

## THE TIMING OF FACE SELECTIVITY AND ATTENTIONAL MODULATION IN VISUAL PROCESSING

Y. OKAZAKI,<sup>a,b,\*</sup> A. ABRAHAMYAN,<sup>c</sup> C. J. STEVENS<sup>c</sup>  
AND A. A. IOANNIDES<sup>a,b</sup>

<sup>a</sup>Department of Brain Science and Engineering, Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu 808-0196, Japan

<sup>b</sup>Laboratory for Human Brain Dynamics, Brain Science Institute, RIKEN, 2-1 Hirosawa, Wako-shi, Saitama, 351-0198, Japan

<sup>c</sup>MARCS Auditory Laboratories, University of Western Sydney, Locked Bag 1797, South Penrith NSW 1797, Australia

**Abstract**—Despite the complete imprint of a visual scene on the retina, the brain selects particular items for further processing. However, there is considerable debate about when and where the first attentional effects take hold in the cortex. We examined the timing of face specificity and attentional influences in the primary/secondary visual cortex (V1/V2) and in the fusiform gyrus (FG) in two experiments using magnetoencephalography (MEG). In experiment 1, using a passive viewing task, we identified three components in response to “Face,” “Hand,” and “Shoe” stimuli bilaterally in the FG:  $M_{FG100}$ ,  $M_{FG170}$ , and  $M_{FG200}$ —all showing a stronger preference for faces. The timing of these three activations of the FG is consistent with earlier studies claiming distinct stages of processing of visual stimuli in the first 300 ms. In experiment 2, subjects performed a gender-discrimination task on either faces or hands, drawing attention to only one of the two object categories. In addition to the previously identified three components in FG, here we found object-selective attentional enhancement first appearing in V1/V2 at around 170 ms, and then in FG at around 200 ms, i.e. concurrent with the third component. No attentional effects were evident on the first or second magnetoencephalography components. These findings may indicate that the visual input for an object is first encoded and matched to an attended “cue” object held in mind. When the attended and encoded objects match, a third stage involving attentive processing is enhanced. © 2008 IBRO. Published by Elsevier Ltd. All rights reserved.

**Key words:** fusiform gyrus, human, MEG, object-based attention, V1.

In our daily lives, we are bombarded by sensory signals, only some of which are potentially useful. We become

aware of a “capacity” problem when we fail to recognize particular information despite receiving all sensory inputs (Lavie and Tsai, 1994; Rees et al., 1997; VanRullen and Koch, 2003). It is generally agreed that a “biased attention selection competition” in early retinotopic visual areas enhances stimulus-induced neuronal responses that match “top-down” influences from the specialized extrastriate visual areas with features highlighted by the current attentional focus (Desimone and Duncan, 1995). For example, functional magnetic resonance imaging (fMRI) studies have found that selective attention to visual features or whole objects enhanced activity in the regions of extrastriate visual cortex that selectively process these same attributes (Corbetta et al., 1990; O’Craven et al., 1999). Types of attention are defined by their targets, what gets selected by attention. In the case of object (feature)-based attention the target is a whole object or parts of object, whereas in a case of spatial attention the target is a part of the visual field (VF).

There is, however, considerable debate on the details of the attention selection process for each type of attention, particularly on when and where the first attentional effects take hold in the cortex and how they develop in space and time thereafter. Evidence for attentional effects in striate cortex has been found in some event-related potential (ERP) studies co-localized with fMRI, particularly in the case of spatial attention (Di Russo et al., 2003; Martinez et al., 2001). In these studies, attentional effects were found first in extrastriate cortical regions between 80 and 130 ms, and then in primary visual cortex (V1) rather late, typically between 150 and 200 ms. Current evidence suggests that attentional modulation in striate cortex is the product of influences from higher visual areas (Martinez et al., 2001). However, the earliest attentional effect in V1 reported by our group (comparing passive and attentive viewing conditions in a GO/NOGO experiment) is considerably earlier than the ones found in those studies. This early attentional modulation in V1 occurred within 100 ms, but still some 40 ms after an earlier V1 activation which was not influenced by attention (Poghosyan et al., 2005). This result suggests that the extrastriate attentional effect reported by many authors after 100 ms may simply be a feedforward effect of an early attentional modulation in V1 within 100 ms. To clarify this issue we use a stimulus that excites a well-defined extrastriate area so that attentional effects can be tracked in this extrastriate area and contrasted with the attentional effect in V1. Specifically, our study uses face selectivity in the human fusiform gyrus (FG) as a tool for studying the timing of attentional effects in primary/secondary visual cortex (V1/V2) and FG.

\*Correspondence to: Y. Okazaki, Laboratory for Human Brain Dynamics, Brain Science Institute, RIKEN, 2-1 Hirosawa, Wako-shi, Saitama, 351-0198, Japan. Tel: +81-48-462-1111x7186; fax: +81-48-467-5973. E-mail address: yuka@brain.riken.jp (Y. Okazaki).

**Abbreviations:** ECG, electrocardiogram; EEG, electroencephalography; EOG, electrooculogram; ERF, event-related field; ERP, event-related potential; FG, fusiform gyrus; fMRI, functional magnetic resonance imaging; ICA, independent component analysis; MEG, magnetoencephalography; MFT, magnetic field tomography; MSR, magnetically shielded room; PET, positron-emission tomography; RAC, regional activation curve; ROI, region of interest; SAC, sensor activation curve; SPM, statistical parametric mapping; VF, visual field; V1/V2, primary/secondary visual cortex.

Face stimuli are very effective and their processing is an indispensable tool in social interactions. Questions regarding the spatiotemporal order of face-induced activity in extrastriate cortex have been examined by techniques which are complementary in their temporal and spatial resolving power, e.g. electroencephalography (EEG)/magnetoencephalography (MEG) and fMRI/positron-emission tomography (PET) studies. An area in the FG has often been shown to be preferentially excited by face stimuli and is referred to by many authors as the fusiform face area (FFA) (Hadjikhani and de Gelder, 2002; Kanwisher et al., 1997; Schwarzlose et al., 2005). In this manuscript, we use the abbreviation FG for this area to acknowledge the fact that this may not be the only area showing preferential activity for faces (Ishai et al., 1999). Using EEG and MEG, face selectivity is evident between 100 and 200 ms, most prominently around 170 ms, and hence the corresponding components are labeled N170 and M170 for EEG and MEG respectively. A few studies have reported early face-specific responses, but only at the level of the EEG or MEG sensor signal (Itier and Taylor, 2004b; Liu et al., 2002).

In many studies, face-selective ERP/event-related field (ERF) are explicitly or implicitly associated with face-selective fMRI hemodynamic responses. However, making such inferences is problematic, because the well-localized fMRI response contains contributions from different times and it is very likely dominated by the late responses that best match the slow hemodynamic response. On the other hand the (average) ERP/ERF signal is very likely dominated by the early responses that are well time-locked to the stimulus onset, with overlapping contributions even at early latencies from different areas. Under such uncertainty, separate electrophysiological and hemodynamic experiments allow us to draw plausible inferences but no definitive conclusions. A better understanding of face-induced brain activity can be obtained using methods that can provide accurate spatial and temporal information simultaneously, such as intracranial recordings or source estimation from MEG data.

