Available online at www.sciencedirect.com
SCIENCE DIRECT®
Brain and Cognition 60 (2006) 176–186
Brain and Cognition
www.elsevier.com/locate/b&c

Hemodynamic and electrophysiological relationship involved in human face processing: Evidence from a combined fMRI–ERP study

Tetsuya Iidaka [a,c,*], Atsushi Matsumoto [a], Kaoruko Haneda [a], Tomohisa Okada [b], Norihiro Sadato [b,c]

[a] *Department of Psychology, Nagoya University, Graduate School of Environmental Studies, Chikusa, Nagoya, Aichi 464-8601, Japan*
[b] *Department of Cerebral Research, National Institute for Physiological Sciences, Okazaki, Aichi 444-8585, Japan*
[c] *Japan Science and Technology Corporation (JST)/Research Institute of Science and Technology for Society (RISTEX), Kawaguchi, Saitama 332-0012, Japan*

Accepted 21 November 2005
Available online 4 January 2006

Abstract

Functional magnetic resonance imaging (fMRI) and event-related potential (ERP) experiments were conducted in the same group of subjects and with an identical task paradigm to investigate a possible relationship between hemodynamic and electrophysiological responses within the brain. The subjects were instructed to judge whether visually presented stimuli were faces or houses and then press the corresponding button. Functional MRI identified face- and house-related regions in the lateral and medial part of the fusiform gyrus, respectively, while ERP showed significantly greater N170 negativity for face than for house stimuli in the temporo-occipital electrodes. Correlation analysis between the BOLD signal in the fusiform gyrus and ERP parameters demonstrated a close relationship between the signal and both latency and amplitude of N170 across the subjects. These correlations may indicate that the variation in cognitive demand and hemodynamic responses during the face/house discrimination task is coupled with the variation of N170 peak latency/amplitude across the subjects. Thus, integrative analysis of spatial and temporal information obtained from the two experimental modalities may help in studying neural correlates involved in a particular cognitive task.

Keywords: Fusiform gyrus; Correlation; N170; Latency; Amplitude

1. Introduction

Multimodal data acquisition and analysis are important strategies for the investigation of higher brain function in the field of cognitive neuroscience. Functional magnetic resonance imaging (fMRI) provides fine spatial information about neuronal activity during cognitive tasks, and event-related potential (ERP) experiments can measure electrophysiological changes in the brain with high temporal resolution. Integrative analysis of spatial and temporal information associated with a particular cognitive function can be achieved by combining these two experimental modalities. Several studies have used a combined fMRI and ERP experiment in the same group of subjects and the same task paradigm to test hypotheses pertaining to auditory attention (Horovitz, Skudlarski, & Gore, 2002; Menon, Ford, Lim, Glover, & Pfefferbaum, 1997; Opitz, Rinne, Mecklinger, von Cramon, & Schroger, 2002), visual attention (Ullsperger & von Cramon, 2001), and face recognition (Horovitz et al., 2002; Puce et al., 2003). Thus, a combination of these two methods for non-invasive examination of the brain and cognition in human subjects seems to be promising; however, the relationship between the blood oxygen level dependent (BOLD) signal obtained by fMRI and the amplitude or latency of the waveform measured by ERP is still unclear.

The neurophysiological basis of the fMRI signal (Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001; Ogawa et al., 1992) and electroencephalograph (EEG)

* Corresponding author. Fax: +81 52 789 4883.
E-mail address: iidaka@info.human.nagoya-u.ac.jp (T. Iidaka).

0278-2626/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.bandc.2005.11.004

(Kutas & Dale, 1997) has been extensively investigated; however the precise mechanisms involved in the transformation of neuronal activity to a BOLD signal or EEG waveform have not been fully elucidated. The hemodynamic responses measured by fMRI fluctuate in a matter of seconds, much more slowly than those monitored by EEG, and are thought to reflect local changes in oxy- and deoxyhemoglobin concentration (Logothetis et al., 2001; Ogawa et al., 1992). On the other hand, EEG waveforms have originated from the synchronized activity of neurons underlying the electrodes on the scalp, and their amplitude is substantially influenced by both the depth of the generator and the orientation of the dipole. Although recent evidence suggests that the BOLD signal has a tighter correlation with local field potentials than with multiunit spiking activity (Logothetis et al., 2001), the manner in which the results of fMRI and those of ERP are integrated in studies on brain activity under a particular task condition is still unclear. To investigate these issues, we adopted a combined fMRI and ERP study and a face/house discrimination task in which the subjects were required to judge whether a visually presented stimulus was a face or a house and then respond by pressing the corresponding button. Although the fMRI and ERP sessions were conducted separately, an identical task paradigm with a different set of stimuli was used, and the order of the sessions was counterbalanced across the subjects. In each session, the run was repeated twice with a randomly intermixed stimulus.

Face stimuli have been shown to elicit clear and large negative deflections in ERP at around 170 ms after the stimulus onset; these deflections are known as the "N170" component (Bentin, Allison, Puce, Perez, & McCarthy, 1996; Sagiv & Bentin, 2001). The N170 component is related to the detection and/or structural encoding of a face that is likely to be mediated by the ventral temporal cortex (Caldara et al., 2003; Itier & Taylor, 2002; Jemel, George, Olivares, Fiori, & Renault, 1999; Schweinberger & Burton, 2003; Schweinberger, Pickering, Jentzsch, Burton, & Kaufmann, 2002). However, several other authors suggest that the source of N170 is located in the lateral part of the temporal lobe (Henson et al., 2003; Itier & Taylor, 2004; Puce et al., 2003). The results of intracranial recordings in epileptic patients indicated that there were two major areas involved in face processing: one in the ventral region including the fusiform gyrus and the other in the lateral region located in the middle temporal gyrus (Allison, Puce, Spencer, & McCarthy, 1999; McCarthy, Puce, Belger, & Allison, 1999; Puce, Allison, & McCarthy, 1999). Furthermore, single cell recordings in monkeys showed that neurons that responded to face stimuli were found both in the superior temporal sulcus and in the inferior temporal gyrus (Rolls, 1992). Therefore, it is still unclear whether N170 originates in the ventral or lateral temporal region or in both of them.

Functional MRI studies have consistently shown that viewing face stimuli activates the lateral part of the fusiform gyrus whereas, viewing houses activates the parahippocampal gyrus and the medial part of the fusiform gyrus (Epstein & Kanwisher, 1998; Gorno-Tempini & Price, 2001; Hadjikhani et al., 2004; Kanwisher, McDermott, & Chun, 1997; Maguire, Frith, & Cipolotti, 2001). On the other hand, some authors (Henson et al., 2003; Puce et al., 2003) have indicated that the lateral temporal region is involved in face processing. These discrepancies with regard to the localization of neural activity and the source of evoked potentials may be attributed to the fact that the BOLD signals and the evoked potential were measured in separate groups of subjects. Therefore, we considered that investigation of a correlation between the data sets obtained from fMRI and ERPs in the same group of subjects would help in resolving these issues. Thus, one goal of the present study is to examine the functional relationship between the BOLD signal and ERP parameters using a combined event-related fMRI and ERP experiment and a face/house discrimination task. We conducted an fMRI subtraction analysis and a comparison of the grand mean waveforms of the N170 component between the face and house conditions. The N170 parameters were then entered into a simple regression analysis to test whether these values correlated with the BOLD signal within the temporo-occipital lobes. We predicted that the N170 amplitude would correlate with the magnitude of the BOLD signal in the temporo-occipital areas under the face condition or the face condition with the house condition variance removed.

