

(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

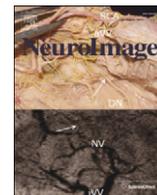
In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: [www.elsevier.com/locate/ynimg](http://www.elsevier.com/locate/ynimg)

## Neuroanatomical substrates involved in true and false memories for face

Tetsuya Iidaka<sup>a,\*</sup>, Tokiko Harada<sup>a</sup>, Jun Kawaguchi<sup>b</sup>, Norihiro Sadato<sup>c</sup>

<sup>a</sup> Nagoya University, Graduate School of Medicine, Nagoya, Japan

<sup>b</sup> Nagoya University, Graduate School of Environmental Studies, Nagoya, Japan

<sup>c</sup> National Institute for Physiological Sciences, Okazaki, Japan

### ARTICLE INFO

#### Article history:

Accepted 21 April 2012

Available online 28 April 2012

#### Keywords:

Amygdala  
Anterior cingulate cortex  
fMRI  
Illusory memory  
Orbito-frontal cortex

### ABSTRACT

We often mistake an unknown person for a familiar person because of the similarities in facial features. This phenomenon, known as false memory, has been investigated mainly using words, pictures, and shapes. Previous neuroimaging studies on false memory have shown that both true and false memories trigger a similar activation in the medial temporal lobe, suggesting that it plays a common role in both. However, no study to date has investigated neural substrates of false memories for faces. In the present fMRI study, we applied a modified version of the standard false memory paradigm, using morphed pictures of faces, to induce false memory in an MRI environment. We found that activity in the amygdala and orbital cortices was associated with the degree of familiarity of items. In particular, false responses to “lure” items evoked a level of activity in the amygdala between that evoked for correct or incorrect responses to “true” items. This indicates a possible role of the amygdala in false memory. A specific region in the anterior cingulate cortex was involved in false recognition; the activity being correlated to reaction times for the response types. These results suggest that the amygdala is involved in determining the relevance of items; therefore, ambiguousness of lure items in terms of familiarity and novelty may be related to decreased activity in the amygdala. The anterior cingulate activity in false memory may be caused not only by increased effort and motor demand but also by higher mnemonic processing of lure items.

© 2012 Elsevier Inc. All rights reserved.

### Introduction

Human memory is often inaccurate, particularly when a person's available cognitive resources are reduced by other distracting factors, or when items lack salient features that would aid effective encoding. In these situations, accurate memory formation can be disturbed and is prone to various types of distortion (Schacter and Slotnick, 2004). Numerous laboratory-based studies have investigated memory distortion by inducing false memory for items presented either visually or audibly (Dodson et al., 2000; Roediger, 1996; Schacter and Dodson, 2001). In these studies, false recognition was defined as a condition in which subjects incorrectly identified novel items during a recognition phase as being familiar from an earlier encoding phase.

The occurrence of false memory is increased in a task developed by Roediger and McDermott (1995), which is based on earlier work by Deese (1959) (Deese–Roediger/McDermott paradigm, or DRM paradigm). In the original DRM paradigm, subjects were presented with lists of associated words (e.g., *candy*, *sugar*, and *bitter*) that

were all related by a critical but unrepresented lure word (e.g., *sweet*). In a subsequent recognition test, the proportion of old responses to the lure item was significantly higher than that for unrelated and unstudied items. A modified version of the DRM paradigm has also been used in studies investigating false memory for pictorial materials (Koutstaal et al., 1999). However, despite the commonly reported experience of false memory for faces, the DRM paradigm has rarely been used in this context.

Similar kinds of memory distortion for face have been investigated by using items constructed by combining the facial features of separately studied items and are therefore composed of previously memorized components (Jones and Bartlett, 2009; Jones et al., 2006; Reinitz et al., 1992). False alarm rates for these combined-face stimuli were typically higher than those for completely new faces and lower than the hit rates of studied faces. In a study using natural face photographs, a conjunction error rate (i.e., false alarms to items composed of studied items) was almost equal to the feature error rate (i.e., false alarms to items composed of half-studied and half-new components) (Jones et al., 2006). In addition, false alarm rates for conjunctive face stimuli did not increase in senior subjects as compared to those in young subjects (Searcy et al., 1999). These results suggest that memory errors for the combined faces are mainly based on familiarity processes in the absence of the precise recollection process of episodic memory (Jones and Bartlett, 2009).

\* Corresponding author at: Department of Psychiatry, Nagoya University, Graduate School of Medicine, 65 Tsurumai, Showa, Nagoya, Aichi, 466-8550, Japan. Fax: +81 52 744 2293.

E-mail address: [iidaka@med.nagoya-u.ac.jp](mailto:iidaka@med.nagoya-u.ac.jp) (T. Iidaka).

Functional neuroimaging has been employed to elucidate neuro-anatomical substrates of false memory for words (Abe et al., 2008; Cabeza et al., 2001; Garoff-Eaton et al., 2007; Kim and Cabeza, 2007; Schacter et al., 1996, 1997), pictures (Gonsalves et al., 2004; Okado and Stark, 2003), and shapes (Garoff-Eaton et al., 2006; Slotnick and Schacter, 2004). Major findings from these studies are as follows: i) levels of neural activity in the medial temporal lobe were almost equal for both true and false memories and ii) primary sensory areas activated during the encoding phase are reactivated during the retrieval phase for “true” items. We hypothesized that perceptual similarity between faces gives rise to false memory for face as observed in the previous behavioral studies (Jones and Bartlett, 2009; Jones et al., 2006; Reinitz et al., 1992; Searcy et al., 1999), while false memory for words is associated with conceptual similarity between words. Therefore, in addition to visual sensory areas, regions involved in detection of facial salience such as the amygdala would be active during false memory paradigm for face.

The face conveys various information about a person such as identity, familiarity, and emotion and is also allows the evaluation of social dimensions such as trustworthiness and attractiveness (Todorov, 2008). Neural responses in the amygdala to facial images have been shown to be linearly or nonlinearly correlated with facial features, particularly those within the social dimension (Said et al., 2010). Several neuroimaging studies have shown that both facial attractiveness (Winston et al., 2007) and trustworthiness (Said et al., 2009; Todorov, 2008) are related to nonlinear responses in the amygdala with greater activity in response to stimuli at the extremes of each dimension as compared to those at the middle of the continuum. Based on these previous reports, we hypothesized that amygdala activation would correlate with the familiarity of face (Jones and Bartlett, 2009). In addition, we predicted that the activation would show a nonlinear pattern with activities in response to stimuli at the 2 extremes, i.e., old responses to studied items and new responses to unstudied items, being greater than those to stimuli from the middle of the continuum, i.e., responses to lure items. This is because a face that is either familiar or unfamiliar is a more salient stimulus than one that is ambiguous regarding familiarity in ordinary situations.

The present functional magnetic resonance imaging (fMRI) study uses a modified version of the DRM paradigm with face pictures to investigate, for the first time to our knowledge, the neural correlates of false memory for human faces. We created face pictures with similar appearances by using computerized morphing techniques to provoke false memories in the experimental setting. Neural activity was measured using event-related fMRI and correlated with the degree of familiarity. Specifically, we hypothesized that the difference in amygdala activity between old- and new-response conditions would be high for old items, middle for lure items, and low for new items. Furthermore, to disentangle the results of correlation analysis, activity from fixation baseline was measured in each item and response condition. The mean activity was plotted as a function of saliency; i.e., old responses to old item and new responses to new item are the 2 extremes of saliency. We predicted that the amygdala, which is known to be involved in face processing (Haxby et al., 2000) and detection of stimulus saliency (Phan et al., 2002), would show the highest activity under high-saliency conditions and lower activity under low-saliency conditions involving the lure items.

In addition, we sought to elucidate the functional role of the anterior cingulate cortex (ACC) in false memory, which has been reported but rarely focused on in previous neuroimaging studies (Okado and Stark, 2003; von Zerssen et al., 2001). In these studies, the role of the ACC is limited to the neural processing of effortful responses to the lure items; however, there is no precise analysis of the relationship between activation and reaction time. In the present study, we conducted a correlation analysis between ACC activity and reaction time differences under false and true memory conditions to elucidate the possible role of ACC in mnemonic processing.

