left inferior gyrus for deceptive responses might be explained by the hypothesis that participants suppressed the language production of the truthful labeling of faces.

#### Late Retrieval of Truth

Within the later 2 intervals (i.e., [425, 510] and [510, 670] ms), greater source intensities were found for familiar faces at the left precuneus. This observation suggests the retrieval of episodic memories of familiar individuals appears later than that of personal traits and is independent of the early processing of deception. The left middle temporal gyrus was also found to

show stronger source intensities for familiar faces within [510, 670] ms. Considering that it was also found to be activated during the early interval for face familiarity recognition, that is, [270, 425] ms, the left middle temporal gyrus might work in both early and late intervals to provide information on face familiarity for further processing by the other brain regions.

#### Late Processing of Deception

Source reconstruction analysis showed stronger source intensities for deceptive responses at the left inferior parietal gyrus within [790, 930] ms. This time interval followed the processing of differentiating familiar and unfamiliar faces, suggesting that deception during this time window involves manipulating

Time (ms)	Scalp area			Cluster	Z
	t1	t2	t-peak		
Main effect of Face					
F > U					
275	645	495	Frontal-central	3963	4.769
F < U					
270	670	490	L temporal	8367	4.938
280	570	510	R temporal	1616	4.079
Main effect of Type of Response					
L > T					
425	510	485	L occipital-parietal	753	3.736
790	930	865	L occipital	1880	4.046
L < T					
None					
Interaction effect between Type of Response and Face					
F(L - T) > U(L - T)					
1765	1800	1775	L temporal	475	4.540
F(L - T) < U(L - T)					
None					

t1 and t2 represent the begin and end time of the significant cluster. t-peak indicates the time point of the maximal significance. Cluster represents the number of voxels within the significant cluster. Z indicates the z value.

F: familiar face; U: unfamiliar face; T: truthful response; L: deceptive response.

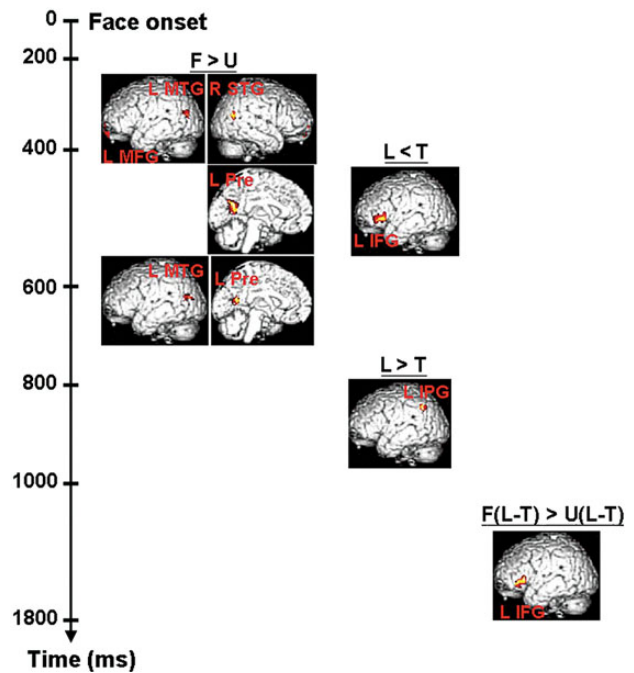
truthful information after it has been completely retrieved. Previous studies have shown that the inferior parietal gyrus plays an important role in working memory updates (Borst and Anderson 2013). Therefore, during this time interval, the left inferior parietal gyrus might help to manipulate truthful information for deceptive aims.