2. Methods

2.1. Subjects

Twelve right-handed healthy subjects (six males, mean age ± *SD*, 20.8 ± 1.6 years) participated in the experiment after providing written informed consent. This study was approved by the ethics committee of the National Institute for Physiological Sciences.

2.2. Experimental procedure

Digitized grayscale pictures of 50 faces with neutral expressions taken from posers (25 male and 25 female) and pictures of 50 houses created using computer graphics software (Aska-Pro, Logic, Japan) served as the stimuli (Fig. 1). These pictures were divided into two sets of stimuli that were assigned to each of the fMRI and ERP sessions. The luminance of the pictures was equated. In each run, 25 faces, 25 houses, and 25 null events with fixation were randomly presented to the subject one at a time with 500 ms duration and 4500 ms interstimulus interval. During the interval, a fixation point was always shown. As a warning, the color of the fixation point turned from black to red 400 ms before stimulus onset. The subject was required to not blink while the red fixation point and experimental stimuli were on the screen.

The stimuli were projected onto the transparent screen hanging on the bore of the magnet that was placed

Fig. 1. Examples of face (upper) and house (lower) stimuli.

approximately 75 cm from the subjects' eyes. They viewed the stimuli through a tilted mirror attached to the head coil of the scanner. The task was to judge whether the presented stimulus was a face or a house and then press the corresponding button with his or her right hand. Each run was repeated twice with the same set of randomly intermixed stimuli. An identical task paradigm was used for each of the fMRI and ERP sessions with a different stimulus set. Half of the subjects performed the fMRI session first, while the other half performed the ERP session first. The fMRI and ERP sessions were separated by approximately two days (mean, 1.8 days).

2.3. ERP data acquisition and analysis

The EEG was recorded from 14 international 10–20 system scalp locations (Fz, F3, F4, Cz, C3, C4, Pz, P3, P4, T5, T6, Oz, O1, and O2) referenced to the tip of the nose. Eye movement was monitored by an electrode placed on the supraorbital ridge of the left eye. Interelectrode impedance was set below 5 kΩ. The EEG and EOG data were filtered using a band-pass of 0.5–60 Hz. The data were digitized with an AD conversion rate of 1000 Hz and sampled from 100 ms before the stimulus onset to 500 ms after it. EEG data were corrected to a 100 ms baseline prior to the stimulus onset. Trials in which the EEG or eye movement exceeded ±50 μV were automatically rejected from the averaging process. The ERP experiment was controlled by SuperLab software (Cedrus, San Pedro, CA, USA).

EEG data were sorted according to each type of response (correct responses for face and correct responses for house were included in the analysis), and the average ERP waveforms were calculated for each condition and subject using EEG Expert software (ver. 4.5, Brain Function Laboratory, Kawasaki, Japan). In each subject, a negative deflection of the ERP waveform was identified at around 170 ms after the stimulus onset (the N170 component) at the temporo-occipital electrodes (Oz, O1, O2, T5, and T6). The peak amplitude and latency of N170 were measured in each subject and entered into a 3-way ANOVA with condition, electrode, and run as the factors. Fig. 2A illustrates the grand mean waveforms for 12 subjects under the face and house conditions. A topographic scalp map of the difference wave between the face and house condition at 187 ms poststimulus onset (Fig. 2B) was computed using EEGLAB ver. 4.4 (Delorme & Makeig, 2004).

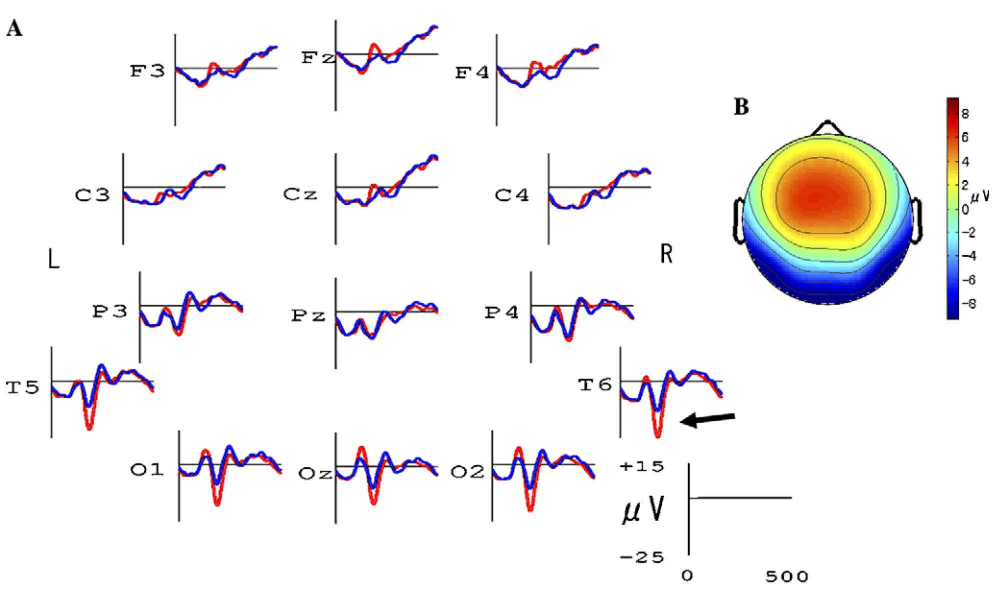

Fig. 2. Results of the ERP experiment are shown. (A) Grand mean waveforms for the face (red) and house (blue) conditions averaged across 12 subjects. The original waveform was smoothed by a low-pass filter of 30 Hz. An arrow indicates a significantly larger N170 for the face condition than for the house condition at the T6 electrode. (B) Topographic scalp mapping of the difference wave between the face and house conditions at 187 ms poststimulus onset. The face-related negative potential (as indicated in blue) is predominantly observed in the bilateral temporo-occipital regions.

2.4. fMRI data acquisition and analysis

Functional images of the whole brain were acquired in an axial orientation using a 3-Tesla Siemens Allegra MRI scanner equipped with single-shot EPI (TR = 2.3 s, TE = 30 ms, flip angle = 80°, 64 × 64 matrix and 26 slices, voxel size = 3 × 3 × 4 mm) sensitive to BOLD contrast (Ogawa et al., 1992). After discarding the first 6 images, the next 166 successive images in each run were subjected to analysis. An anatomical T1-weighted image was also acquired (MPRAGE, TR = 3 s, TE = 4.6 ms, flip angle = 90°, 256 × 256 matrix and 26 slices, voxel size = 0.75 × 0.75 × 4 mm) for each subject. The fMRI experiment was controlled using Presentation software (Neurobehavioral Systems, Albany, CA, USA).