## Materials and methods

### Subjects

Twenty healthy volunteers with normal or corrected-to-normal vision were recruited for this fMRI study. The subjects were undergraduate or graduate students recruited from Nagoya University community. Subjects provided written informed consent for the procedure. An investigator carefully checked their physical and mental condition according to a standard institutional protocol. This study was approved by the ethics committee of the National Institute for Physiological Sciences. Data from 1 subject was excluded from the analysis because the rate of old responses for lure items did not exceed that for new items (0.39 and 0.39, respectively), precluding an interpretation of the subject's fMRI data. The results obtained using the remaining 19 subjects (8 men and 11 women; right-handed; mean age,  $21.2 \pm 1.1$  years) are presented in this study.

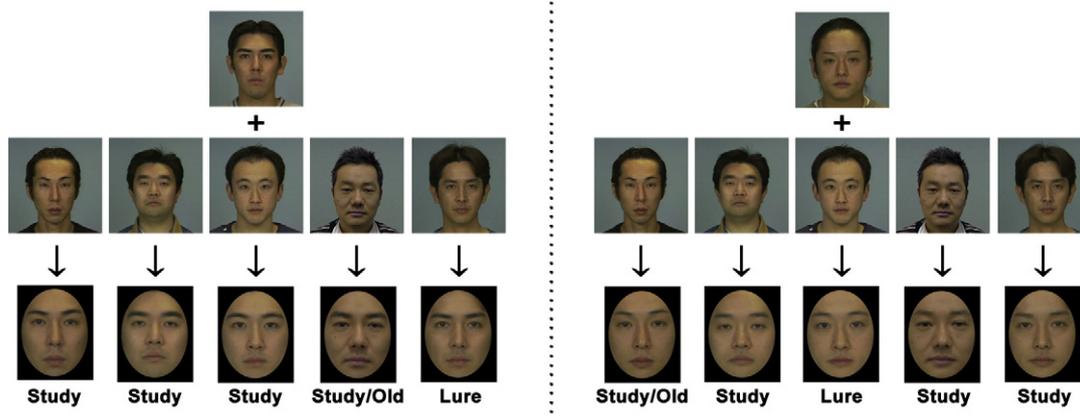
### Experimental stimuli

We prepared 384 colored pictures of faces by morphing and averaging original face pictures. The original face pictures were selected from a Japanese database of face pictures provided by Softopia Japan Foundation (<http://www.softopia.or.jp/rd/facedb.html>). The database consisted of 150 photographs each of male and female faces (age, 15–64 years); the photographs showed the front faces with neutral expressions. The photographs were divided into 5-year groups according to gender and age (15- to 19-year-old male faces, 20- to 24-year-old male faces, and so on); 20 groups of 15 face pictures each were formed. Each of the 20 groups was further divided into 2 subgroups of 7 or 8 face pictures, thus forming 40 subgroups in total.

We randomly chose 2 face pictures in each subgroup for combining with the remaining 5 or 6 face pictures by using a face-morphing software (Sqirlz Morph, <http://www.xiberpix.net/SqirlzMorph.html>). These face pictures were cut in the shape of an ellipse and showed the brows, eyes, nose, and mouth, with a black background. Two sets of 5 morphed face pictures with similar facial appearances were created from 1 subgroup (Fig. 1). One picture from a set of 6 morphed face pictures was discarded to equate the number of items between the sets. One of the 5 morphed faces was used as a “LURE” item and shown only during the test session. The remaining 4 faces were used as “STUDY” items and presented during the study session. One of the “STUDY” items was used as an “OLD” item and presented both in the study and test sessions. Thus, from the 32 subgroups of face pictures, 256 STUDY items (64 of which were used as OLD items) and 64 LURE items were created. From the remaining 8 subgroups of face pictures, 64 NEW items were created using the same morphing procedures. These NEW items were presented only during the test sessions and never during the study session. These face pictures (256 STUDY, 64 LURE, and 64 NEW items) were equally divided into 2 stimulus sets for 2 sessions of the fMRI experiment.

### Pilot study for the assessment of face similarity

Before the fMRI experiment, we conducted a pilot study to assess the subjective similarity of the face pictures between the item types. This pilot study included 80 native Japanese university students (48 men and 32 women; mean age [s.d.] = 20.9 [2.2] years) who did not participate in the fMRI experiment. The study was conducted in an ordinary classroom using a pencil-and-paper test. The test consisted of 24 questions, each consisting of 1 target face picture and 4 or 5 probe face pictures that were compared with the target. There were 8 different patterns displayed on the test sheets, each showing 136 morphed faces. All the 384 morphed faces created in the previous



**Fig. 1.** A schema showing the process of creation of morphed face pictures. A stimulus set consisted of 7 face pictures of persons of the same gender and age group. Two individual faces (top row) were randomly chosen for combining with each of the remaining faces (middle row) to create 5 morphed face pictures with similar facial appearances (bottom row). In the fMRI experiment, 1 of the 5 morphed faces was randomly selected as a “LURE” item and shown only during the test session. The other faces were used as “STUDY” items and shown during the study session. One of the “STUDY” items was used as an “OLD” item and shown during both the study and test sessions. Another set of faces, created in the same manner but never shown during the study session, was used as “NEW” items during the test session. Face pictures shown in this figure are not those used in the present study due to copyright restrictions on the use of the photographs from the database.

section were randomly sorted in these 8 patterns. Eighty subjects were assigned a single pattern, with 10 subjects for each pattern. The target face pictures were randomly chosen from the “OLD,” “LURE,” or “NEW” items, with 8 faces for each item type. The probe face pictures were chosen from the “STUDY” items. The subject rated the similarity of the target and the probe faces by using the 7-point Likert scale (from 1, for completely different persons, to 7, for identical persons).

The mean similarity score for comparison of the probe and target faces within the same gender and age range was computed. The first comparison was made between the LURE and STUDY items created using the same original face. We anticipated that those faces would be rated as highly similar to each other. The second comparison was made between the NEW and STUDY items created using different sets of faces with the same gender and age range; we anticipated that these faces would be rated as less similar. The results showed that the mean (s.d.) similarity score for the 2 comparisons was 4.98 (1.05) and 3.71 (1.15), respectively (see supplementary Fig. 1). A paired *t*-test showed a significant ( $p < 0.01$ ) difference in the mean similarity score between the 2 comparisons. The results indicate that the morphing and averaging procedures used in the present study could create sets of faces that were highly similar to each other.

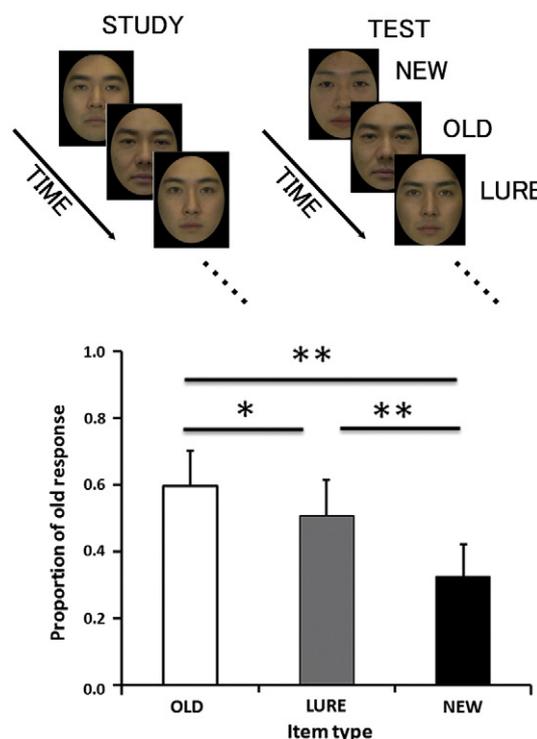
*Experimental procedure*

*Study sessions*

A study session, not scanned by fMRI, was conducted inside the scanner unit and consisted of 16 task blocks separated by 5-s rest periods. At the beginning of each block, a red fixation point was presented for 1 s as a start cue. Within each block, 8 morphed face pictures were successively presented 1 at a time for 2 s with an inter-stimulus interval of 1 s (see supplementary Fig. 2). The 8 faces were 6 STUDY items and 2 OLD items from a subgroup with the same gender and age range as described in the previous section. The blocks were presented in a fixed order, but the order of face pictures was randomized within the block. During the study session, the subjects were presented with 128 face pictures (16 blocks by 8 faces), which they were requested to memorize (Fig. 2). The results of a pilot behavioral study have indicated that a single study session with morphed faces was not sufficient to provoke false memory. Therefore, in the fMRI experiment, the study session was immediately repeated with the presentation order randomized only within the block.