### Preparation Before Response

Familiar and unfamiliar faces are treated differently during deception. This idea was supported by the behavioral data that familiar faces were associated with higher accuracy and shorter reaction time relative. fMRI data also showed that familiar faces were associated with stronger activations at the left inferior frontal gyrus for the contrast L - T. Deception involving unfamiliar faces only needs a response contrary to the truth, whereas deception involving familiar faces further requires inhibiting the brain network activated for retrieval of information about familiar individuals. Previous studies have shown that the right inferior frontal gyrus plays crucial roles in inhibition of truthful responses (Aron et al. 2004; Aron and Poldrack 2005).

The current finding suggested that, to compensate for the function of the right inferior frontal gyrus, the left inferior frontal gyrus is employed to provide extra cognitive resources for inhibition. Thus, more neural processing of inhibition might be recruited for deception involving familiar faces. This point of view was further supported by the positive correlation between brain activations for the contrast U(L - T) extracted from the left inferior frontal gyrus and the reaction time for the same contrast. That is, the more difficulty participants experienced in telling lies about unfamiliar faces, the longer reaction time, and more cognitive resources for inhibition were spent to complete the task. Considering that it is much more difficult to tell lies about familiar faces, the inhibition resources should be exhausted by deception involving familiar faces across participants.

This finding also explains why there is no significant correlation between fMRI activations and behavioral performance for the contrast F(L - T). Both the ERP data and results of



**Figure 5.** Source reconstruction results for ERP data with fMRI spatial priors. Source intensities for contrasts of interest within the time windows and brain regions predefined by fMRI and ERP analyses were calculated. Within the SPM T-maps, voxels in the significant clusters were height threshold  $P < 0.05$  (no correction) with the extent of  $> 25$  voxels. F: familiar face; U: unfamiliar face; L: response of lie; T: response of truth.

source analysis suggest that this interaction effect appeared within the interval of [1765, 1800] ms, which was just before the presentation of the text cue. The fMRI activation extracted from the left inferior frontal gyrus and the ERP extracted from the channel at the scalp left temporal site showed positive correlation. These results suggest that, within a short interval just before response, deception involving familiar and unfamiliar faces recruits different strategies as well as different recruitment styles of the same neural circuits.

### Limitations

Several limitations existed in the present research. The first is that only male volunteers were recruited. The lack of females excludes generality of the results. Secondly, we did not find significant source intensities in some brain regions indicated by fMRI results for the same contrast. This may be explained by the hypothesis that the ERP source reconstruction analyses might ignore some brain activities, because the ERP (but not fMRI) method focuses on time- and phase-locked signals. Thirdly, the task paradigm of directed lying does not investigate the other cognitive/affective processes that occur during deception in reality, such as decision-making (to deceive or to be honest) before action and feedback evaluation after outcome onset (being detected or not). Finally, the fMRI and ERP data from the same participant were not recorded simultaneously. The time lag might have elicited variability in brain activations even though the participants were the same. Future studies on the neural processing of deception should investigate neural responses simultaneously and should recruit samples from different sample groups as well as task paradigms sensitive to investigating the other cognitive/affective functions involved in deception.



Despite the fact that the small sample size may limit the generalization of the current findings, the results were significant and consistent with the findings of previous studies. Only 12 of 17 participants consented to participate in both experiments. To ensure the validity of the conclusions drawn, we compared the data of these 12 participants with that reported here and noticed that the findings of the 2 data sets are largely consistent. Source reconstruction analyses showed significant results at brain regions of interest and within the time windows of interest, suggesting that both fMRI and ERP data were associated with similar neural processing.

### Conclusions

The findings supported a priori model of deception, which suggests the representation of truthful information is first aroused and then manipulated during deception. The data helped delineate the neural activity associated with the following temporal sequence of the neurocognitive processes of deception: Early retrieval of truth, early processing of deception, late retrieval of truth, late processing of deception, and preparation before executing the response. Our results provide a model describing the dynamic interaction between temporal and spatial neural processing in the act of lying about face familiarity.

### Funding

This work was supported by the May Endowed Professorship of The University of Hong Kong and the Research Grant Council General Research Fund (Ref:HKU747612H). The funders have no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Funding to pay the Open Access publication charges for this article was provided by the the KKHo International Charitable Foundation.