Data were analyzed by SPM99 (the Wellcome Department of Imaging Neuroscience, London, UK). First, all volumes were realigned spatially to the last volume, and the signal in each slice was realigned temporally to that obtained in the middle slice using a sinc interpolation. The resliced volumes were normalized to the standard space of Talairach and Tournoux (1988) using a transformation matrix obtained from the normalization process of the T1-weighted anatomical image of each individual subject to the T1 template image. The T1-weighted anatomical image was coregistered to the mean EPI image in each subject. The normalized images were spatially smoothed with an 8-mm Gaussian kernel. Following preprocessing, statistical analysis of each individual subject was conducted using the general linear model. At the first level, each single event was modeled as a hemodynamic response function (HRF) and its temporal derivative. Low-pass (HRF) and high-pass (mean cutoff period of 75.2 s) frequency filters were applied to the timeseries data. The images were scaled to a grand mean of 100 over all voxels and scans within a session. In the subtraction analysis, three conditions (correct response for face, correct response for house, and incorrect response) were modeled separately. Parameter estimates for each condition and for the difference between the conditions were calculated from the least mean square fit of the timeseries data to the model. Images of parameter estimates representing event-related activity at each voxel for each condition and each subject were created. In a region of interest (ROI) analysis, the region in the left and right fusiform gyrus where activation was larger under the face than under the house condition was identified in each subject. The statistical threshold was set individually in each subject (T value ranged from 1.03 to 8.02). The mean parameter estimate in a spherical ROI ($r = 8$ mm) created at the peak voxel was then extracted from the contrast image for the face minus house subtraction condition.

At the second level, the results for each individual subject were entered into the random effects model by applying t tests between the contrast images to create a group statistical parametric map (SPM). An SPM of voxels showing a significant response to stimulus presentation versus baseline and differences in the response between the conditions were created. Two analyses were conducted: (1) subtraction analysis between the conditions to identify face- and house-related regions and (2) correlation analysis with N170 latency and amplitude to investigate possible relationship between the BOLD signal and electrophysiological responses. The first analysis was done by entering contrast images containing parameter estimates representing the difference in activity between the face and house conditions into a one-sample t test. The statistical threshold was set at $p = .001$ (uncorrected), and clusters larger than 5 voxels were listed. Region names, coordinates, and Z values are tabulated in Table 1. For the purpose of presentation, the distribution of face- and house-related regions is illustrated in Fig. 3 using MRIcro software (Rorden & Brett, 2000).

The second analysis addressed the question of whether the N170 amplitude and latency would correlate with the BOLD signal obtained during the face/house discrimination task. The contrast images of 12 subjects that pertained to activity under the face condition, house condition or the difference in activity between the face and house conditions were used. The ERP parameters obtained from each subject and condition were separately entered into a simple regression analysis implemented in SPM99. These parameters are: (1) the peak amplitude of N170 under the face condition, (2) the peak amplitude of N170 under the house condition, (3) the difference in the N170 peak amplitude between the face and house conditions, (4) the peak latency of N170 under the face condition, and (5) the peak latency of N170 under the house

Table 1
Face- and house-related regions as revealed by fMRI subtraction analysis

Region name	L/R	Coordinates	Z value	Size (voxels)
Face minus house				
Hippocampus	L	−14, −14, −20	4.16	132
Hippocampus	R	18, −18, −16	4.79	82
		34, −12, −18	4.22	7
Med. front. gy.	L	−6, 32, 48	4.01	21
		−2, 56, 26	3.35	5
	R	10, 42, 38	3.92	43
		10, 16, 56	3.65	9
Inf. front. gy.	L	−44, 30, −18	3.86	6
Inf. temp. gy.	R	30, −12, −34	3.53	5
Inf. pari. lob.	R	54, −48, 46	3.45	6
Mid. front. gy.	R	42, 28, 18	3.44	6
Fusiform gy.	L	−44, −62, −24	3.28	5
House minus face				
Fusiform gy.	L	−28, −52, −18	4.58	666
Fusiform gy.	R	30, −50, −16	5.51	1266
SMA	L	−10, −6, 52	4.54	21
S/M	L	−30, −32, 60	4.34	27
		−42, −16, 50	3.5	15
		−40, −22, 42	3.31	7
Lingual gy.	R	12, −88, 12	4.28	140
Precuneus	R	20, −74, 50	4.24	94
Occi. gy.	R	40, −80, 24	4.12	202
Mid. occi. gy.	L	−32, −90, 10	4.07	178
Cuneus	L	−12, −86, 12	3.56	23

L/R: left/right hemisphere, inf.: inferior, mid.: middle, front.: frontal, occi.: occipital, temp.: temporal, gy.: gyrus, SMA: supplementary motor area, S/M: primary sensory/motor area.

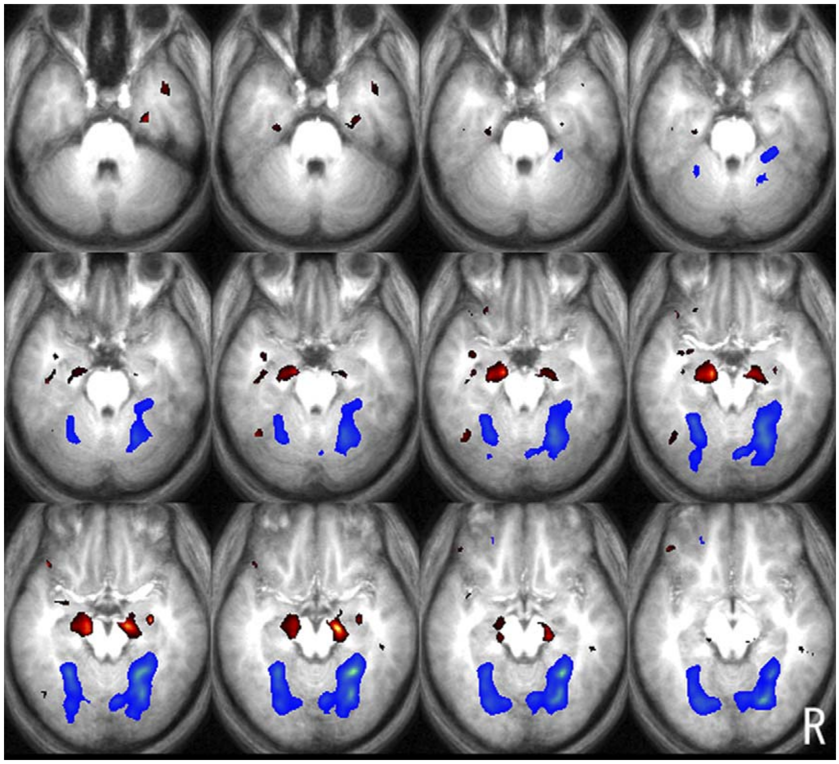

Fig. 3. Results of subtraction analysis in the fMRI experiment. Face- (red) and house-related (blue) regions are superimposed on a mean normalized T1-weighted anatomical image of 12 subjects using MRIcro software. To present the distributed pattern of activation in the temporo-occipital lobe, the statistical threshold is set at $p = .005$ (uncorrected). Horizontal images are shown from $z = -34$ mm (upper-left) to $z = -12$ mm (lower-right) in 2 mm increments. For statistical assessment thresholded at $p = .001$ (uncorrected) see Table 1.

condition. The difference in the N170 peak latency between the face and house condition was not included because this parameter did not differ significantly between the conditions. Specifically, the difference in the peak amplitude was computed by partialling out the house condition from the face condition by regression and saving the residual. The parameters were measured at the O1, O2, Oz, T5, and T6 electrode sites. The statistical threshold was set at $p = .001$ (uncorrected, $T = 4.14$), and clusters larger than 10 voxels within the temporo-occipital lobe were reported in Tables 2–4. Figs. 4 and 5 show the clusters in the inferior temporal lobe where a significant fMRI–ERP correlation was observed. In Fig. 6, parameter estimates from the peak voxel of four clusters and the latency in each electrode are plotted.