*Test sessions*

During the test sessions, the subjects were presented with 96 morphed face pictures, including 32 OLD items, 32 LURE items, and 32 NEW items (Fig. 2). After the presentation of a fixation point with 1-s duration, the face pictures were presented for 2 s in a



**Fig. 2.** Top: During the study session, 128 morphed faces were presented 1 at a time, in  $8 \times 16$  item blocks, and subjects were instructed to memorize these faces for the later test session. During the test session, 32 OLD, 32 LURE, and 32 NEW faces were presented in a random order, and subjects provided a recognition judgment for each item. The study/test session was repeated twice with different sets of morphed faces used as memory items. See Materials and methods for details. Bottom: The mean (column) and s.d. (bar) of the proportion of old responses to OLD, LURE, and NEW items during the recognition test are shown. One-way ANOVA showed a significant difference in the proportion of old response between the item types ( $p < 0.001$ ). The post-hoc *t*-test showed a significant difference between the OLD and LURE ( $p < 0.05$ ) and between the LURE and NEW ( $p < 0.01$ ) items. The greater old response for the LURE than for NEW items indicates the occurrence of false memory in the present experiment. Asterisk indicates significance (\* $p < 0.05$ , \*\* $p < 0.01$ ).

randomized order for each subject. A fixation point interleaved the trials with jittering, ranging from 2 s to 10 s (mean, 6 s). The subjects were asked to indicate whether the presented face picture had been shown in the study session by pressing a corresponding button as quickly as possible. The subjects made a 4-choice judgment with regard to familiarity and confidence: surely old, maybe old, maybe new, and surely new. The participants responded by pressing a button of 2, 2-button boxes, 1 held in each hand. Each of the 4 buttons was assigned to surely old, maybe old, maybe new, and surely new response, respectively. The subjects pressed the button by their index and middle fingers of the both hands. The assignment of the response buttons was counterbalanced across the subjects. Because the mean number of “surely new” and “surely old” responses was small (8.3 and 9.6 events per subject, respectively), responses were integrated into “old” and “new” for image data analysis. The results for the behavioral data are shown in Table 1 and the mean proportion of old response is plotted in Fig. 2. The fMRI experiment consisted of 2 sets of study/test sessions with different face pictures used for each set. The total number of faces for which the recognition judgment was made throughout the 2 test sessions was 64 for each item type.

*fMRI acquisition and preprocessing*

Functional images of the brain were obtained during the test sessions in an axial-oblique position by using a 3-T MRI scanner (Allegra, Siemens, Erlangen, Germany) equipped with single-shot echo planar imaging (EPI; TR = 2.3 s, TE = 30 ms, flip angle = 80°, 64 × 64 matrix, 36 slices, voxel size = 3 mm × 3 mm × 3 mm) that was sensitive to blood oxygen level-dependent contrast. A total of 384 images were obtained in each test session. After discarding the first 4 images, the successive images were subjected to analysis. A high-resolution anatomical T1-weighted image was also acquired (MPRAGE; TR = 2.5 s, TE = 4.38 ms, flip angle = 8°, 256 × 256 matrix, 192 slices, voxel size = 0.75 mm × 0.75 mm × 1 mm) for each subject. The data were analyzed using SPM8 (Wellcome Department of Imaging Neuroscience, London, UK). First, the signal in each slice was temporally realigned to that obtained in the middle slice by using sinc interpolation, and all the volumes were spatially realigned to the mean volume. The resliced volumes were then normalized to the standard Montreal Neurological Institute (MNI) space by using a transformation matrix obtained from the normalization process of the mean EPI image of each subject to the EPI template image. The normalized images were spatially smoothed using an 8-mm Gaussian kernel.

*First-level analysis*

After preprocessing, statistical analysis of the data obtained for each subject was conducted by using a general linear model. At the first level (a fixed-effects model), each event was modeled as a hemodynamic response function. High-pass-frequency filters (128 s) were

**Table 1**  
Mean number of events and RT (s.d. in the parentheses) for each experimental condition.

		Item type		
		OLD	LURE	NEW
Response	Old	Event 37.7 (6.7)	32.0 (6.9)	20.6 (6.0)
		RT 1713 (476)	1776 (513)	1779 (455)
New	Event	25.5 (6.7)	31.1 (6.6)	42.3 (5.6)
	RT	1861 (571)	1846 (565)	1778 (530)

Event; mean number of events, RT; reaction time in ms. Total number of events for each item type was 64.

applied to the time-series data. An autoregressive AR(1) model was used to estimate the temporal autocorrelation. The images were scaled to a grand mean of 100 overall voxels and scans within a session. In the subsequent analysis, the following conditions were separately modeled as regressors: 1, old response to OLD item; 2, new response to OLD item; 3, old response to LURE item; 4, new response to LURE item; 5, old response to NEW item; and 6, new response to NEW item. In addition, 6 movement parameters obtained during the realignment were entered as regressors. The parameter estimates for each condition and for the differences between the conditions were computed from the least mean square fit of the model to the time-series data. Images of the parameter estimates representing event-related activity at each voxel were created for each condition and subject. Particularly, in the main analysis, we used the subtraction between the old and new responses for each of the 3 item conditions because the activity indicated by the contrast images reflects the neural correlates of successful retrieval (Schacter and Slotnick, 2004; Spaniol et al., 2009). In addition, to measure neural activity from the baseline fixation condition in each of the 6 experimental conditions (i.e., old response to OLD item, old response to LURE item, old response to NEW item, new response to OLD item, new response to LURE item, and new response to NEW item), separate one-sample *t*-tests were conducted for these conditions.

*Second-level analysis*

At the second level, the contrast images of each subject were entered into a group analysis (a random-effects model). The contrast images pertaining to the differences in activation between the old and new responses were used for the group analysis. First, the brain regions where activation was significantly greater for the old responses than for the new responses were investigated separately for the OLD, LURE, and NEW items by using a one-sample *t*-test. The results are listed in Table 2 at the threshold of  $p = 0.001$ , uncorrected, and  $k = 10$  voxels. Second, we hypothesized that the sense of familiarity, as indexed by the proportion of old response in each item condition, would be associated with the neural correlates of familiarity as measured by old minus new subtraction of the fMRI data. To test this, we performed ANCOVA analysis of SPM8 by using the 3 item conditions (OLD, LURE, and NEW) as factors and the degree of familiarity as covariates (contrast value: 1 for OLD, 0 for LURE, and -1 for NEW). A significant positive correlation between activity and these values was computed. The statistical threshold was set at  $p = 0.001$ , uncorrected for multiple comparisons and  $k = 10$  voxels (Table 3, Fig. 3 top).

The signal was extracted from the significant cluster (10-mm spherical VOI at the peak voxel listed in Table 3) as identified by the

**Table 2**  
The brain regions where activity was significantly greater for “old” responses than for “new” responses for each item type.