### Notes

*Conflict of Interest:* None declared.

### References

Aron AR, Poldrack RA. 2005. The cognitive neuroscience of response inhibition: relevance for genetic research in attention-deficit/hyperactivity disorder. *Biol Psychiatry*. 57:1285–1292.

Aron AR, Robbins TW, Poldrack RA. 2004. Inhibition and the right inferior frontal cortex. *Trends Cogn Sci*. 8:170–177.

Bhatt S, Mbwana J, Adeyemo A, Sawyer A, Hailu A, Vanmeter J. 2009. Lying about facial recognition: an fMRI study. *Brain Cogn*. 69:382–390.

Boly M, Garrido MI, Gosseries O, Bruno MA, Boveroux P, Schnakers C, Massimini M, Litvak V, Laureys S, Friston K. 2011. Preserved feed-forward but impaired top-down processes in the vegetative state. *Science*. 332:858–862.

Borst JP, Anderson JR. 2013. Using model-based functional MRI to locate working memory updates and declarative memory retrievals in the fronto-parietal network. *Proc Natl Acad Sci USA*. 110:1628–1633.

Christ SE, Essen DC, Watson JM, Brubaker LE, McDermott KB. 2009. The contributions of prefrontal cortex and executive control to deception: evidence from activation likelihood estimate meta-analyses. *Cereb Cortex*. 19:1557–1566.

Costafreda SG, Fu CHY, Lee L, Everitt B, Brammer MJ, David AS. 2006. A systematic review and quantitative appraisal of fMRI studies of

verbal fluency: role of the left inferior frontal gyrus. *Hum Brain Mapp*. 27:799–810.

Eimer M. 2000a. Effects of face inversion on the structural encoding and recognition of faces. Evidence from event-related brain potentials. *Brain Res Cogn Brain Res*. 10:145–158.

Eimer M. 2000b. Event-related brain potentials distinguish processing stages involved in face perception and recognition. *Clin Neurophysiol*. 111:694–705.

Farrow TFD, Reilly R, Rahman TA, Herford AE, Woodruff PWR, Spence SA. 2003. Sex and personality traits influence the difference between time taken to tell the truth or lie. *Percept Mot Skills*. 97:451–460.

Friston K, Harrison L, Daunizeau J, Kiebel S, Phillips C, Trujillo-Barreto N, Henson R, Flandin G, Mattout J. 2008. Multiple sparse priors for the M/EEG inverse problem. *Neuroimage*. 39:1104–1120.

Gobbini MI, Haxby JV. 2006. Neural response to the visual familiarity of faces. *Brain Res Bull*. 71:76–82.

Gobbini MI, Haxby JV. 2007. Neural systems for recognition of familiar faces. *Neuropsychologia*. 45:32–41.

Gratton G, Coles MG, Donchin E. 1983. A new method for off-line removal of ocular artifact. *Electroencephalogr Clin Neurophysiol*. 55:468–484.

Henson RN, Flandin G, Friston KJ, Mattout J. 2010. A parametric empirical Bayesian framework for fMRI-constrained MEG/EEG source reconstruction. *Hum Brain Mapp*. 31:1512–1531.

Henson RN, Mouchlianitis E, Matthews WJ, Kouider S. 2008. Electrophysiological correlates of masked face priming. *Neuroimage*. 40:884–895.

Herzmann G, Schweinberger SR, Sommer W, Jentsch I. 2004. What's special about personally familiar faces? A multimodal approach. *Psychophysiology*. 41:688–701.

Hoffmann H, Kessler H, Eppel T, Rukavina S, Traue HC. 2010. Expression intensity, gender and facial emotion recognition: women recognize only subtle facial emotions better than men. *Acta Psychol*. 135:278–283.