3. Results

3.1. Behavioral data

The two runs in each of the ERP and fMRI sessions were collapsed as there was no significant difference in the RT and error rate between the first and second runs (3-way ANOVA, $F(1, 88) = .57$, $p = .45$). The mean RTs (ms, *SD* in parentheses) for the face and house conditions were 436 (64) and 437 (63) in the ERP session, and 427 (57) and 452 (62) in the fMRI session, respectively. A 2-way ANOVA showed that there was no main effect of condition ($F(1, 44) = .56$, $p = .45$), session ($F(1, 44) = .02$, $p = .87$), or the interaction ($F(1, 44) = .44$, $p = .51$). The mean error rate (%) was 1.6 (1.6) for the face condition and 2.6 (3.4) for the house condition in an ERP session, and 0.3 (0.7) for the face condition and 1.0 (1.5) for the house condition in an fMRI session. The subjects made more errors in the ERP session ($F(1, 44) = 6.06$, $p < .05$) than in the fMRI session, and there was no main effect of condition ($F(1, 44) = 1.87$, $p = .17$) or the interaction ($F(1, 44) = .07$, $p = .78$).

Table 2
Temporo-occipital regions with significant correlation between BOLD signal and N170 amplitude

Electrode	Region name	L/R	Coordinates	*Z* value	Size (voxels)
(a) Face condition					
O2:	Lingual gy.	R	18, −88, −10	3.77	15
(b) House condition					
T5:	Middle occipital gy.	L	−14, −94, 18	4.13	32
	Lingual gy.	R	14, −94, −10	3.92	17
T6 :	Lingual gy.	R	12, −94, −10	3.66	13
Oz:	Lingual gy.	L	16, −88, −10	3.96	16
O1:	Lingual gy.	R	14, −90, −10	4.21	43
	Middle occipital gy.	L	−14, −94, 18	3.95	31
(c) Face minus house condition					
T6:	Fusiform gy.*	R	34, −60, −12	4.28	48
O2:	Lingual gy.*	R	30, −76, −12	3.72	34

An asterisk indicates a cluster shown in Fig. 4.
The statistical threshold is $p = .001$ (uncorrected) and $k = 10$ voxels.

Table 3
Temporo-occipital regions with significant correlation between BOLD signal and N170 latency for face

Electrode	Region name	L/R	Coordinates	Z value	Size (voxels)
T5:	Fusiform gy.*	L	−40, −56, −18	3.63	45
	Inf. temp. gy.	R	44, −6, −26	3.71	29
	Fusiform/mid. occi. gy.	R	32, −78, −6	3.62	37
T6:	Fusiform gy.	L	−42, −60, −20	4.53	66
	Fusiform gy.	L	−38, −30, −18	3.59	36
	Fusiform/mid. occi. gy.*	R	44, −66, −4	3.67	50
	Inf. temp. gy.	R	36, −6, −26	3.52	83
O1:	Fusiform gy.*	L	−42,−60,−18	3.76	27
	Fusiform gy.	R	34, −60, −24	3.73	46
O2:	Fusiform gy.	L	−42, −60, −22	4.01	44
	Fusiform gy.	L	−40, −30, −18	3.38	11
	Inf. temp. gy	R	36, −6, −26	3.43	35
Oz:	Fusiform gy.*	L	−42, −58, −16	3.68	14
	Fusiform gy.*	R	32, −62, −20	3.72	17

An asterisk indicates a cluster shown in Fig. 5.
The statistical threshold is $p = .001$ (uncorrected) and $k = 10$ voxels.

Table 4
Temporo-occipital regions with significant correlation between BOLD signal and N170 latency for house

Electrode	Region name	L/R	Coordinates	Z value	Size (voxels)
T6:	Parahippocampal gy.	L	−26, −2, −28	3.86	21
	Parahippocampal gy.	R	24, −22, −30	3.86	13
O1:	Cuneus	R	24, −98, 2	4.03	18
O2:	Amygdala–hippocampus	L	−20, −10, −18	3.75	31
	Amygdala–hippocampus	R	24, −6, −18	3.69	10
	Fusiform gy.	L	−24, −28, −28	3.65	25
Oz:	Middle temporal gy.	L	−54, 0, −22	3.71	17

The statistical threshold is $p = .001$ (uncorrected) and $k = 10$ voxels.

3.2. ERP data

As shown in Fig. 2A, N170 is clearly defined in the temporo-occipital electrodes (T5, T6, Oz, O1, and O2), and the peak amplitude is more negative under the face condition than under the house condition. The topographic scalp map of the difference wave between the face and house condition (Fig. 2B) also shows that face-related negativity is dominant in the bilateral temporal regions at 187 ms poststimulus onset. A 3-way ANOVA for N170 peak amplitude conducted with electrode, condition, and run as the factors revealed a significant main effect of condition ($F(1, 220) = 50.5$, $p < .001$) and no significant effect of electrode ($F(4, 220) = 1.85$, $p = .11$). The amplitude tended to be more negative in the first run than in the second run ($F(1, 220) = 3.41$, $p = .06$). There was no significant 2-way interaction effect between condition and electrode ($F(4, 220) = .16$, $p = .95$), between run and electrode ($F(4, 220) = .03$, $p = .99$), or between condition and run ($F(1, 220) = 2.74$, $p = .09$). There was no significant 3-way interaction effect on N170 amplitude ($F(4, 220) = .08$, $p = .98$). A 3-way ANOVA conducted on N170 peak latency did not show a significant main effect or interaction. There was no significant correlation between mean RT and N170 latency (T6; $r = .08$, $p = .79$) or between mean RT and N170 amplitude (T6; $r = -.22$, $p = .47$) under the face condition.