Region name (BA)	Hem	T value	p value	x, y, z	Voxels
<b>OLD</b>					
Inferior frontal gyrus (44)	Lt	4.81	0.001	-44, 16, 28	135
Orbital gyrus (11)	Rt	4.36	0.001	28, 32, -22	15
Fusiform gyrus (36)	Rt	4.29	0.001	38, -46, -26	18
Periamygdaloid cortex	Lt	4.23	0.001	-22, 10, -28	12
<b>LURE</b>					
Superior parietal lobule (7)	Lt	4.72	0.001	-40, -62, 54	35
Inferior frontal gyrus (10)	Lt	4.28	0.001	-44, 48, 4	15
<b>NEW</b>					
Inferior parietal lobule (40)	Lt	5.13	0.001	-34, -38, 46	19
Inferior frontal gyrus (45)	Lt	5.00	0.001	-38, 24, 22	23
Cuneus (18)	Lt	3.99	0.001	-8, -100, 14	14
Superior parietal lobule (7)	Lt	3.95	0.001	-32, -72, 50	16

Hem, hemisphere; Rt, right; Lt, left; BA, Brodmann area, p value is uncorrected for multiple comparisons,  $k = 10$  voxels.

**Table 3**

The brain regions where activity significantly related with the proportion of old responses.

Region name (BA)	Hem	T value	p value	x, y, z	Voxels
Orbital gyrus (11)	Rt	4.26	0.001	26, 28, -20	29
Orbital gyrus (11)	Lt	4.25	0.001	-30, 30, -22	41
Amygdala*	Lt	3.69	0.001	-20, 4, -22	17

Hem, hemisphere; Rt, right; Lt, left; BA, Brodmann area, p value is uncorrected for multiple comparisons,  $k = 10$  voxels.

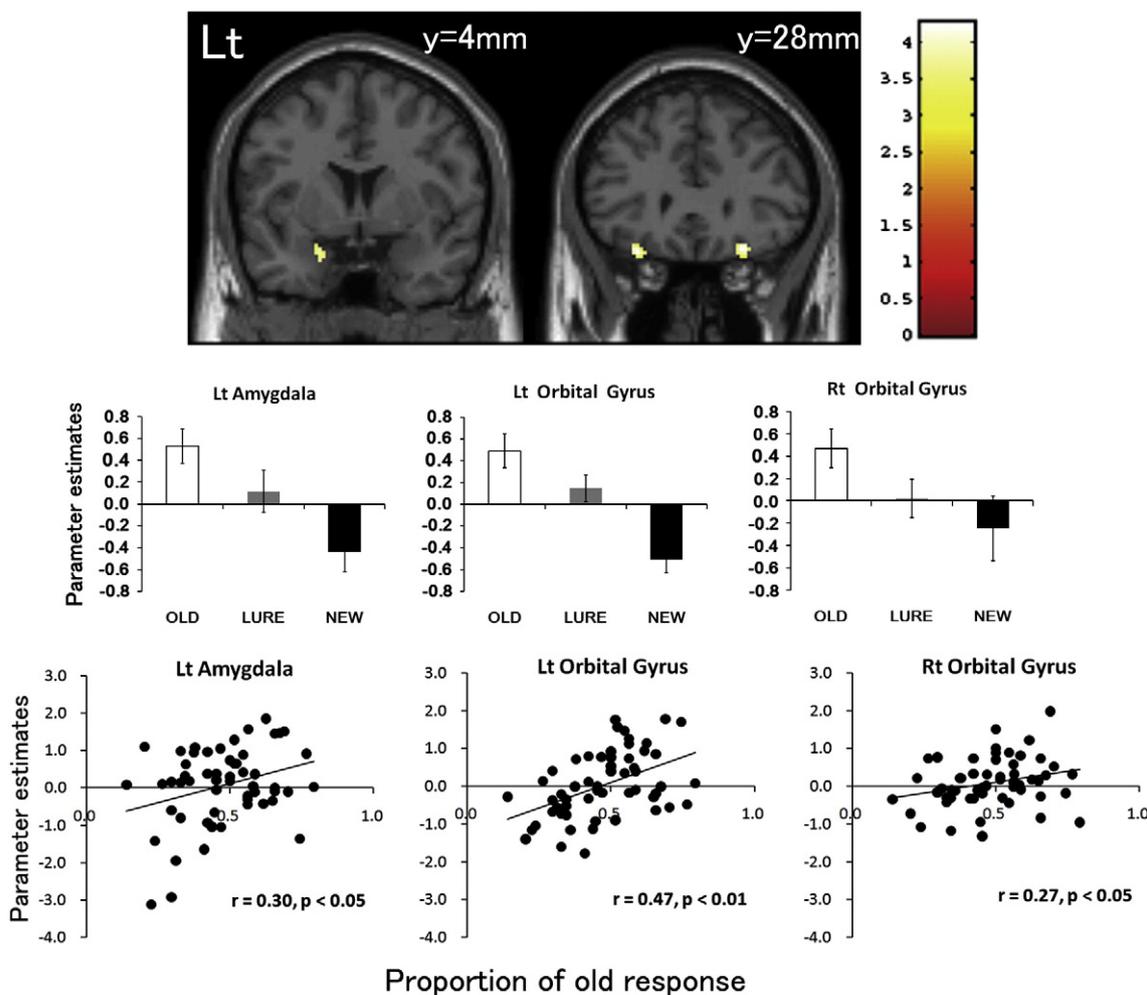
An asterisk indicates significance of small volume correction at  $p = 0.05$  (FEW correction) within the sphere of 8 mm radius at the peak voxel.

ANCOVA model that investigated the relationship between neural activity and degree of familiarity. A principal eigenvariate of parameter estimates in a given VOI was computed using a function implemented in SPM8. The mean  $\pm$  s.e. of the value associated with old minus new subtraction across subjects and for each item type were plotted (Fig. 3, middle). The same value in each subject and condition was plotted against the proportion of old responses (Fig. 3, bottom), and a Pearson's correlation coefficient was computed (threshold at  $p < 0.05$ ). Furthermore, to explore the relationship between amygdala activity and experimental condition, activation from the baseline

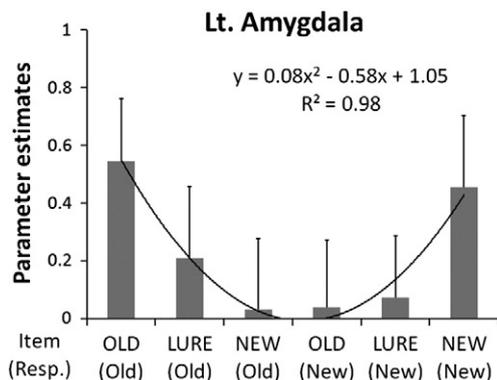
fixation condition was extracted from the separate one-sample  $t$ -tests for each of the 6 experimental conditions (10-mm spherical VOI at  $x, y, z = -20, 4, -22$ , respectively) by using a VOI function in SPM8. The mean value was plotted as a function of the experimental condition and curve fitted by using a quadric polynomial function (Fig. 4).

In addition, to investigate the brain region that was specifically involved in processing false memory for face, the LURE condition was contrasted with each of the OLD or NEW conditions by using an ANOVA model. The contrast images pertaining to the subtraction between old and new responses for OLD, LURE, and NEW items were used for the analysis. The statistical threshold was set at  $p = 0.001$ , uncorrected for multiple comparisons and  $k = 10$  voxels. The results are listed in Table 4, and the significant cluster in the ACC is shown in Fig. 5. The signal extracted from the ACC cluster (10-mm spherical VOI) by using a VOI function in SPM8 was correlated with the difference in reaction time (RT) between the old and new responses in the LURE condition (Fig. 6). A Pearson's correlation coefficient between ACC activity and RT was computed (threshold at  $p < 0.05$ ).