Intracranial recordings have identified early and late face responses. For example, Seeck et al. (1997) localized early source activity (50–90 ms) elicited by faces in the mid- and infero-temporal cortex. Also, a face-specific N200 was intracranially recorded in a ventral region including the FG and a lateral region centered in the middle temporal gyri including the inferior temporal gyri (Allison et al., 1999). MEG studies on M170 with limited sensor coverage claimed face specific-responses in the FG for the M170 component (Liu et al., 1999; Watanabe et al., 1999), at latencies intermediate to the early and late components identified with intracranial recordings. The inconsistency in timing between extracranially recorded M170 and the intracranially recorded signal has not been clarified yet. MEG studies have also reported early components induced by face stimuli, some in early visual areas (Itier et al., 2006; Linkenkaer-Hansen et al., 1998; Watanabe et al., 1999), while others in the FG, as early as 70–80 ms after stimulus onset (Liu and Ioannides, 2006). Unfortunately, the last study did not use non-face control stimuli, so it

could only demonstrate that face stimuli can elicit early activity in FG, but could not help determine whether this component was face specific. What is therefore needed is accurate tomographic localization and precise timing in the measurement of face and non-face stimuli within the same experiment.

The present study investigates how selective attention to faces modulates face-selective processing. We firstly identified face-specific activity using tomographic source analysis of MEG signals to map the response to face and non-face stimuli in both space and time. We addressed three key questions. 1) Where does face specificity appear, and specifically, which of the FG components shows face specificity? 2) How early and where does attentional modulation appear in the visual hierarchy? 3) When does selective attention to a preferred stimulus or non-preferred stimulus, modulate the FG response? We selected hands and shoes as non-face stimuli because they are meaningful, non-biological, approximately symmetrical about their vertical axis, three-dimensional, and dynamic. So despite not being faces, they share some of the gross features of faces. We performed two experiments using as stimuli, images of human faces, hands, and shoes. In the first experiment, all three stimulus types were passively presented separately to the central and peripheral VF. For each VF location, we compared the activation patterns obtained with Faces to the other stimulus categories, Hands and Shoes, to test the face-selectivity of each component elicited by our stimulus sets in V1/V2 and FG. In experiment 2, we used only the face and hand stimuli of experiment 1 and presented them only to the peripheral VF while adopting a continuous target detection task to minimize recording time. Since subjects counted the number of gender-specific targets from one category within a block of trials, their attention was drawn to the entire object (face or hand) to identify their gender. We expected that subjects would attend more to all objects of the target category even if they were not actual targets, than to the objects of the other category. We compared the activations elicited by same-category stimuli, when they were used as targets versus non-targets, to identify the latency of the first attentional effect in V1/V2 and FG. We hypothesized that face-selective responses would be differentially modulated depending on whether attention was drawn to faces or hands.

## EXPERIMENTAL PROCEDURES

### Subjects

Seven, healthy right-handed Caucasian male subjects participated (mean age,  $30.0 \pm 5.0$  years). All subjects were in good health with no history of psychiatric or neurological disease, and gave informed, written consent. The MEG protocol had been approved by the Research Ethics Committee of RIKEN. Fifty-seven (34 women and 23 men, mean age  $26.0 \pm 12.0$ ) subjects from the University of Western Sydney participated in a behavioral experiment corresponding to the task of experiment 2. The study was approved by the Human Research Ethics Committee at the University of Western Sydney.

## Visual task

**Experiment 1: face specificity.** As stimuli, we used gray-scale images of six faces and six hands. The faces and hands consisted of three males and three females, each with positive, neutral, or negative expressions. All faces were taken from a standard set of pictures prepared by Ekman and Friesen (1976). A set of hand stimuli was prepared and tested in a separate behavioral study (Abrahamyan et al., 2005). As non-biological control stimuli, two shoes, one in male and the other in female form/style were introduced in addition to the face and hand stimuli, based on their similarity to faces and hands in their ability to reflect gender, their rough symmetry about the vertical axis, and their potential for motion. Shoe images were provided courtesy of the website <http://www.shoes.com>. Stimulus size and mean luminance values were adjusted to be as similar as possible.

Each stimulus was presented for 300 ms in one of five locations: either in the center or at 10.7° eccentricity from fixation across the diagonal, (i.e. upper left, upper right, lower left, and lower right quadrants of the VF). In each run, three of the five locations were used, alternating the choices to equally cover all five locations in the different runs. As there were five runs in this experiment, each location was chosen three times. Each stimulus was presented five times in each of three locations in one run. In central presentations, faces, hands, and shoes were presented at sizes of 5.5×4.1°, 4.5×3.6°, and 4.8×3.5° respectively. In the periphery, face, hand and shoe images were 8.2×6.1°, 6.7×5.3° and 7.1×5.2° in size. The larger stimulus size in the periphery was chosen to partially compensate for cortical magnification differences between centrally and peripherally presented stimuli. The subject's task in experiment 1 was to fixate on the central cross and to respond to the subtle change in its color with a quick button press. The task, which occurred a few times in each run, was introduced in order to maintain subject concentration from beginning to end (see Fig. 1a, b).

**Experiment 2: effect of attention.** The same face and hand stimuli as in experiment 1 were used. We had already confirmed the efficacy of the stimuli in a task with 57 subjects by Abrahamyan et al. (2005). A behavioral paradigm was suitably adopted for transition into the MEG setting. The experiment consisted of eight runs, each containing 12 blocks of 13 stimuli each. A block

started with a cue (e.g. the words, “male face”) defining the task. In each run, 18 non-target face and hand stimuli and a variable number of target stimuli were randomly presented for 300 ms in one of the quadrants. The task was to count the target and to report the number of targets, i.e. 0, 1 or 2, at the end of each block by pressing response buttons (see Fig. 1c). Targets were used to establish the task, and they occurred on only about 8% of a run. These trials were not used in the analysis of evoked fields to rule out any effects of counting on the target. We contrasted non-target faces or hands when attended ( $F_{stim}F_{att}$  and  $H_{stim}H_{att}$ ), with the same unattended ( $F_{stim}H_{att}$  and  $H_{stim}F_{att}$ ).

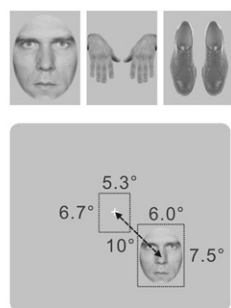
## MEG setup

Magnetic fields were measured with the CTF 151-channel MEG systems (Omega 151, CTF Systems Inc., Vancouver, BC, Canada). The subject was seated in a chair in front of a screen, and performed the task in the otherwise pitch darkness of the magnetically shielded room (MSR). Subject responses were recorded using an optical sensor (FU-38V, Keyence Co., Osaka, Japan). Stimuli were presented on the screen by back-projection with a DLP projector with a 96 Hz refresh rate (HL8000Dsx+, NEC Viewtechnology Ltd., Tokyo, Japan) located outside the MSR.

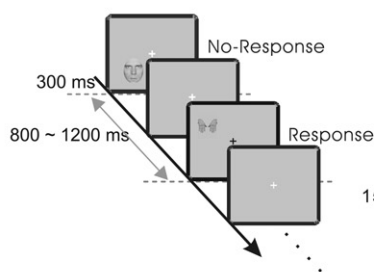
Using conventional projectors, the observed luminance may differ more than fourfold between distant screen locations. Uniform luminance is critical in studying VF asymmetries of brain function. Therefore, in order to correct for this non-uniformity, we introduced two procedures. First, we used neutral density filters (MC Center-ND Filter, Kenko Co. Ltd., Tokyo Japan) in front of the projector to roughly reduce the non-uniformity and to adjust brightness. Second, we fine-tuned the output signal of the projector according to the observed luminance through the filter at grid points on the screen. As a result, luminance non-uniformity of the screen was improved to within 10% with an obtained on-screen luminance of about 45 cd/m<sup>2</sup>. Stimulus delivery was controlled by a program from Neurobehavioral Systems (Presentation, Neurobehavioral Systems, Inc., Albany, CA, USA). The exact onset time of each stimulus was determined by luminance detection with a photodiode on the screen.