Johnson MM, Rosenfeld JP. 1992. Oddball-evoked P300-based method of deception detection in the laboratory. II: Utilization of non-selective activation of relevant knowledge. *Int J Psychophysiol*. 12:289–306.

Johnson R Jr, Barnhardt J, Zhu J. 2004. The contribution of executive processes to deceptive responding. *Neuropsychologia*. 42:878–901.

Johnson R Jr, Barnhardt J, Zhu J. 2003. The deceptive response: effects of response conflict and strategic monitoring on the late positive component and episodic memory-related brain activity. *Biol Psychol*. 64:217–253.

Johnson R Jr, Henkell H, Simon E, Zhu J. 2008. The self in conflict: the role of executive processes during truthful and deceptive responses about attitudes. *Neuroimage*. 39:469–482.

Joyce C, Rossion B. 2005. The face-sensitive N170 and VPP components manifest the same brain processes: the effect of reference electrode site. *Clin Neurophysiol*. 116:2613–2631.

Katznelson RD. 1981. EEG recording, electrode placement, and aspects of generator localization. In: Nunez PL, editor. *Electric fields of the brain: the neurophysics of EEG*. New York: Oxford University Press. pp. 76–213.

Kiebel SJ, Friston KJ. 2004a. Statistical parametric mapping for event-related potentials: I. Generic considerations. *Neuroimage*. 22:492–502.

Kiebel SJ, Friston KJ. 2004b. Statistical parametric mapping for event-related potentials (II): a hierarchical temporal model. *Neuroimage*. 22:503–520.

Langen DD, Schroeder L, Maldjian JA, Gur RC, McDonald S, Ragland JD, O'Brien CP, Childress AR. 2002. Brain activity during simulated deception: an event-related functional magnetic resonance study. *Neuroimage*. 15:727–732.

Lee TM, Liu HL, Chan CC, Ng YB, Fox PT, Gao JH. 2005. Neural correlates of feigned memory impairment. *Neuroimage*. 28:305–313.

Lee TM, Liu HL, Tan LH, Chan CC, Mahankali S, Feng CM, Hou J, Fox PT, Gao JH. 2002. Lie detection by functional magnetic resonance imaging. *Hum Brain Mapp*. 15:157–164.

Litvak V, Friston K. 2008. Electromagnetic source reconstruction for group studies. *Neuroimage*. 42:1490–1498.

Oldfield RC. 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*. 9:97–113.

- Phan KL, Magalhaes A, Ziemlewicz TJ, Fitzgerald DA, Green C, Smith W. 2005. Neural correlates of telling lies: a functional magnetic resonance imaging study at 4 Tesla. *Acad Radiol.* 12:164–172.
- Pierce LJ, Scott LS, Boddington S, Droucker D, Curran T, Tanaka JW. 2011. The n250 brain potential to personally familiar and newly learned faces and objects. *Front Hum Neurosci.* 5:111.
- Proverbio AM, Brignone V, Matarazzo S, Del Zotto M, Zani A. 2006. Gender differences in hemispheric asymmetry for face processing. *BMC Neurosci.* 7:44.
- Schweinberger SR, Pickering EC, Jentsch I, Burton AM, Kaufmann JM. 2002. Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions. *Brain Res Cogn Brain Res.* 14:398–409.
- Spence SA, Farrow TFD, Herford AE, Wilkinson ID, Zheng Y, Woodruff PWR. 2001. Behavioural and functional anatomical correlates of deception in humans. *Neuroreport.* 12:2849–2853.
- Spence SA, Hunter MD, Farrow TF, Green RD, Leung DH, Hughes CJ, Ganesan V. 2004. A cognitive neurobiological account of deception: evidence from functional neuroimaging. *Philos Trans R Soc Lond B Biol Sci.* 359:1755–1762.
- Talairach J, Tournoux P, Musolino A. 1988. Anatomical stereotaxic studies of the frontal-lobe in the management of the epilepsies. *Epilepsia.* 29:205.