In previous neuroimaging studies of false memory, the medial temporal lobe activity did not significantly differ between the false and true memory conditions (Cabeza et al., 2001; Schacter et al., 1996, 1997; Slotnick and Schacter, 2004). In addition, significantly



**Fig. 3.** Top: The results of fMRI data analysis by using an ANCOVA model show that activity in the left amygdala and bilateral orbital gyri is associated with the degree of conscious memory across item types, i.e., high for OLD, middle for LURE, and low for NEW items. Brain activity in these regions was measured by subtracting new response from old response for each item type. The detailed information from each cluster is shown in Table 3. Middle: The mean  $\pm$  s.e. activation extracted from the left amygdala, left orbital gyrus, and right orbital gyrus are plotted for each item type (OLD, LURE, and NEW). Bottom: A significant correlation between the proportion of old responses (x-axis) and neural activity (y-axis) in the left amygdala and the left and right orbital gyri is plotted. The neural activity on the y-axis indicates the difference in activity between old and new responses. Pearson's correlation coefficients and statistical significance are also shown.



**Fig. 4.** The mean (column) and one s. e. (bar) of the activation levels in the left amygdala (y-axis) are plotted as a function of item and response types (x-axis). The amygdala activity represents the increase in activity from the baseline fixation condition, which was extracted from the separate one-sample *t*-test for each item and response type. The mean value of each condition was fitted using a quadric polynomial function. The expression and *r*-square value are shown.

greater responses in the primary visual areas were observed for old items than for new items, indicating reactivation of the neural signature of previously observed items (Okado and Stark, 2003; Slotnick and Schacter, 2004). To investigate whether these findings have been replicated in the present study, activation from the baseline condition was compared between the conditions using an ANCOVA model (contrast value: OLD, 1; LURE, 1; NEW, -2). The contrast images involving old response to OLD items, old response to LURE items, and new response to NEW items were used for the analysis. The statistical threshold was set at  $p=0.001$ , uncorrected for multiple comparisons and  $k=10$  voxels. The signal changes from the baseline fixation condition were extracted from the 10-mm spherical ROI at the right parahippocampal gyrus and the superior occipital gyrus, and plotted for each condition (Fig. 7). Furthermore, the differences in activity between the OLD and NEW conditions were computed when the response type (old and new) was matched between the conditions.

**Results**

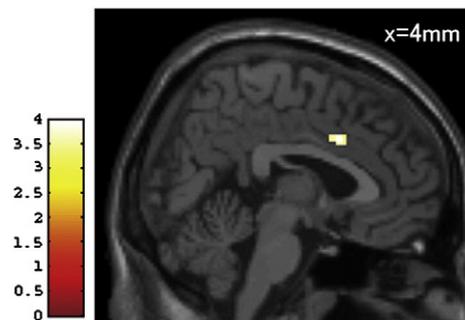
*Behavioral data*

The mean proportion of old responses (s.d. in parentheses) for each of the OLD, LURE, and NEW item sets was 0.60 (0.11), 0.51 (0.11), and 0.33 (0.09), respectively. The results of a one-way ANOVA showed that the proportion of old responses significantly differed between each type of memory item ( $F(2, 54) = 34.2, p < 0.001$ ). Post-hoc *t*-tests with Bonferroni correction for multiple comparisons revealed that the proportion of old responses was significantly ( $p < 0.05$ ) higher for OLD

**Table 4**  
The brain regions specifically related to false memory condition.

Region name (BA)	Hem	T value	p value	x, y, z	Voxels
LURE minus OLD					
Anterior cingulate cortex (24)	Rt	3.97	0.001	4, 14, 34	64
LURE minus NEW					
Fusiform gyrus (37)	Lt	4.08	0.001	-52, -54, -22	14
Middle frontal gyrus (10)	Rt	3.53	0.001	26, 52, -6	11

Hem, hemisphere; Rt, right; Lt, left; BA, Brodmann area. The signal difference between old and new responses was compared between the LURE and other item conditions, *p* value is uncorrected for multiple comparisons,  $k=10$  voxels.

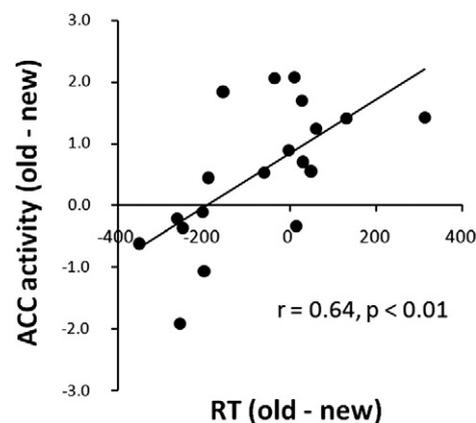


**Fig. 5.** A single cluster in the anterior cingulate cortex (ACC) was predominantly involved in the processing of lure items when the LURE condition was contrasted with the OLD condition. Detailed information of the cluster is shown in Table 4.

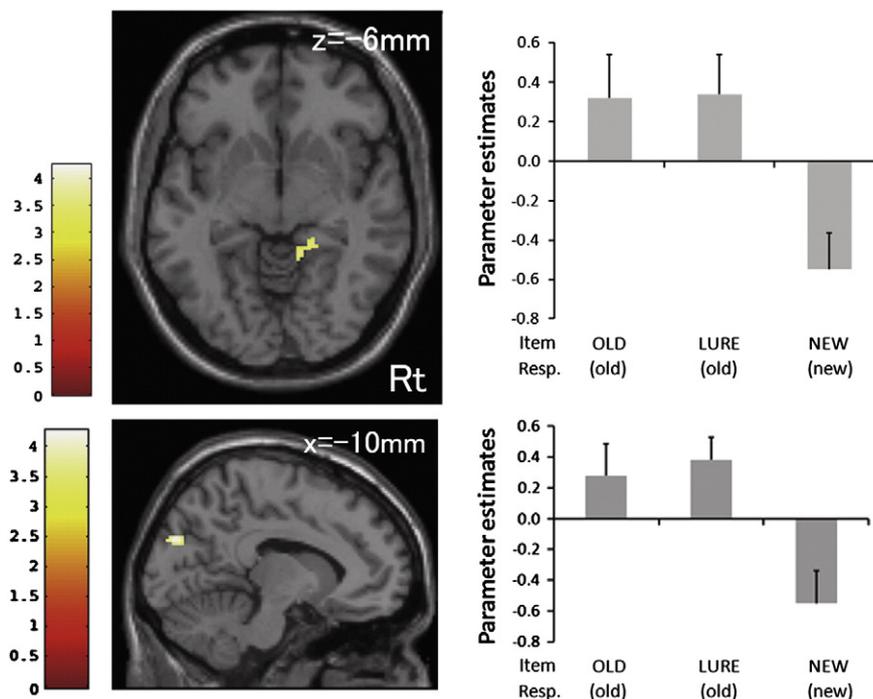
than for LURE, and more critically, for LURE than for NEW conditions ( $p < 0.01$ , Fig. 2). These results indicate that false memory was successfully induced for the LURE items in the present study. There was sufficient number of events to detect neural activity for old responses to the LURE items (mean 32 events, Table 1). The mean number of “sure” and “maybe” responses did not significantly differ between the item conditions (sure,  $F(2, 54) = 0.51, p = 0.59$ ; maybe,  $F(2, 54) = 0.50, p = 0.61$ ; see supplementary table 1). For the reaction time (RT, Table 1), the effect of item type was not significant ( $F(1, 70) = 1.75, p = 0.19$ ) and the effect of response time was trending ( $F(1, 18) = 4.2, p = 0.054$ ) toward faster RTs for old responses than for new responses. There was no significant item-by-response interaction effect on the RT ( $F(1, 35) = 3.1, p = 0.08$ ). The mean and s.d. of the RT for each item type and response type (surely old, maybe old, surely new, and maybe new) are shown in supplementary table 2 and supplementary Fig. 3.

*fMRI data*

Significantly greater activity was observed mainly in the prefrontal and parietal cortices for old responses than for new responses for each item type (Table 2). The results of ANCOVA using the 3 item types as factors and the degree of familiarity as covariates reveal a significant positive relationship between familiarity and neural activity in the left amygdala and the bilateral orbital cortices (Fig. 3, top, middle, Table 3). The significant activation of the left amygdala persisted following a small volume correction at  $p=0.05$  (FWE correction) within a spherical ROI with 8 mm radius at the peak voxel.