In preparing the subjects, three head localization coils were attached to the nasion and the left and right pre-auricular points.

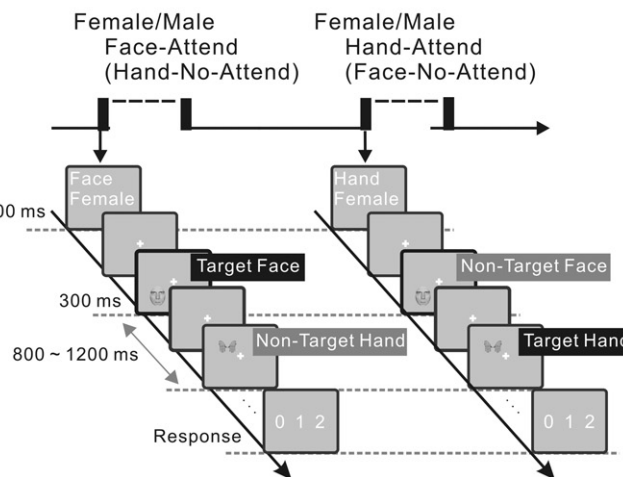
### a: stimuli and location



### b: Experiment 1



### c: Experiment 2



**Fig. 1.** In each run of experiment 1, Face, Hand and Shoe stimuli were presented randomly at the center and quadrants (Upper, Lower Left, and Right VFs). In each run of experiment 2, Face and Hand stimuli were presented in random order to quadrants of the VFs. In experiment 1, while subjects fixated the center cross, they identified a subtle change in the color of the fixation cross and responded as quickly as possible. In experiment 2, in alternating blocks of trials, subjects either discriminated target face gender or target hand gender. They counted the number of targets and reported this number (e.g., 0, 1 or 2) at the end of the block.



Electrodes were also attached to record the electrooculogram (EOG) and electrocardiogram (ECG). The three head localization coils were activated and the resulting MEG signal was recorded at the beginning and end of each experimental recording (head-localization data). Each set of head-localization data was used to calculate the subject's head position relative to the sensors. The electrical activity generated by eye movements and cardiac systole was monitored by simultaneously recording EOG and ECG. The EOG and ECG signals were later used to eliminate unwanted strong magnetic fields generated by eye movement and heart activity (see MEG signal processing). The signal from all channels, including subjects' responses, the photodiode, and their trigger from the stimulus PC, was collected through a 200 Hz low-pass filter and digitized at a sampling-rate of 625 Hz by the native acquisition program of our MEG hardware.

### Co-registration of MEG and MRI

Anatomical brain images of each subject were collected with a T1-weighted MRI (1.5-T Siemens MRI system, voxel size of  $1 \times 1 \times 1$  mm<sup>3</sup>). In order to localize the source of the signals, MEG results were registered with MRI images using the following procedure. Before the MEG experiment, five coils (the three main head-localization coils described above, and two additional coils added for precision), were attached to the subject's head. The positions of the coils with respect to the subject's head shape were digitized using a 3D digitizer (FASTRAK, Polhemus, Colchester, VT, USA) and a 3D camera system (VIVID 700, Konica Minolta Holdings, Inc., Tokyo, Japan). The positions of the five coils were calculated relative to the subject's head coordinate system by activating the five coils in a separate MEG recording before the main experiments. The digitized head shape with head coordinates was fitted to the MRI contour, allowing for small deformations, so as to obtain a transformation matrix from the head shape to the MRI coordinate system (Hironaga and Ioannides, 2002).

### MEG signal processing (noise removal, filter, independent component analysis (ICA))

The collected MEG data were segmented into single trials from –300–700 ms relative to the stimulus onset marked by the photodiode. Noisy channels were removed. Trials with eye movements or blinks exceeding 50  $\mu$ V EOG signal change during the stimulus presentation period were discarded. Those data were filtered with a bandwidth of 3–200 Hz and with notches at 50 Hz and its harmonics to eliminate power-line noise. The concatenated single-trial signal for each run was analyzed with ICA (Jahn et al., 1999). Remaining artifacts related to vital signals were identified by strong ICA components correlated with either EOG or ECG and were removed. After cleaning by ICA, the data were averaged for each condition (i.e. Faces, Hands, Shoes, Attend and No-Attend), and VF location.

### Magnetic field tomography (MFT)

The localization of neuronal activity was computed using MFT (Ioannides et al., 1990; Taylor et al., 1999). We applied MFT to the averaged data of each condition after cleaning by ICA in individual subjects. An independent MFT computation was done for each time slice of averaged data. For each time slice, separate MFT calculations were made for four source spaces ( $17 \times 17 \times 11$  grid points in size), defined to be partially overlapping and completely covering the left, right, top, and back part of the brain for each individual subject. The primary current density was computed separately using the 90 channels closest to each of the four source spaces. The results from the four separate MFT computations were finally combined to produce a three-dimensional distribution of primary current density,  $\mathbf{J}(\mathbf{r}, t)$  in an array of ( $17 \times 17 \times 17$ ) grid points, covering the entire brain. The combina-

tion of the reconstructions from different source spaces used weights defined from the sensitivity profiles (lead fields) of the sensors.

### Post-MFT statistical analysis

**SPM analysis.** For each subject, voxel-by-voxel statistical parametric mapping (SPM) analysis was performed in each time-slice (1.6 ms). Bonferroni-correction was used to account for type I errors due to multiple voxel-by-voxel comparisons. SPMs were generated by comparing the sample distribution for the modulus of the MFT solution by two methods: a condition vs. condition comparison called the *active test*, and the pre- and post-stimulus period comparison, called the *baseline test*. In the active test, the distributions of two conditions (e.g. Faces and Hands) were generated from samples in 19.2 ms windows taken from all three runs with stimuli in the same part of the VF. After comparing these distributions, the centers of the windows were moved by 1.6 ms to generate new distributions. The baseline test compared the distribution of samples taken from 1.6 ms post-stimulus latency windows of the three runs with the distribution of random samples from the baseline. The baseline distribution (taken from –250 ms to –50 ms), consisted of samples in 1.6 ms time windows separated by at least 6.4 ms. After comparing these distributions, the center of the post-stimulus window was shifted by 1.6 ms for the next comparison.

**Region of interest (ROI) definition.** We used anatomical criteria and the baseline test of SPM results to define ROIs with a radius of 10 mm for V1/V2 and for the FG in each subject. Having fixed the ROI center (see below), we used circular statistics (Fisher, 1993; Ioannides et al., 2005) to define the “main direction,” i.e. the dominant direction of the MFT current density inside the given ROI.

We used the runs with stimuli in one quadrant of the periphery to define the retinotopically corresponding V1/V2 ROI, in the diagonally opposite part of the calcarine. For example, the left dorsal V1/V2 was determined from the MFT solutions for stimuli in the lower right VF. The center and main direction of the V1/V2 ROIs were determined from the MFT solutions in the range of 40–100 ms. We used the same locations to define four ROIs for central presentation, but determined the direction independently using the MFT solutions for stimuli in the center of the VF.