3.3. fMRI data

Subtraction analysis between the face and house conditions revealed the brain regions that were predominantly activated under each condition as listed in Table 1. Fig. 3 illustrates the distributed patterns of the activated clusters in the temporo-occipital lobe. Face-dominant regions are found in the bilateral medial temporal lobe including the hippocampus/amygdala and in the lateral part of the fusiform gyrus, while house-dominant regions are located in the occipital lobe and the medial part of the fusiform gyrus that extends to the parahippocampal gyrus. The topographic patterns of activation under each condition are similar to the results of previous neuroimaging studies suggesting that face- and house-related regions are found in the lateral and medial part of the fusiform gyrus, respectively. In a ROI analysis of each individual subject, the left and right fusiform gyri where the signal was greater under the face condition than under the house condition was identified. The mean coordinates (x, y, z) of these peaks were (−38, −58, −18) for the left hemisphere and (43, −58, −16) for the right hemisphere, respectively. The standard deviation of coordinates, particularly in the y-axis, was larger in the right than in the left hemisphere (left vs. right: x-axis, 4.3 vs. 3.9; y-axis, 5.4 vs. 8.4; z-axis, 7.9 vs. 8.2). The mean parameter estimates extracted from the ROIs in the left and right fusiform gyri did not differ between the hemispheres (left vs. right, 0.11 ± 0.08 vs. 0.12 ± 0.12, $F = .02$, $p = 0.88$).

3.4. fMRI–ERP correlation

First, the BOLD signal obtained under the face or house condition (i.e., activation from the fixation baseline) was correlated with the N170 amplitude of each condition (Table 2(a) and (b)). The results showed that the regions with significant BOLD-amplitude correlation were located in the occipital cortices. Second, the difference in BOLD signal magnitude between the face and house conditions was correlated with the difference in the N170 peak amplitude between the conditions. There were two clusters in the fusiform and lingual gyri of the right hemisphere that showed a significant negative correlation between the signal and amplitude measured at the T6 and O2 electrodes (Table 2(c), Fig. 4). As plotted in Fig. 4, the larger the N170 negativity, the greater the increase in the face-related signal in these regions under the face minus house condition.

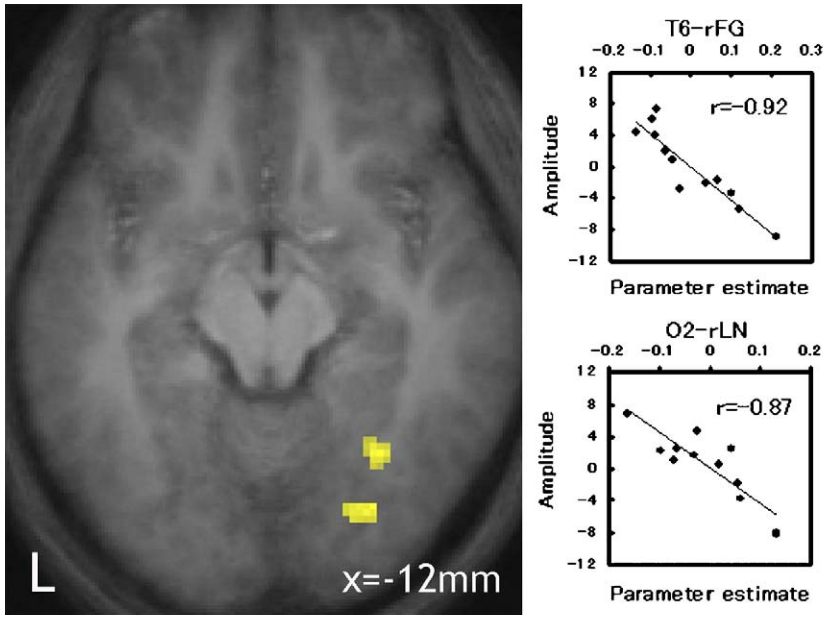

Fig. 4. Results of the fMRI–ERP correlation analysis are shown in the horizontal image. The difference in the BOLD signal between the face and house conditions had a significant negative correlation with the difference in N170 amplitude between the face and house conditions. Due to the negativity of the component, the negative correlation indicates that the larger the difference in amplitude, the greater the difference in the BOLD signal. Left: The yellow areas are clusters in the right fusiform ($x, y, z = 34, -60, -12$) and lingual gyrus ($x, y, z = 30, -78, -14$) that have a significant correlation between the signal and amplitude measured at the T6 and O2 electrodes, respectively. For more details see Table 2(c). Right-top: Significant correlation between BOLD signal (x-axis: the difference in parameter estimate between the conditions, arbitrary unit) in the right fusiform gyrus (rFG) and N170 amplitude (y-axis, the residuals of the regression between the conditions, μV) in the T6 electrode. Right-bottom: Significant correlation between the signal in the right lingual gyrus (rLN) and O2 electrode. A regression line and correlation coefficient are shown in the figure. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

A significant fMRI–ERP correlation was observed when the peak latency of each individual N170 was regressed with the magnitude of activation from the baseline under the face or house condition. As shown in Table 3, the correlation under the face condition was found in every temporo-occipital electrode, and the clusters tend to converge in the bilateral fusiform gyrus (Figs. 5A and B). N170 latency in the T5, T6, O1, and O2 electrodes and parameter estimates of significant voxels in the temporal lobe are plotted in Fig. 6. The correlation was also found in temporo-occipital electrodes under the house condition; however, these clusters did not overlap with those found under the face condition (Table 4). There was no significant correlation between the mean RT during the face condition and the BOLD signal in the fusiform gyrus. When the N170 latency under the face condition was correlated with the BOLD signal under the face minus house condition, there was no significant region with latency–signal correlation in the temporo-occipital areas. The mean signal change between the face and house conditions extracted from ROIs in each subject did not correlate with either amplitude (T6; $r = -.12, p = .71$) or latency (T6; $r = .01, p = .96$) of N170. For the ROI data there was no significant correlation between the BOLD signal for the face condition with variability of the house condition partialled out and the N170 amplitude for the face condition with variability of the house condition partialled out ($r = .24, p = .45$).

4. Discussion

The present study uses the face/house discrimination task to address the question of whether any relationship exists between the two noninvasive measures of brain function, i.e., the BOLD signal of fMRI and the electrophysiological activity of ERP. Specifically, correlating the BOLD signal in the fusiform gyrus and N170 parameters in a combined fMRI and ERP experiment where the same procedure and same group of subjects are used for both sessions would answer the question. The major finding was that voxel-by-voxel correlation between the BOLD signal and ERP parameters (i.e. amplitude and latency) in the temporo-occipital lobe did show a strong relationship between these measures.