**Fig. 6.** A significant correlation between ACC activity (y-axis) and the RT difference for the LURE item condition (x-axis) is plotted. The ACC activity was extracted from the old minus new responses subtraction for the LURE items. The RT difference was computed by subtracting the RT for the new response from the RT for old response in the LURE condition. The correlation coefficient and statistical significance are also shown.



**Fig. 7.** Top: There were significantly higher levels of activation during old responses to both the OLD and LURE items, than during new responses to the NEW items in the right parahippocampal gyrus. The mean and s.e. of signal changes from the baseline fixation condition are plotted. Bottom: There were significantly higher levels of activation during old responses to both the OLD and LURE items than during new responses to the NEW items in the superior occipital gyrus. The mean and s.e. of signal changes from the baseline fixation condition are plotted. Resp., response type of the subject. See results section for details of the significant clusters.

The signals extracted from these regions were significantly positively correlated with the proportion of old responses across the 3 item types (Fig. 3, bottom, left amygdala,  $r = 0.30$ ,  $p < 0.05$ , left orbital gyrus,  $r = 0.47$ ,  $p < 0.01$ , right orbital gyrus,  $r = 0.27$ ,  $p < 0.05$ ).

To illustrate the relationship between the neural activity and each condition, the activation in the amygdala relative to the baseline fixation condition (measured by a separate one-sample  $t$ -test) was plotted against the 6 experimental conditions. This revealed a U-shaped relationship between the amygdala activity and the experimental condition, with the quadric polynomial function explaining 98% of the variance (Fig. 4). This indicates that activity was the highest under correct conditions (i.e., old response to OLD item or new response to NEW item) and lowest under incorrect conditions (i.e., new response to OLD item or old response to NEW item), and the level of amygdala activity measured during the experimentally induced false memory condition (i.e., old response to LURE item) fell in between.

To investigate the brain region specifically involved in processing false memory for faces, the LURE condition was compared with both the OLD and NEW conditions (Table 4). The comparison between the LURE and OLD conditions showed that a single cluster in the ACC had greater activity in the LURE condition than in the OLD condition (Fig. 5). The signal extracted from the ACC cluster had a significant positive correlation with the difference in the RT between the old and new responses for LURE items ( $r = 0.64$ ,  $p < 0.01$ , Fig. 6). There was no significant correlation between the ACC activity and the RT for either the OLD ( $r = -0.25$ , n.s.) or NEW ( $r = -0.15$ , n.s.) conditions. The comparison between the LURE and NEW conditions showed significant differences in the activation of the fusiform gyrus and the lateral prefrontal cortex (Table 4).

Previous studies have shown that activity levels in the medial temporal lobe are similar for both true and false memories, and significant activity occurs in primary sensory areas for true memories during retrieval (Schacter et al., 1996; Slotnick and Schacter, 2004). To investigate whether these findings hold true for memory for faces, activation under the baseline condition was compared to the

experimental conditions using an ANCOVA module. In line with the previous studies described, the right parahippocampal gyrus ( $x$ ,  $y$ ,  $z = 20$ ,  $-38$ ,  $-6$ ,  $T = 3.51$ ,  $p = 0.001$ , uncorrected,  $k = 26$  voxels) had significantly greater activation levels for both OLD and LURE items than for NEW items (Fig. 7, top). Both OLD and LURE items evoked greater activity in the occipital visual area ( $x$ ,  $y$ ,  $z = -10$ ,  $-86$ ,  $28$ ,  $T = 4.24$ ,  $p = 0.001$ , uncorrected,  $k = 48$  voxels) than NEW items did (Fig. 7, bottom). Furthermore, there was a significantly greater activity in the same region in the occipital visual area ( $x$ ,  $y$ ,  $z = -12$ ,  $-84$ ,  $28$ ,  $p = 0.001$ , uncorrected,  $k = 182$  voxels) for old responses to the OLD items than for old response to the NEW items. A similar result was obtained in the contrast between new responses to the OLD items and new responses to the New items ( $x$ ,  $y$ ,  $z = -12$ ,  $-86$ ,  $28$ ,  $p = 0.001$ , uncorrected,  $k = 42$  voxels).

## Discussion

The present study demonstrates 3 major findings concerning behavior and neuroanatomical substrates of false memory in normal human subjects. First, our procedure using a modified DRM paradigm with morphed faces successfully evoked false memory for faces in the fMRI environment. There have been many behavioral (Dodson et al., 2000; Jones and Bartlett, 2009; Jones et al., 2006; Reinitz et al., 1992; Roediger, 1996; Schacter and Dodson, 2001; Schacter and Slotnick, 2004; Searcy et al., 1999) and neuroimaging studies that have investigated false memory for words (Abe et al., 2008; Cabeza et al., 2001; Dennis et al., 2008; Garoff-Eaton et al., 2007; Heun et al., 2004; Kim and Cabeza, 2007; Paz-Alonso et al., 2008; Schacter et al., 1996, 1997; von Zerssen et al., 2001), pictures (Gonsalves et al., 2004; Okado and Stark, 2003), and shapes (Garoff-Eaton et al., 2006; Slotnick and Schacter, 2004); however, this is the first fMRI study to use the modified DRM paradigm and morphed face pictures to investigate false memory for faces in healthy subjects. The behavioral results of the present study clearly demonstrate that the proportion of old responses to the lure face stimuli (which had similar appearances to the previously presented, studied faces) was

significantly higher than that of the unstudied faces; these results confirm the occurrence of false recognition for the morphed faces.

Second, activity in the left amygdala and the bilateral orbital gyri, recorded during the recognition test, was significantly associated with the sense of familiarity with test items across the item conditions. Specifically, activity levels in the left amygdala were greatest when subject's responses were correct, lowest when responses were incorrect, and activity related with false memory fell somewhere in between. The finding that both old responses to the OLD items and new responses to NEW items (i.e., correct responses) elicited significant activation in the amygdala may relate to the notion that the amygdala plays a critical role in detecting the saliency (Phan et al., 2002) and the relevance (Ousdal et al., 2008) of a stimulus. Both the old and new correct conditions are salient and relevant to subjects in terms of mnemonic processing; the former involves familiarity and the latter novelty. It is perhaps noteworthy that the level of amygdala activity in the experimentally induced false memory condition (that evoked ambiguous and combined feelings with regard to familiarity and novelty) fell between the 2 extremes. This indicates that activation of the amygdala could partially differentiate between true and false memories, in contrast to parahippocampal activation that could not distinguish between these memories.

Predominant involvement of the amygdala in the detection of stimulus saliency has been previously reported in studies that used tasks in other domains of mental processing. Several reviews have shown that the human amygdala is activated in response to both positive and negative emotional stimuli (Costafreda et al., 2008; Fusar-Poli et al., 2009; Phan et al., 2002; Wager et al., 2003); similar patterns have also been observed in the primate amygdala (Paton et al., 2006). In other studies using fMRI and an evaluation task of face trustworthiness, both extremely trustworthy and untrustworthy faces elicited significant activation in the amygdala (Said et al., 2009; Todorov, 2008; Todorov et al., 2011). Furthermore, the amygdala activation is elicited by a threatening avoidance cue, as well as by a rewarding cue that predicts monetary loss or gain (Schlund and Cataldo, 2010). The amygdala activity in the present study may be enhanced in the 2 extremes of familiarity, i.e., old responses to old items and new responses to new items, because whether or not a person is familiar to the subject is important information in ordinary situations. The distribution of face stimuli including the 2 extremes and the middle of the continuum may be critical for observing such nonlinear amygdala responses. Therefore, it is believed that the amygdala plays a role in detecting salient stimuli features in several fields of human emotion and cognition, including face evaluation, reward, and memory.