We defined two ROIs for the FG—right and left—separately for the center and each right and left side of the peripheral location (always using the contralateral stimulus). We first projected the Talairach coordinates for the common FG activations across subjects back to each individual MRI. ROIs were confined to areas close to the back-transformed location and anatomical landmarks: the collateral sulcus and the temporal occipital sulcus. For each subject, we identified the focal beginning of significant SPMs (after 100 ms) satisfying the aforementioned criteria, as the FG ROI center. The direction of the current density was defined from the MFT solutions elicited by face stimuli for the ROI-defining runs in the 120 ms to 180 ms range.

### Regional activity

**Regional activation curve (RAC).** After ROI definition, an activation curve was calculated for each stimulus condition at each time-slice (1.6 ms) by projection of their current density vector onto the main direction induced by the face stimulus. The RAC was generated for all conditions in each run of each subject for the six ROIs (two for the FG and four for the V1/V2).

**Analysis of RAC.** Momentary amplitudes of the RAC were analyzed using ANOVA, following the same conceptual steps as for the SPM analysis—a 4.8 ms running window stepped every 1.6 ms. In the RAC active test of experiment 1, an ANOVA was performed with Stimulus type (Faces, Hands, Shoes), Hemi-

sphere (left, right), and VF (upper, lower) as fixed factors, and Subject (seven subjects) as a random factor, in order to identify any statistically significant main effect of Stimulus type or an interaction. For experiment 2, Attention (Attend, No-Attend) was added as a fixed factor to the case of experiment 1. In order to statistically determine whether our stimuli activated FG and V1/V2, a baseline test (RAC baseline test) was created by adding Period (pre-stimulus, post-stimulus) as an additional fixed factor to those of experiment 1.

## RESULTS

### Behavioral experiment

Fifty-seven subjects participated in behavioral experiments designed for stimulus selection and validation. Data from three participants (two women and a man) were excluded due to inadvertent events that impeded successful data collection. The dependent measure, gender identification accuracy, was converted into  $d'$  values separately for face and hand stimuli (Tanner and Swets, 1954). Four participants (two women and two men) were excluded from further analysis because their average  $d'$  value was less than zero, which is an indicator of poorer than chance performance. The gender identification accuracy of faces (mean  $d'=1.68$ ) was significantly better compared with hands (mean  $d'=1.49$ ;  $P<0.05$ ; two-tailed paired  $t$ -test). However, the difference in the gender identification accuracy in the behavioral experiment should not alter the effect of attentional modulation by attended stimuli in the present neuroimaging experiment, because we did not directly compare responses to face and hand stimuli. Instead, we compared the responses to attended face/hand stimuli with the responses to unattended face/hand stimuli.

### Experiment 1

**Signal on the MEG sensor.** The MEG signal waveforms recorded in our experiment with face stimuli show the previously described face-selective M170 response. Fig. 2A shows the butterfly plot for the average MEG signal from 30 trials for central presentation of face stimuli for a typical subject. The M170 component can be clearly seen with a peak at 150 ms. Fig. 2B shows a magnetic field map on a scalp at the face-sensitive M170 latency (left topography) and a computer-generated field map computed with a dipolar source placed in the FG (right topography). The similarity in the two maps is evident for the topography generated by the occipito-temporal sensors. This result supports the often-made claim that the MEG signal around the M170 latencies of occipito-temporal sensors reflects activities in the FG. However, closer examination of the brain activity around M170 reveals statistically significant activity in other areas on the ventro-lateral occipital cortex, in addition to the FG (Fig. 2D; see Activation in the FG for detail). Fig. 2C shows sensor activation curves (SAC), constructed from a linear combination of the MEG sensors defining the M170 dipolar pattern (see below for definition). This linear combination provides a quantification of the overall pattern for face and hand responses in one curve. Each curve is obtained by adding the five most positive

sensors ( $p_1$ – $p_5$ ) and the five most negative ( $n_1$ – $n_5$ ) sensors of the clear dipolar pattern of the M170 topography:

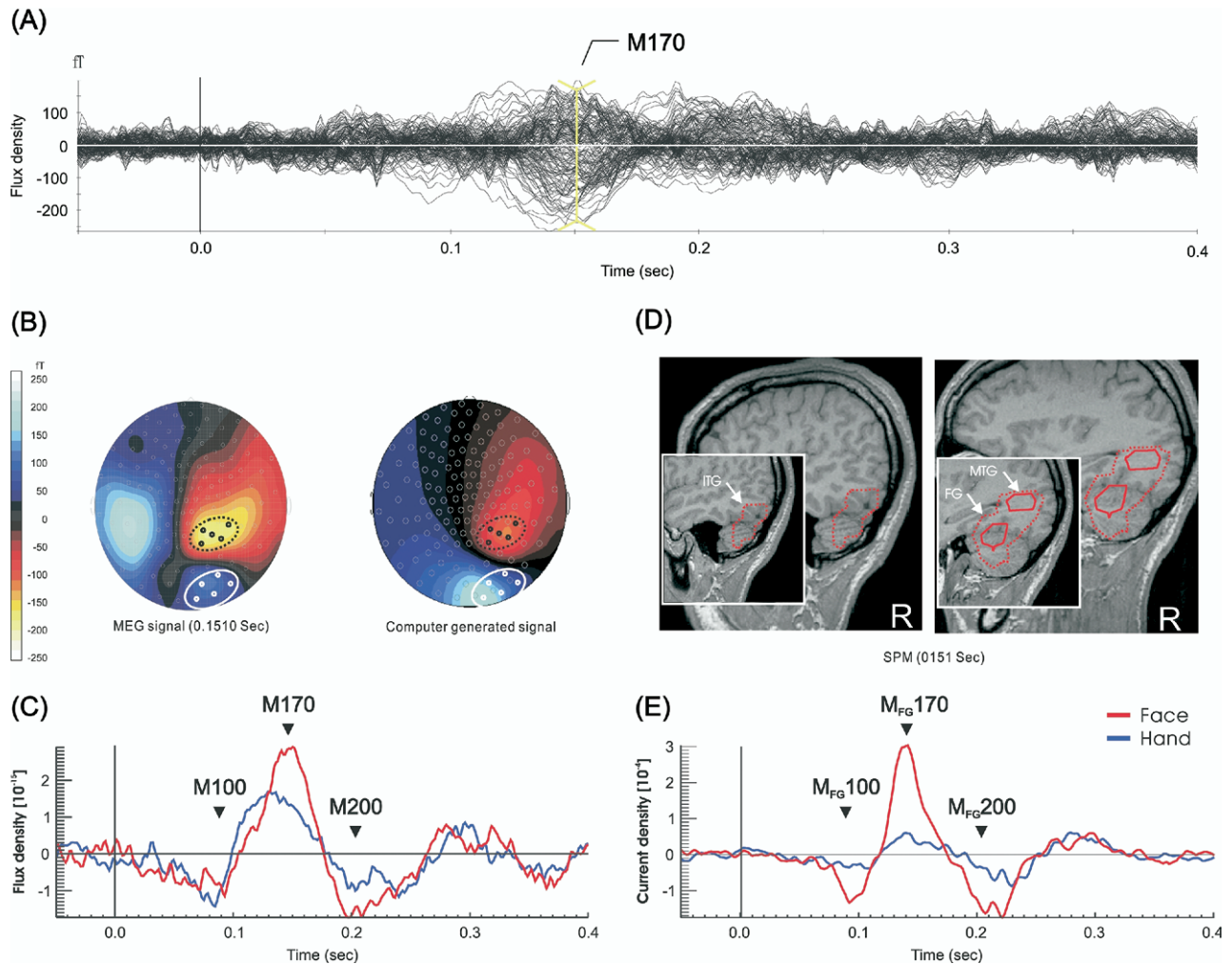
$$SAC = \frac{1}{5} \left[ \sum_{i=1}^5 s_{p_i}(t) - \sum_{i=1}^5 s_{n_i}(t) \right] \quad (1)$$

A direct comparison of SAC for face and hand stimuli shows clear difference for the M170 and the M200 components, but critically no such difference is seen for the M100 component. In contrast, all three components of the FG RAC showed higher magnitude of current density for face (Fig. 2E; see A Face-Selective Response for further results). The difference between the SAC and RAC measures is due to the dependence of the raw MEG signal and the SAC on the activity of many sources some of which are not face sensitive.