4.1. Separate analyses for ERP and fMRI experiments

In an ERP experiment, consistent with the results of previous studies (Bentin et al., 1996; Carmel & Bentin, 2002; Eimer, 2000a; Sagiv & Bentin, 2001), the amplitude of N170 was significantly larger when viewing faces than when viewing houses, while there was no significant difference in N170 peak latency between the two conditions. The face-related negativity was predominantly observed in the bilateral temporo-occipital regions as illustrated in the topographic map (Fig. 2B). N170 has been found to be elicited during tasks involving face processing; however, it should be noted

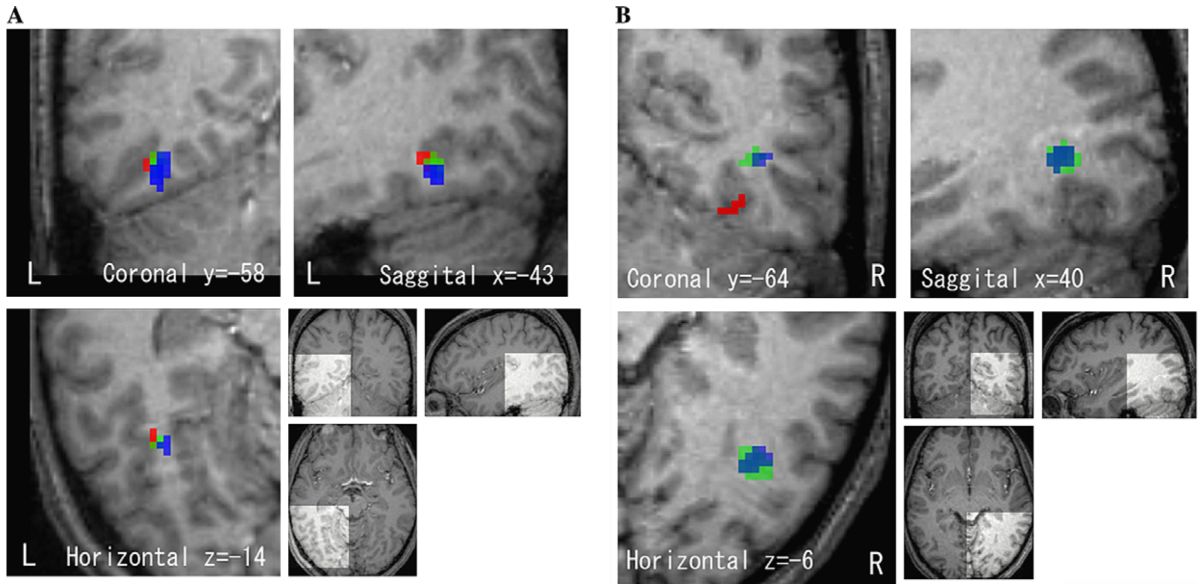

Fig. 5. Results of the fMRI–ERP correlation analysis are shown in three orthogonal views. The magnitude of the BOLD signal increase from the baseline measured in the bilateral temporo-occipital clusters had a significant positive correlation with N170 latency while viewing a face. Highlighted areas in the bottom-right quadrant in each panel are magnified in other quadrants. (A) Colored areas are clusters in the left fusiform gyrus with a significant correlation between signal and latency measured at three electrodes. Red: Oz, green: O1, and blue: T5. (B) Clusters in the right fusiform/middle occipital gyrus. Red: Oz, green: O2, and blue: T6. For region names, coordinates, and other information see Table 3. Note that figure (B) shows a cluster in the right fusiform/middle occipital gyrus (green) that is not listed in Table 3 due to its small voxel size (5 voxels).

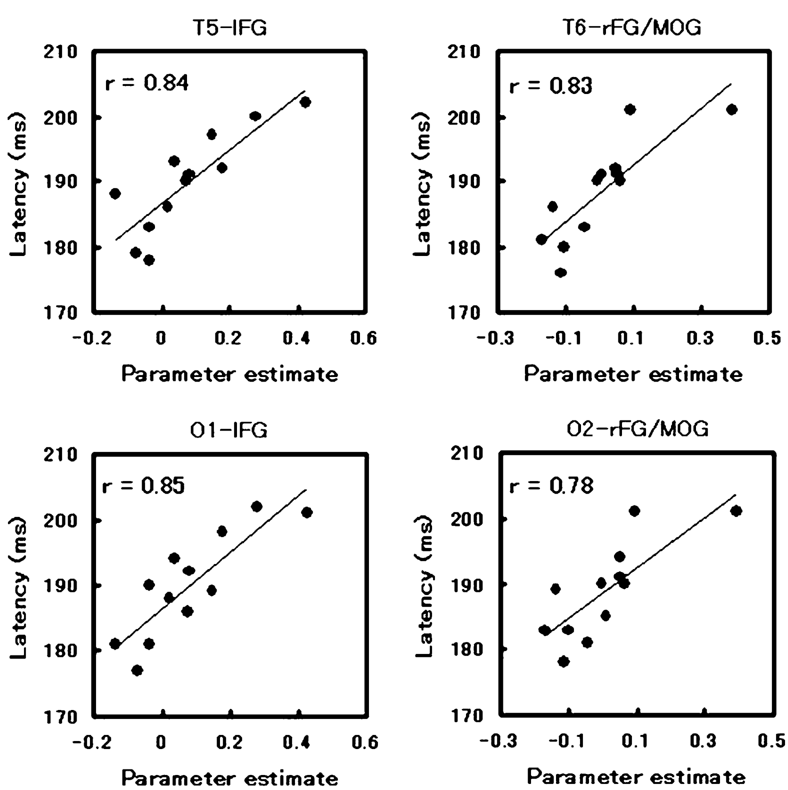

Fig. 6. Significant positive correlations between the BOLD signals (*x*-axis: parameter estimate in an arbitrary unit) from four representative clusters (lFG: left fusiform gyrus, rFG: right fusiform gyrus, and MOG: middle occipital gyrus) and the N170 latency (*y*-axis, ms) in each electrode (O1, O2, T5, and T6) are plotted. A regression line and correlation coefficient are shown in the figure.

that the appearance of N170 of much lower amplitude has also been reported for other visual stimuli such as houses, cars, and birds (Bentin et al., 1996; Carmel & Bentin, 2002). The finding that the amplitude of N170 is unaffected by face repetition (Eimer, 2000b) suggests that it reflects structural encoding or detection of a facial pattern in which the stimulus is classified as a face, rather than recognition of an individual face.

In an fMRI experiment, face- and house-related regions are dissociated in the temporo-occipital lobes as shown in Fig. 3. Apparently, the face-related regions (red) were distributed in the bilateral amygdala/hippocampus and the lateral part of the fusiform gyrus, whereas the house-related regions (blue) were located in the medial part of the fusiform gyrus that extended to the parahippocampal gyrus and the occipital cortices. The results are in accordance with neuroimaging studies (Gorno-Tempini & Price, 2001; Hadjikhani et al., 2004; Kanwisher et al., 1997; Maguire et al., 2001) and show that the lateral part of the fusiform gyrus is particularly active while viewing a face. However, in contrast to the previous studies (Epstein & Kanwisher, 1998; Gorno-Tempini & Price, 2001; Hadjikhani et al., 2004; Kanwisher et al., 1997; Maguire et al., 2001), the difference in the signal between the face and house conditions in the right fusiform gyrus did not survive the statistical threshold in the present fMRI experiment. The results of ROI analysis indicate that the location of the face-related activity along the *y*-axis may be more variable in the right than in the left fusiform areas in our group of subjects, although the level of activity is comparable in the two hemispheres. We speculate that the variability of localization in the right fusiform activity reduced the significance in the group random-effect analysis.

To date, several combined fMRI and ERP studies using face stimuli have been published (Henson et al., 2003; Horovitz et al., 2002; Puce et al., 2003); however, two of these studies did not elucidate the correlation between BOLD responses and ERP parameters. Although the focus of the study by Puce et al. (2003) was to investigate neural activity associated with viewing facial motion and not static faces, they suggested that N170 would be elicited by activation in the lateral temporo-occipital cortex. The tasks employed in their study (Puce et al., 2003) emphasized visual attention to moving aspects of facial structures such as the mouth and eyes; therefore, the predominant cortical activity might have been evoked in the lateral part of the temporo-occipital lobes.