The left and right orbital gyri have shown similar activation patterns to the amygdala; these results could be predicted from the dense anatomical connectivity evident between the amygdala and orbital cortex (Ghashghaei et al., 2007). A lesion study in humans showed that patients with damage to the orbito-frontal cortex (OFC) made more false positive responses to pictures in a memory task than those with damage outside the OFC (Schnider and Ptak, 1999). A positron emission tomography study on healthy humans also demonstrated that the OFC is significantly activated when subjects were presented with both relevant and irrelevant pictures, and suppressed the response to the irrelevant pictures (Schnider et al., 2000). These results indicate that the OFC is predominantly involved in a mechanism that selects memories of current relevance by suppressing irrelevant memories. Thus, the OFC activity measured in the present study may reflect successful memory processing associated with the suppression of irrelevant items during the false memory paradigm.

Thirdly, false memory for faces specifically involved activation of the ACC that significantly and positively correlates with the difference in RT observed for the LURE items. A number of studies have shown a significant relationship between activity in the ACC and false memory for words (Garoff-Eaton et al., 2007; Kim and Cabeza, 2007; von

Zerssen et al., 2001), shapes (Slotnick and Schacter, 2004), faces (Hofer et al., 2007), and pictures (Duarte et al., 2010; Okado and Stark, 2003); however, these findings have generally been regarded as of little importance. The anterior cingulate activity observed during false memory of words was interpreted as an association between performance monitoring and response conflict (von Zerssen et al., 2001). In another study that used pictures as memory items, the authors showed that the longer RT and higher ACC activity are associated under false memory conditions (Okado and Stark, 2003).

Several reviews of the literature have shown the ACC to be involved in handling increasing cognitive demands and RT differences when subjects were monitoring conflicts between competing representations (Barch et al., 2001; Carter and van Veen, 2007). Greater cingulate cortex activity associated with the memory paradigm has been reported in an fMRI study that evoked a tip-of-tongue state in the subjects (Maril et al., 2001). The authors of this study concluded that ACC activation may reflect the difficulty of the response decision. In another study that investigated the neural correlates of retrieval-related forgetting, the change in ACC activation was coupled with a conflict between competing mnemonic representations (Kuhl et al., 2007). In line with these findings, the present study has revealed that the ACC is specifically engaged in the false recognition of LURE items that have a perceptual resemblance to the OLD items, thus causing a response conflict between the items.

Another important point is that the engagement of ACC in false memory for faces is associated with a difference in RT between the old and new responses only for LURE items. The mean RT was significantly shorter for old than for new responses ( $p < 0.02$ , post-hoc paired  $t$ -test) only for OLD items; these results present a trend for the LURE ( $p = 0.08$ ) and are not significant for the NEW items ( $p = 0.99$ ). Therefore, it could be assumed that the greater the difference in the old minus new response RT, the more the subjects had a sense of familiarity for items. In the correlation analysis, the negative RT difference value, which could be a reflection of familiarity for the LURE items, correlated with less conflict for the LURE items as indexed by the ACC activity. This may suggest that ACC engagement in false memory is not only caused by an increasing effort and conflict but is also associated with higher mnemonic processing of the lure items. No significant difference was observed in the mean number of "sure" and "maybe" responses between the item types. Similar patterns of RT were observed in the OLD and LURE conditions when the response type was further divided into "sure" and "maybe" responses (see supplementary materials). Therefore, it is unlikely that specific involvement of the ACC in the LURE condition is caused by the difference in the RT between the item type conditions.

We partially replicated findings observed in previous neuroimaging studies on false memory. In our investigation, the parahippocampal gyrus was significantly activated under both true and false memory conditions. These results are in line with those of previous studies (Cabeza et al., 2001; Schacter et al., 1996, 1997; Slotnick and Schacter, 2004) indicating that involvement of the medial temporal lobe is common to both true and false memories. In addition, in the present study, both the OLD and LURE items evoked significantly greater activation than the NEW items in the occipital cortex during the test phase, presumably reflecting the fact that there were substantial physical and perceptual similarities between the OLD and LURE items. Activation in the occipital visual area was greater for the OLD items than for the NEW items when the response type was matched (i.e., both "old" and both "new"). This partially replicates a sensory reactivation hypothesis suggested by the previous studies (Okado and Stark, 2003; Schacter and Slotnick, 2004). In a study by Slotnick et al., prototypes and their exemplars of abstract shapes were created as non-verbal memory items (Slotnick and Schacter, 2004). During the test phase, neural activity measured in the occipital cortex was almost equivalent between previously seen shapes and related but novel shapes.

There are several limitations in the present study. First, although we considered that items with similar appearance would not be presented as NEW items in the test session, there was a single item for each NEW item that was created from the same original face due to the limited number of pictures in the face image database. This may confound the behavioral and neuroimaging results for the NEW condition. Second, in the behavioral data, the number of “surely old” and “surely new” responses was small, indicating a weak confidence level of responses during the recognition test. The subtraction between the old and new responses to each item type yielded significant activation in the prefrontal and parietal cortices (Table 2), a finding which was similar to that reported in previous studies (Duarte et al., 2010; Hofer et al., 2007). However, the activated regions associated with the old responses were relatively small, probably due to a weak confidence level during the task.

In conclusion, the present study has shown that false memory for faces could be successfully evoked in the fMRI experimental setting by using a modified DRM paradigm and morphed face pictures. The neural correlates for false memory for face were partially overlapping with those for true memories in terms of localization of the brain structures, namely, the left amygdala and the bilateral orbital cortices. However, levels of neural activity associated with false memory were between the memories that recognized the items correctly and incorrectly as either old or new. In contrast, activity in a specific region of the ACC may represent some aspects of response conflict and mnemonic processing of falsely recognized memory items as the activity level is significantly related with the performance speed.

## Funding

This work was supported by a Grant-in-Aid for Scientific Research (No. 20119004, face perception and recognition, and No. 21220005) and a Grant-in-Aid for Academic Frontier Project for Private Universities from the Japanese Ministry of Education, Science, Sports and Culture.

## Acknowledgments

The authors are grateful to Yohei Yamada, Yuka Nishiyama, and Takaaki Ogawa for their valuable help for the experiment.

No conflict of interest.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.neuroimage.2012.04.044>.