**Activation in the FG (baseline test: pre-stimulus vs. post-stimulus).** Highly significant SPM foci were identified in V1/V2 within 100 ms, with this early focal activity then extending anteriorly to FG. Fig. 2D shows widely distributed face-activated areas, which were significantly larger than their baselines at the latency corresponding to M170. Increased response for faces was found in occipito-temporal regions FG, middle temporal gyrus (MTG), and the inferior temporal gyrus which corresponds to the inferior occipital gyral area known as the occipital face area (Gauthier et al., 2000). In this paper, we will focus on the two most prominent areas; V1/V2 which is the main gate of visual input to the cortex and the posterior FG which has consistently shown face-preferential activity (Kanwisher et al., 1997) and selective attentional modulation by face stimulus (O'Craven et al., 1997). In our study, we also found that the FG was the area most consistently excited by face stimuli: The FG was the only area surviving at high thresholds ( $P<0.001$ ) when SPM analysis was performed separately for each run with a short sample window size (1.6 ms). In these SPM analyses the FG preference for faces was found in each run at latencies 150 ms.

The ROIs for individual subjects were defined from the functional data of each subject, guided by the SPM results and anatomical landmarks in the individual subject's MRI. Fig. 3a shows an example map for the significant change of activity ( $P<0.000005$ ) in FG for one subject. For this subject, the FG ROIs for the center and periphery were indistinguishable. The ROI centers of each subject were displayed with active test results (see Fig. 3b, blue dots) after transformation to the common Talairach space and back-transformation to the coordinates of the displayed MRI (active test results will be described later). The Talairach coordinates of V1/V2 and FG ROIs for center and peripheral locations are listed in Table 1. Statistical comparison showed no significant differences between center and peripheral FG ROIs.

The significance of FG activations was assessed by an RAC baseline test. Fig. 4a, b shows the FG activation time courses for the grand average across subjects in the central and peripheral presentations. Significant FG activation periods were at 71.0–121.0 ms ( $F(1, 6)=47.6$ ,  $P<0.05$ ),



**Fig. 2.** MEG signals and activity in brain sources for a typical subject. (A) The butterfly plot of the averaged MEG raw signal exhibits the large ERF for face stimuli corresponding to the M170 component. The signal is aligned on stimulus onset and 30 trials with a central presentation of face stimuli are averaged. (B) The left part shows the magnetic field topography on a flat projection of the surface containing the sensors; the location of each MEG sensor is indicated by a yellow marker. The right topography is derived from a computer-generated signal, with a dipole source placed in the same location and with the same direction as the FG source identified by MFT. (C) The SAC shows that the magnitude of flux density for face stimuli is notably higher than for hand stimuli at M170 and M200, but not at M100. SAC is derived from linear combinations of the existing MEG channels, specifically, the relative differences between the average of the positive and the average of the negative MEG signals of the dipole pattern on the marked part of the topography. (D) The SPM with baseline test shows a brain source in the visual ventral pathway on the right hemisphere corresponding to the magnetic field distribution in B, left. Circumscribed areas show significant changes of activity ( $P < 0.000001$ ). (E) The RAC with FG averaged across runs shows a higher value of current density for face stimuli than for hand stimuli for all components, including  $M_{FG100}$ .

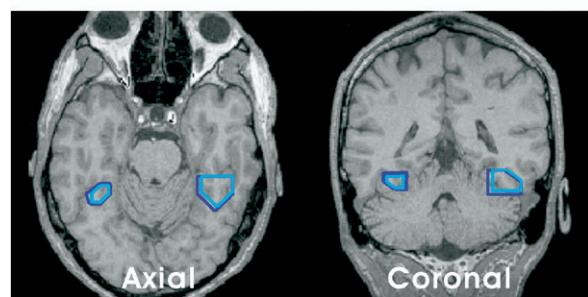
135.0–186.0 ms ( $F(1, 6) = 22.4$ ,  $P < 0.05$ ), and 204.0–263.0 ms ( $F(1, 6) = 20.2$ ,  $P < 0.05$ ) for central presentation. For the peripheral presentation, significant contralateral FG activation periods were at 65.0–113.0 ms ( $F(1, 6) = 30.2$ ,  $P < 0.05$ ), 127.0–177.0 ms ( $F(1, 6) = 45.3$ ,  $P < 0.05$ ), and 207.0–263.0 ms ( $F(1, 6) = 35.3$ ,  $P < 0.05$ ). In Fig. 4, significant activation periods are represented by the shaded bands.

*A face-selective response (active test: face vs. the other objects).* SPM results for the active test identified significantly larger activity in the FG for Faces than Hands. The contours in Fig. 3b show the common face-selective FG activations across subjects ( $P < 0.005$ ) around 140 ms, back-transformed and displayed with the MRI of one sub-

ject. On the axial slices of Fig. 3b the small squares show the FG SPM result for baseline test for individual subject (after transformation to the anatomical space of the MRI used for the display). As can be seen, the SPMs from active (contours) and baseline (squares) tests produced consistent loci.

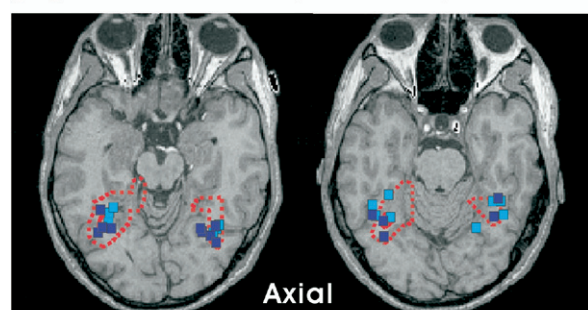
Three peaks were obtained, at  $<100$  ms,  $<200$  ms, and  $<300$  ms, by comparing the pre- and post-stimulus period of the RAC in the FG ROI. We performed a post hoc test for the most significant main effects of stimuli at these peaks using Tukey's method. No main effect of Hemisphere or Stimulus type  $\times$  Hemisphere interaction was found. The bar graph in Fig. 4 shows the amplitude of current density at the latencies that showed the most sig-



**(a) Baseline Test** (single subject)

Stimulus: Face

□ Statistically significant Activity

**(b) Active Test** (7 subjects)Face > Hand ( $p < 0.005$ )

5/7 6/7 ROI (from Baseline Test)

**Fig. 3.** (a) Baseline test (single subjects). Circumscribed areas show a significant change of activity induced by a face stimulus presented to the side contralateral to the right and left FG (light blue), and at the center (blue). (b) Active test (all seven subjects). Circumscribed areas indicate regions of higher activation ( $P < 0.005$ ) for faces than hands common to five of seven subjects (dash), and common to six of seven subjects (solid). Blue and light blue small squares indicate the ROI center identified by the baseline test in central and peripheral presentations for all subjects.