A study by Henson et al. (2003) examined the recognition and priming of static face pictures using fMRI and ERP in a different cohort of subjects. The authors suggested that the superior temporal sulcus (STS) is the more likely generator of N170. Although no dipole estimation was performed in their study, this hypothesis is inconsistent with the results of previous ERP studies (Caldara et al., 2003; Itier & Taylor, 2002; Jemel et al., 1999; Schweinberger et al., 2002) that identified dipoles for a static face in the inferior temporal areas. In our study, no significant activation was found in the right STS region under the face condition or face minus house condition when the statistical threshold was reduced to $p = .01$ (uncorrected) level. A discrepancy between the two studies may have arisen from the fact that the mean RT for behavioral tasks is much longer in the study by Henson et al. (approximately 1500 ms) than in our study (less than 500 ms). Henson et al. (2003) required the subjects to rate the left–right symmetry of facial pictures, and this procedure possibly encouraged the subjects to pay more attention to detailed structures such as the eyes and mouth, which consistently activate the STS (Haxby, Hoffman, & Gobbini, 2000). Thus, we speculate that their results are complicated by a long RT and cognitive demands that are specific to their task paradigm.

4.2. *fMRI–ERP correlation*

To investigate the functional relationship between the hemodynamic and electrophysiological measures more extensively, a correlation analysis of fMRI and ERP parameters was conducted. A voxel-by-voxel analysis showed that the difference in the BOLD signal in the right fusiform gyrus between the conditions was negatively correlated with the difference in N170 peak amplitude. A ROI analysis in the fusiform face area did not show such a correlation. These results indicate that the areas showing a signal–amplitude correlation are located in the border or surrounding region of the face-related areas but not in the face-related region per se. The extent to which this area was activated by the presentation of faces varied among the subjects and was critically associated with the difference in ERP amplitude. We used the residuals of the regression between the conditions to remove the variance of the N170 amplitude for house condition from the amplitude for face condition. Therefore, it is suggested that the correlation is mainly derived from the face condition rather than from a generalized N170 response.

In addition, we found a positive correlation between the magnitude of increase in the BOLD signal in the bilateral inferior temporal areas and the peak latency of N170 under the face condition. The regions with a signal–latency correlation were distributed extensively in the bilateral inferior temporal lobe including the fusiform gyri. These results imply that hemodynamic activity in these regions possibly reflected electrophysiological activity, which represents timing and the degree of maximum neuronal synchronization associated with face processing. The signal–latency correlation was also observed in the temporal lobe under the house condition; however, these regions did not overlap with the regions found under the face condition.

An advantage of using the same group of subjects in both experiments is that correlation analysis can investigate possible hemodynamic-electrophysiological relationships; however, to date this analysis has been successfully conducted in only a few studies. Opitz et al. (2002) found a weak correlation between the amplitude of mismatch negativity and the BOLD signal in the fronto-temporal foci. In a parametric study of P300 in an auditory oddball task (Horovitz et al., 2002), the authors manipulated the probability of the target and observed a significant positive correlation between the magnitude of P300 and the BOLD signal in the supramarginal gyrus. Mathalon, Whitfield, and Ford (2003) found that the amplitude of the error-related negativity was correlated with the BOLD response in the anterior cingulate gyrus. The present finding indicates that neurophysiological responses obtained from two independent modalities during the face/house discrimination task converge in the right fusiform gyrus.

A recent study by Horovitz, Rossion, Skudlarski, and Gore (2004) used a parametric approach to investigate the relationship between the N170 and fusiform activity in the same group of subjects. The authors found that correlations between the fMRI signal and the N170 amplitude for faces were significant in the bilateral fusiform gyrus. The results of the study by Horovitz et al. as well as those in the present study are consistent with the view that N170 originates in the fusiform gyrus. In addition, they suggested that other regions such as the superior temporal gyrus may also contribute to the N170 scalp potential. However, the authors used a block design for measuring the BOLD signal, while the evoked potential was obtained in a random design. The present study adopted the event-related design in both experiments and further supports the functional relationship between face-related activity measured by ERP and that measured by fMRI.

Although the present result of signal–amplitude correlation is predictable from those of previous studies (Horovitz et al., 2002; Horovitz et al., 2004; Mathalon et al., 2003; Opitz et al., 2002), the neurophysiological basis of the signal–latency correlation is unclear. We speculate that individual differences in cognitive challenge during face processing are the source of the variation in ERP and the BOLD signal across the subjects. It is known that pictures of faces presented upside-down are more difficult to recognize than those presented in an upright position (Valentine, 1988). Several ERP and fMRI experiments have shown that the inverted face elicits a longer N170 peak latency and greater fusiform activation than the upright face (Aguirre, Singh, & D'Esposito, 1999; Eimer, 2000b; Haxby et al., 1999; Kanwisher, Tong, & Nakayama, 1998; Rossion et al., 2000; Sagiv & Bentin, 2001), indicating a coupling of cognitive demand and physiological measurements within an individual subject. In a study using an oddball task, the latency factor in the fronto-central region negatively correlated with the subject's neuropsychological performance (Fjell & Walhovd, 2001), suggesting that longer peak latency is associated with cognitive slowing across the subjects. These studies may support the present hypothesis that demanding and/or slowing cognitive processes would cause prolonged latency and enhanced activation in the temporal regions. In the present study, there was no significant correlation between the BOLD signal and the RT, perhaps due to the fact that the BOLD signal in the fusiform gyrus represents neural activity involving perceptual processes rather than motor responses. The effect of across-subjects difference on other brain regions such as the amygdala, anterior cingulate, and primary motor areas may be closely related with the difference in RT.

There are several caveats in the present study. Due to a technical difficulty, we were unable to record fMRI and ERP simultaneously. Even though the same subjects and task paradigm were used in the experiment, the fact that data acquisition for the experiments was done separately might affect the behavioral and neurophysiological data. For example, the result showing that a higher error rate was observed in the ERP session than in the fMRI session has arisen from this experimental limitation. Although the order of the session was counterbalanced across the subjects, the temporal difference between the sessions may have affected the results. In addition, we should be cautious about the distribution of scalp electrical activity because the electrode positions in the present study were relatively sparse, and high-density EEG mapping is needed in future studies.

5. Conclusions

The present study investigated the relationship between hemodynamic and electrophysiological information obtained from fMRI and ERP using the same task paradigm and the same group of subjects. A significant correlation between the N170 parameters and the BOLD signal further supports the functional role of the inferior temporal lobe in face processing. The close relationship between the N170 latency and the hemodynamic responses may be explained by the evidence showing that the demanding cognitive process of facial stimuli as found in the face inversion effect evokes a greater activation and latency delay. In other words, the magnitude of BOLD activation in a particular region of the brain may be correlated with the time needed for effective processing within that region. The relationship between the data from two experimental modalities, as observed in this study suggests that integrative analysis of spatial and temporal information associated with a particular cognitive function may be achieved by a combined fMRI–ERP study.