## References

- Abe, N., Okuda, J., Suzuki, M., Sasaki, H., Matsuda, T., Mori, E., Tsukada, M., Fujii, T., 2008. Neural correlates of true memory, false memory, and deception. *Cereb. Cortex* 18, 2811–2819.
- Barch, D.M., Braver, T.S., Akbudak, E., Conturo, T., Ollinger, J., Snyder, A., 2001. Anterior cingulate cortex and response conflict: effects of response modality and processing domain. *Cereb. Cortex* 11, 837–848.
- Cabeza, R., Rao, S.M., Wagner, A.D., Mayer, A.R., Schacter, D.L., 2001. Can medial temporal lobe regions distinguish true from false? An event-related functional MRI study of veridical and illusory recognition memory. *Proc. Natl. Acad. Sci. U. S. A.* 98, 4805–4810.
- Carter, C.S., van Veen, V., 2007. Anterior cingulate cortex and conflict detection: an update of theory and data. *Cogn. Affect. Behav. Neurosci.* 7, 367–379.
- Costafreda, S.G., Brammer, M.J., David, A.S., Fu, C.H., 2008. Predictors of amygdala activation during the processing of emotional stimuli: a meta-analysis of 385 PET and fMRI studies. *Brain Res. Rev.* 58, 57–70.
- Deese, J., 1959. On the prediction of occurrence of particular verbal instructions in immediate recall. *J. Exp. Psychol.* 58, 17–22.
- Dennis, N.A., Kim, H., Cabeza, R., 2008. Age-related differences in brain activity during true and false memory retrieval. *J. Cogn. Neurosci.* 20, 1390–1402.
- Dodson, C.S., Koutstaal, W., Schacter, D.L., 2000. Escape from illusion: reducing false memories. *Trends Cogn. Sci.* 4, 391–397.
- Duarte, A., Graham, K.S., Henson, R.N., 2010. Age-related changes in neural activity associated with familiarity, recollection and false recognition. *Neurobiol. Aging* 31, 1814–1830.
- Fusar-Poli, P., Placentino, A., Carletti, F., Landi, P., Allen, P., Surguladze, S., Benedetti, F., Abbamonte, M., Gasparotti, R., Barale, F., Perez, J., McGuire, P., Politi, P., 2009. Functional atlas of emotional faces processing: a voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. *J. Psychiatry Neurosci.* 34, 418–432.
- Garoff-Eaton, R.J., Slotnick, S.D., Schacter, D.L., 2006. Not all false memories are created equal: the neural basis of false recognition. *Cereb. Cortex* 16, 1645–1652.
- Garoff-Eaton, R.J., Kensinger, E.A., Schacter, D.L., 2007. The neural correlates of conceptual and perceptual false recognition. *Learn. Mem.* 14, 684–692.
- Ghashghaie, H.T., Hilgetag, C.C., Barbas, H., 2007. Sequence of information processing for emotions based on the anatomic dialogue between prefrontal cortex and amygdala. *NeuroImage* 34, 905–923.
- Gonsalves, B., Reber, P.J., Gitelman, D.R., Parrish, T.B., Mesulam, M.M., Paller, K.A., 2004. Neural evidence that vivid imagining can lead to false remembering. *Psychol. Sci.* 15, 655–660.
- Haxby, J.V., Hoffman, E.A., Gobbini, M.I., 2000. The distributed human neural system for face perception. *Trends Cogn. Sci.* 4, 223–233.
- Heun, R., Jessen, F., Klöse, U., Erb, M., Granath, D.O., Grodd, W., 2004. Response-related fMRI of veridical and false recognition of words. *Eur. Psychiatry* 19, 42–52.
- Hofer, A., Siedentopf, C.M., Ischebeck, A., Rettenbacher, M.A., Verius, M., Golaszewski, S.M., Felber, S., Fleischhacker, W.W., 2007. Neural substrates for episodic encoding and recognition of unfamiliar faces. *Brain Cogn.* 63, 174–181.
- Jones, T.C., Bartlett, J.C., 2009. When false recognition is out of control: the case of facial conjunctions. *Mem. Cogn.* 37, 143–157.
- Jones, T.C., Bartlett, J.C., Wade, K.A., 2006. Nonverbal conjunction errors in recognition memory: support for familiarity but not for feature bundling. *J. Mem. Lang.* 55, 138–155.
- Kim, H., Cabeza, R., 2007. Trusting our memories: dissociating the neural correlates of confidence in veridical versus illusory memories. *J. Neurosci.* 27, 12190–12197.
- Koutstaal, W., Schacter, D.L., Johnson, M.K., Galluccio, L., 1999. Facilitation and impairment of event memory produced by photograph review. *Mem. Cogn.* 27, 478–493.
- Kuhl, B.A., Dudukovic, N.M., Kahn, I., Wagner, A.D., 2007. Decreased demands on cognitive control reveal the neural processing benefits of forgetting. *Nat. Neurosci.* 10, 908–914.
- Maril, A., Wagner, A.D., Schacter, D.L., 2001. On the tip of the tongue: an event-related fMRI study of semantic retrieval failure and cognitive conflict. *Neuron* 31, 653–660.
- Okado, Y., Stark, C., 2003. Neural processing associated with true and false memory retrieval. *Cogn. Affect. Behav. Neurosci.* 3, 323–334.
- Ousdal, O.T., Jensen, J., Server, A., Hariri, A.R., Nakstad, P.H., Andreassen, O.A., 2008. The human amygdala is involved in general behavioral relevance detection: evidence from an event-related functional magnetic resonance imaging Go-NoGo task. *Neuroscience* 156, 450–455.
- Paton, J.J., Belova, M.A., Morrison, S.E., Salzman, C.D., 2006. The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature* 439, 865–870.
- Paz-Alonso, P.M., Ghetti, S., Donohue, S.E., Goodman, G.S., Bunge, S.A., 2008. Neurodevelopmental correlates of true and false recognition. *Cereb. Cortex* 18, 2208–2216.
- Phan, K.L., Wager, T., Taylor, S.F., Liberzon, I., 2002. Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage* 16, 331–348.
- Reinitz, M.T., Lammers, W.J., Cochran, B.P., 1992. Memory-conjunction errors: miscombination of stored stimulus features can produce illusions of memory. *Mem. Cogn.* 20, 1–11.
- Roediger III, H.L., 1996. Memory illusions. *J. Mem. Lang.* 35, 76–100.
- Roediger III, H.L., McDermott, K.B., 1995. Creating false memories: remembering words not presented in lists. *J. Exp. Psychol. Learn. Mem. Cogn.* 21, 803–814.
- Said, C.P., Baron, S.G., Todorov, A., 2009. Nonlinear amygdala response to face trustworthiness: contributions of high and low spatial frequency information. *J. Cogn. Neurosci.* 21, 519–528.
- Said, C.P., Dotsch, R., Todorov, A., 2010. The amygdala and FFA track both social and non-social face dimensions. *Neuropsychologia* 48, 3596–3605.
- Schacter, D.L., Dodson, C.S., 2001. Misattribution, false recognition and the sins of memory. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 356, 1385–1393.
- Schacter, D.L., Slotnick, S.D., 2004. The cognitive neuroscience of memory distortion. *Neuron* 44, 149–160.
- Schacter, D.L., Reiman, E., Curran, T., Yun, L.S., Bandy, D., McDermott, K.B., Roediger III, H.L., 1996. Neuroanatomical correlates of veridical and illusory recognition memory: evidence from positron emission tomography. *Neuron* 17, 267–274.
- Schacter, D.L., Buckner, R.L., Koutstaal, W., Dale, A.M., Rosen, B.R., 1997. Late onset of anterior prefrontal activity during true and false recognition: an event-related fMRI study. *NeuroImage* 6, 259–269.
- Schlund, M.W., Cataldo, M.F., 2010. Amygdala involvement in human avoidance, escape and approach behavior. *NeuroImage* 53, 769–776.
- Schneider, A., Ptak, R., 1999. Spontaneous confabulators fail to suppress currently irrelevant memory traces. *Nat. Neurosci.* 2, 677–681.
- Schneider, A., Treyer, V., Buck, A., 2000. Selection of currently relevant memories by the human posterior medial orbitofrontal cortex. *J. Neurosci.* 20, 5880–5884.

- Searcy, J.H., Bartlett, J.C., Memon, A., 1999. Age differences in accuracy and choosing in eyewitness identification and face recognition. *Mem. Cogn.* 27, 538–552.
- Slotnick, S.D., Schacter, D.L., 2004. A sensory signature that distinguishes true from false memories. *Nat. Neurosci.* 7, 664–672.
- Spaniol, J., Davidson, P.S., Kim, A.S., Han, H., Moscovitch, M., Grady, C.L., 2009. Event-related fMRI studies of episodic encoding and retrieval: meta-analyses using activation likelihood estimation. *Neuropsychologia* 47, 1765–1779.
- Todorov, A., 2008. Evaluating faces on trustworthiness: an extension of systems for recognition of emotions signaling approach/avoidance behaviors. *Ann. N. Y. Acad. Sci.* 1124, 208–224.
- Todorov, A., Said, C.P., Oosterhof, N.N., Engell, A.D., 2011. Task-invariant brain responses to the social value of faces. *J. Cogn. Neurosci.* 23, 2766–2781.
- von Zerssen, G.C., Mecklinger, A., Opitz, B., von Cramon, D.Y., 2001. Conscious recollection and illusory recognition: an event-related fMRI study. *Eur. J. Neurosci.* 13, 2148–2156.
- Wager, T.D., Phan, K.L., Liberzon, I., Taylor, S.F., 2003. Valence, gender, and lateralization of functional brain anatomy in emotion: a meta-analysis of findings from neuroimaging. *NeuroImage* 19, 513–531.
- Winston, J.S., O'Doherty, J., Kilner, J.M., Perrett, D.I., Dolan, R.J., 2007. Brain systems for assessing facial attractiveness. *Neuropsychologia* 45, 195–206.