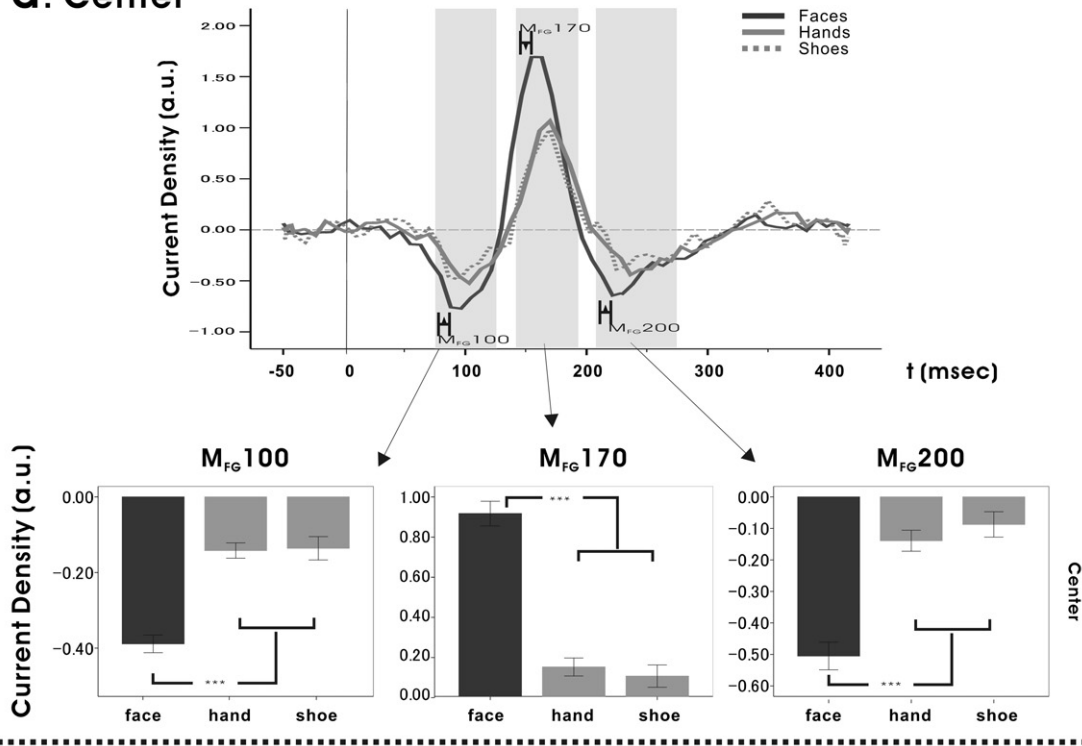
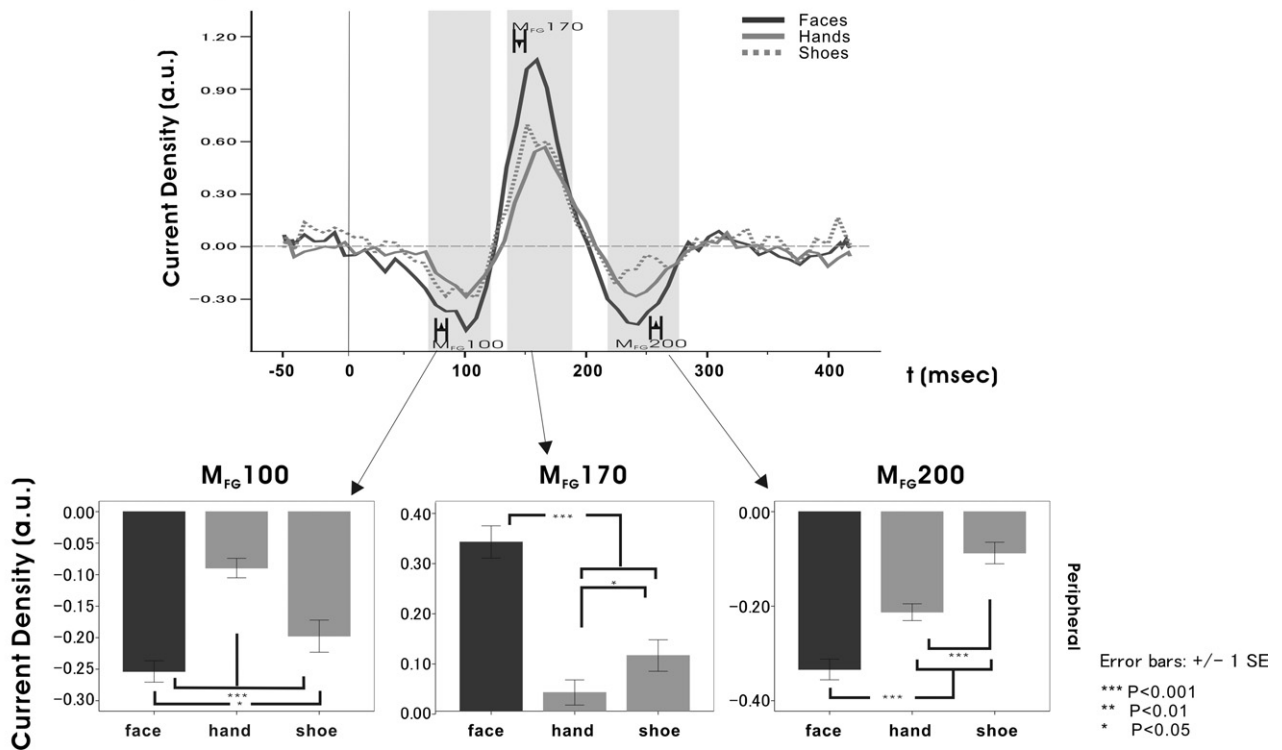
nificant differences between objects for the three periods in central (a) and peripheral (b) presentations. The bar graph demonstrates the expected face selectivity at  $M_{FG}170$ . The amplitude of face stimuli became significantly stronger than the other objects at 135.0 ms with a main effect of Stimulus type according to ANOVA ( $F(2, 12) = 8.76$ ,  $P < 0.005$ ), and showed an amplitude peak at 153.0 ms.

Similarly, during peripheral presentation, contralateral FG showed a stronger response to faces at 126.0 ms ( $F(2, 12) = 7.11$ ,  $P < 0.01$ ), with an amplitude peak at 150.0 ms. Other face-selective responses were observed after  $M_{FG}170$ . Subsequent significant differences were found at 207.0 ms for the central presentations ( $F(2, 12) = 4.97$ ,  $P < 0.05$ ) and 246.0 ms for peripheral presentations ( $F(2, 12) = 13.81$ ,  $P < 0.001$ ). The results show responses to objects (not just faces) within 100 ms, which agrees with findings by Liu and Ioannides (2006). Face preference on early components around 100 ms has not been examined despite the implications of some studies showing a preference to faces in the MEG signal (Liu et al., 2002). The results of this study showed that the  $M_{FG}100$  was stronger for face stimuli compared with other objects. The amplitude for face stimuli was significantly higher at the latency showing a main effect of Stimulus Type according to ANOVA, peaking at 73.0 ms (range, 63.0–81.0 ms), and 65.0 ms (range, 60.0–81.0 ms), for central ( $F(2, 12) = 13.15$ ,  $P < 0.001$ ) and peripheral presentations ( $F(2, 12) = 11.14$ ,  $P < 0.05$ ), respectively.