Acknowledgments

We thank Dr. Kazuhiko Kakehi, Graduate School of Human Informatics, Nagoya University for valuable help for the experiment. The study was supported by JST/RISTEX and JSPS.

References

Aguirre, G. K., Singh, R., & D'Esposito, M. (1999). Stimulus inversion and the responses of face and object-sensitive cortical areas. *Neuroreport, 10*, 189–194.

Allison, T., Puce, A., Spencer, D. D., & McCarthy, G. (1999). Electrophysiological studies of human face perception. I: Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cerebral Cortex, 9*, 415–430.

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience, 8*, 551–565.

Caldara, R., Thut, G., Servoir, P., Michel, C. M., Bovet, P., & Renault, B. (2003). Face versus non-face object perception and the 'other-race' effect: A spatio-temporal event-related potential study. *Clinical Neurophysiology, 114*, 515–528.

Carmel, D., & Bentin, S. (2002). Domain specificity versus expertise: Factors influencing distinct processing of faces. *Cognition, 83*, 1–29.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134*, 9–21.

Eimer, M. (2000a). Event-related brain potentials distinguish processing stages involved in face perception and recognition. *Clinical Neurophysiology, 111*, 694–705.

Eimer, M. (2000b). Effects of face inversion on the structural encoding and recognition of faces. Evidence from event-related brain potentials. *Cognitive Brain Research, 10*, 145–158.

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature, 392*, 598–601.

Fjell, A. M., & Walhovd, K. B. (2001). P300 and neuropsychological tests as measures of aging: Scalp topography and cognitive changes. *Brain Topography, 14*, 25–40.

Gorno-Tempini, M. L., & Price, C. J. (2001). Identification of famous faces and buildings: A functional neuroimaging study of semantically unique items. *Brain, 124*, 2087–2097.

Hadjikhani, N., Joseph, R. M., Snyder, J., Chabris, C. F., Clark, J., Steele, S., et al. (2004). Activation of the fusiform gyrus when individuals with autism spectrum disorder view faces. *NeuroImage, 22*, 1141–1150.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4*, 223–233.

Haxby, J. V., Ungerleider, L. G., Clark, V. P., Schouten, J. L., Hoffman, E. A., & Martin, A. (1999). The effect of face inversion on activity in human neural systems for face and object perception. *Neuron, 22*, 189–199.

Henson, R. N., Goshen-Gottstein, Y., Ganel, T., Otten, L. J., Quayle, A., & Rugg, M. D. (2003). Electrophysiological and haemodynamic correlates of face perception, recognition and priming. *Cerebral Cortex, 13*, 793–805.

Horovitz, S. G., Rossion, B., Skudlarski, P., & Gore, J. (2004). Parametric design and correlational analyses help integrating fMRI and electrophysiological data during face processing. *Neuroimage, 22*, 1587–1595.

Horovitz, S. G., Skudlarski, P., & Gore, J. C. (2002). Correlations and dissociations between BOLD signal and P300 amplitude in an auditory oddball task: A parametric approach to combining fMRI and ERP. *Magnetic Resonance Imaging, 20*, 319–325.

Itier, R. J., & Taylor, M. J. (2002). Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: A repetition study using ERPs. *Neuroimage, 15*, 353–372.

Itier, R. J., & Taylor, M. J. (2004). Source analysis of the N170 to faces and objects. *Neuroreport, 15*, 1261–1265.

Jemel, B., George, N., Olivares, E., Fiori, N., & Renault, B. (1999). Event-related potentials to structural familiar face incongruity processing. *Psychophysiology, 36*, 437–452.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*, 4302–4311.

Kanwisher, N., Tong, F., & Nakayama, K. (1998). The effect of face inversion on the human fusiform face area. *Cognition, 68*, B1–B11.

Kutas, A., & Dale, A. (1997). Electrical and magnetic readings of mental functions. In M. Rugg (Ed.), *Cognitive neuroscience* (pp. 197–242). Hove: Psychology Press.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature, 412*, 150–157.

Maguire, E. A., Frith, C. D., & Cipolotti, L. (2001). Distinct neural systems for the encoding and recognition of topography and faces. *NeuroImage, 13*, 743–750.

Mathalon, D. H., Whitfield, S. L., & Ford, J. M. (2003). Anatomy of an error: ERP and fMRI. *Biological Psychology, 64*, 119–141.

McCarthy, G., Puce, A., Belger, A., & Allison, T. (1999). Electrophysiological studies of human face perception. II: Response properties of face-specific potentials generated in occipitotemporal cortex. *Cerebral Cortex, 9*, 431–444.

Menon, V., Ford, J. M., Lim, K. O., Glover, G. H., & Pfefferbaum, A. (1997). Combined event-related fMRI and EEG evidence for temporal-parietal cortex activation during target detection. *Neuroreport, 8*, 3029–3037.

Ogawa, S., Tank, D. W., Menon, R., Ellermann, J. M., Kim, S. G., Merkle, H., et al. (1992). Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping with magnetic resonance imaging. *Proceedings of the National Academy of Sciences of the USA, 89*, 5951–5955.

Opitz, B., Rinne, T., Mecklinger, A., von Cramon, D. Y., & Schroger, E. (2002). Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results. *NeuroImage, 15*, 167–174.

Puce, A., Allison, T., & McCarthy, G. (1999). Electrophysiological studies of human face perception. III: Effects of top-down processing on face-specific potentials. *Cerebral Cortex, 9*, 445–458.

Puce, A., Syngeniotis, A., Thompson, J. C., Abbott, D. F., Wheaton, K. J., & Castiello, U. (2003). The human temporal lobe integrates facial form and motion: Evidence from fMRI and ERP studies. *NeuroImage, 19*, 861–869.

Rolls, ET. (1992). Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical visual areas. *Philosophical Transaction of the Royal Society of London, Series B: Biological Sciences, 335*, 11–20.

Rorden, C., & Brett, M. (2000). Stereotaxic display of brain lesions. *Behavioral Neurology, 12*, 191–200.

Rossion, B., Gauthier, I., Tarr, M. J., Despland, P., Bruyer, R., Linotte, S., et al. (2000). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: An electrophysiological account of face-specific processes in the human brain. *Neuroreport, 11*, 69–74.

Sagiv, N., & Bentin, S. (2001). Structural encoding of human and schematic faces: Holistic and part-based processes. *Journal of Cognitive Neuroscience, 13*, 937–951.

Schweinberger, S. R., & Burton, A. M. (2003). Covert recognition and the neural system for face processing. *Cortex, 39*, 9–30.

Schweinberger, S. R., Pickering, E. C., Jentzsch, I., Burton, A. M., & Kaufmann, J. M. (2002). Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions. *Brain Research Cognitive Brain Research, 14*, 398–409.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. New York: Thieme.

Ullsperger, M., & von Cramon, D. Y. (2001). Subprocesses of performance monitoring: A dissociation of error processing and response competition revealed by event-related fMRI and ERPs. *NeuroImage, 14*, 1387–1401.

Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology, 79*, 471–491.