However, the difference in the FG  $M_{FG}100$  component between objects was found as early as the first large component occurring in striate cortex. This could be the result of earlier levels of processing not specific to faces but rather to the physical feature differences of the stimuli (e.g. luminance, size, contrast, etc.). We used the RAC active test to compare the activations elicited by the different images within 100 ms in V1/V2 and in the FG. No interactions between Stimulus type, Hemisphere and VF were found, so the RACs were averaged. Fig. 5a, b shows the RACs in grand averaged left/right–dorsal/ventral V1/V2 for peripheral and central presentations, respectively. There was no preference for faces in the  $M_{V1}100$  component at any location. These activities in V1/V2 and FG very likely reflect feedforward input via the ventral pathway (Mishkin et al., 1983), because the onset of  $M_{FG}100$  was slightly later than  $M_{V1}100$  and the  $M_{FG}100$  is too slow for subcortical inputs that are not V1-mediated. Given the significant difference between faces and the other object stimuli in  $M_{FG}100$  and the absence of such preference in V1 we conclude that the most parsimonious explanation for these findings is that the separation of objects is established early in higher levels in the visual hierarchy and that it is not merely the consequence of physical feature differences of the stimulus. Although there

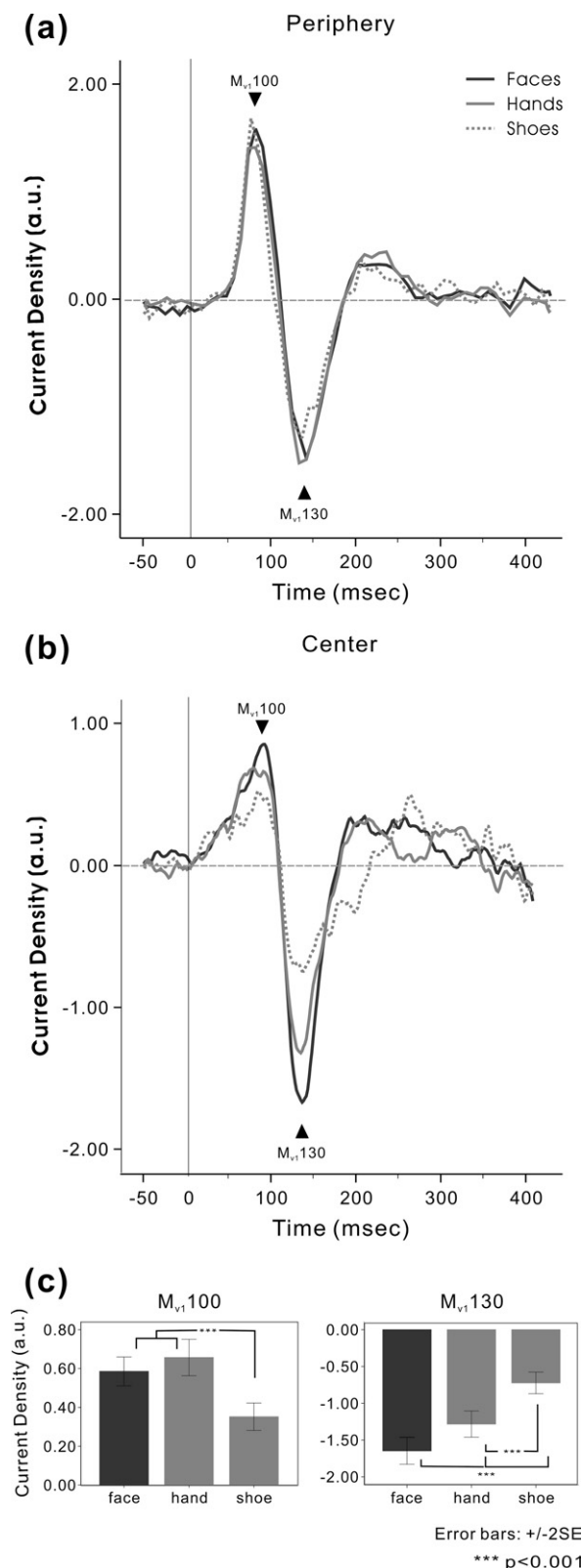
**Table 1.** Talairach coordinates x, y, z (mean  $\pm$  S.D.) in mm for right and left fusiform gyrus ROIs defined by the SPM baseline test in central and peripheral presentations, and left/right–dorsal/ventral V1/V2 ROIs defined by MFT solutions for peripheral presentations

Location	Region	ROI	x	y	z
Periphery	V1/V2	Right-ventral V1/V2	10 $\pm$ 4	–81 $\pm$ 6	–9 $\pm$ 6
		Right-dorsal V1/V2	12 $\pm$ 2	–81 $\pm$ 6	–6 $\pm$ 4
		Left-ventral V1/V2	–10 $\pm$ 2	–86 $\pm$ 4	–11 $\pm$ 7
		Left-dorsal V1/V2	–10 $\pm$ 3	–90 $\pm$ 3	–6 $\pm$ 6
	FG	Right FG	32 $\pm$ 5	–50 $\pm$ 7	–14 $\pm$ 5
		Left FG	–34 $\pm$ 4	–50 $\pm$ 4	–11 $\pm$ 7
Center		Right FG	31 $\pm$ 5	–56 $\pm$ 8	–11 $\pm$ 5
		Left FG	–35 $\pm$ 4	–56 $\pm$ 7	–12 $\pm$ 4

**a: Center****b: Periphery**

**Fig. 4.** The RACs are averaged across subjects in the left and right FG for central (a) and peripheral (b) VF presentation. Black, faces; gray, non-faces. The gray, shaded areas indicate significance higher than baseline (–250 ms to –50 ms). Three object-specific components were obtained:  $M_{FG100}$ ,  $M_{FG170}$ , and  $M_{FG200}$  in both the center and periphery. Bar graphs are amplitudes of the grand-averaged current density corresponding to  $M_{FG100}$ ,  $M_{FG170}$ , and  $M_{FG200}$  responses to faces, hands, and shoes for both the center and periphery. In all responses, face stimuli elicited higher activation than non-face stimuli at both central and peripheral presentation. Error bars, S.E.M.; a.u., arbitrary units.





**Fig. 5.** The activation curve averaged across subjects in the right/left-dorsal/ventral V1/V2, induced by presentation to quadrants diagonally opposite to localized regions surrounding the calcarine sulcus (a) and presentation to center (b). V1/V2 shows no preferential response for

was no preference for faces in central presentation, biological stimuli such as faces and hands are significantly higher than non-biological shoe stimuli ( $P < 0.001$ ), as shown in Fig. 5c. Additionally, subsequent components ( $M_{v1}130$ ) to  $M_{v1}100$  showed a significant difference between objects and face preference ( $P < 0.001$ ).

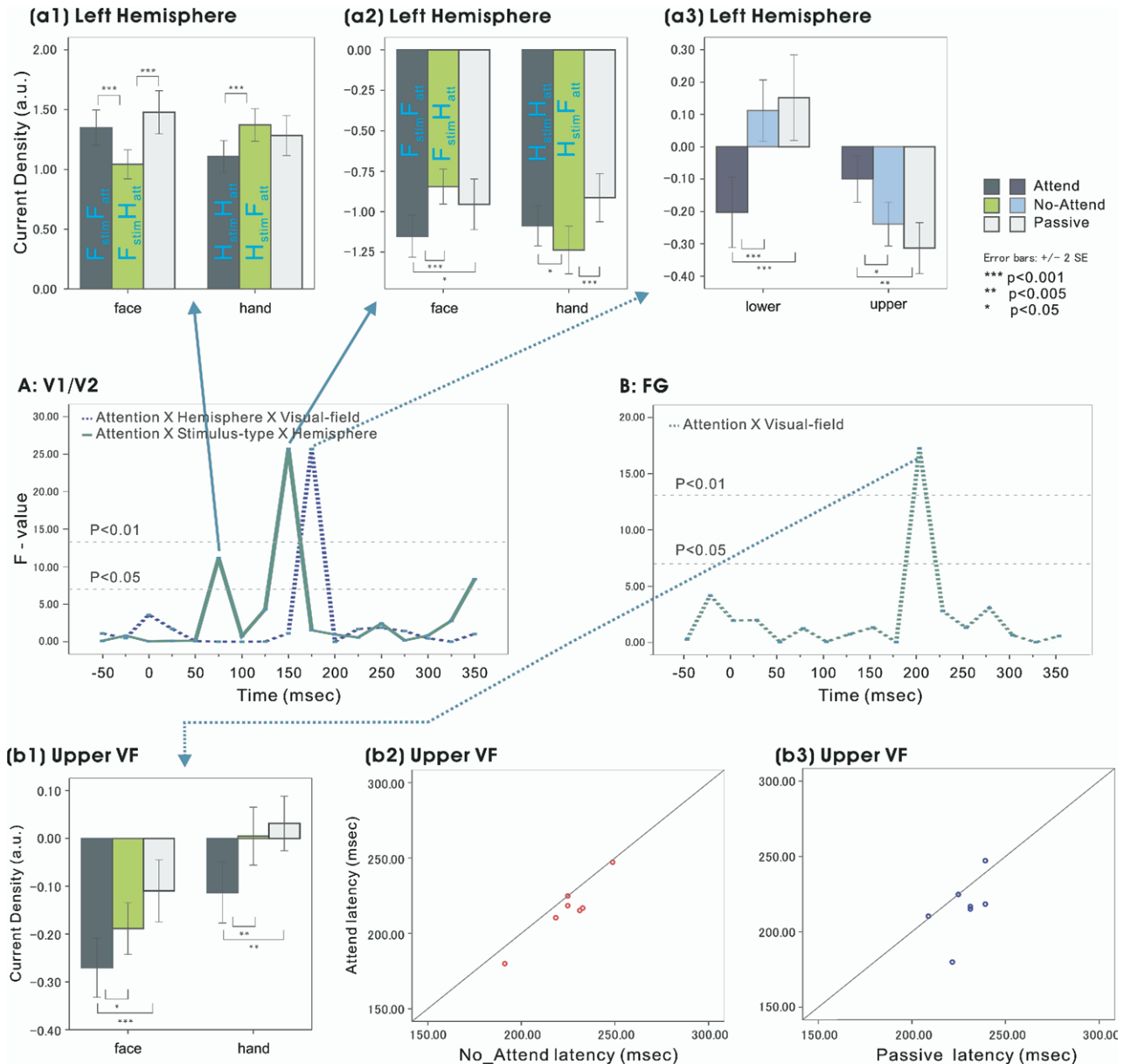
## Experiment 2

### Attention effect (active test: attend vs. no-attend).

The MEG data from experiment 1 were first used to define ROIs for each subject. These regions produced significantly higher responses for faces compared with other objects. In these ROIs, both the amplitude and latency of peak responses to the new conditions of experiment 2 were quantified in the same subject. We asked how selective attention modulates the neural responses which were identified in experiment 1. Four responses to non-target stimuli were compared: responses to non-target faces when subjects directed attention to a given target face ( $F_{stim}F_{att}$ ); when they attended to the target hand ( $F_{stim}H_{att}$ ); responses to non-target hands when subjects directed attention to the target hand ( $H_{stim}H_{att}$ ); and lastly, when they attended to the target face ( $H_{stim}F_{att}$ ). The RAC active test was applied to the current-density curves. For V1/V2, Attention (Attend, No-Attend) × Hemisphere (Right, Left) × VF (Upper, Lower), and Attention × Stimulus-type (Face, Hand) × Hemisphere interactions were statistically significant. For the FG, Attention × VF interactions were significant. Fig. 6A, B shows the amplitude of the  $F$ -value for each significant interaction, sampled with 12 time-slices, for V1/V2 and FG respectively.  $F$ -value peaks in the bar graph denote significant differences between Attend and No-Attend conditions for over five time-slices at 85.0 ms (Fig. 6A, a1), 150.0 ms (Fig. 6A, a2), and 170.0 ms (Fig. 6A, a3) in V1/V2, and at 200.0 ms (Fig. 6B, b1) in FG. Each bar graph shows responses to stimuli in Attend, No-Attend, and Passive conditions.

We identified the first attentional modulation in left striate cortex at 85.0 ms (Fig. 6A, a1), corresponding to the  $M_{v1}100$  component. It was rather weak and showed a tendency similar to next marked attentional effects. The next attention modulation was found in left V1/V2 at 150.0 ms (Fig. 6A, a2). For the Face stimuli, the Attend condition ( $F_{stim}F_{att}$ ) gave a significantly higher response than the No-Attend condition ( $F_{stim}H_{att}$ ), while the Attend condition for Hand ( $H_{stim}H_{att}$ ) was significantly lower than the No-Attend condition ( $H_{stim}F_{att}$ ). This implies that attending to faces ( $F_{stim}F_{att}$  and  $H_{stim}F_{att}$ ) always produces higher activation than attending to hands ( $F_{stim}H_{att}$  and  $H_{stim}H_{att}$ ) regardless of stimulus type. In fact, this modulation is seen not as a decrease in activity caused by attending to hands but as an increase caused by attending to faces, in comparison to responses in the passive viewing task of exper-

faces in the  $M_{v1}100$  component in both peripheral and central presentation. However  $M_{v1}100$  and  $M_{v1}130$  during central presentation shows that faces and hands are significantly higher than shoes (c). Error bars, S.E.M.; a.u., arbitrary units.



**Fig. 6.** (A, B) Amplitude of the  $F$ -values of the four-way interaction between Hemisphere (right, left), Visual-field (upper, lower), Stimulus-type (Faces, Hands), and Attention (Attend, No-Attend), sampled with 12 time slices, for V1/V2, and FG. The  $F$ -value peaks referring to the bar graphs denote significant differences between the Attend and No-Attend conditions lasting over five time-slices (8 ms). The results of the experiment 1 passive viewing condition have been added. In V1/V2,  $F$ -value peaks signifying attentional modulation can be seen at: 85.0 ms, 150.0 ms, and 170.0 ms (a1, a2, and a3). In the FG only one  $F$ -value peak is present signifying attentional modulation at 200.0 ms (b1). The relationship between the peak latency for  $F$ -value corresponding to  $M_{FG200}$  between the Attend and No-Attend (b2), or Passive (b3) conditions. Error bars, S.E.M.; a.u., arbitrary units.

iment 1. In other words, a comparison between the passive viewing condition and the attention to hands revealed no significant activation differences. However, attending to faces produced significantly higher activation than the passive viewing condition. Therefore, object responses for both stimuli increased so long as faces were attended. A third modulation in left V1/V2 was identified at 170.0 ms without interaction between Attention and Stimulus type but rather between Attention and VF (Fig. 6A, a3). Fig. 6A, a3 shows responses for both stimuli in the Attend, No-Attend, and Passive conditions. The Attend condition

( $F_{stim}F_{att}$  and  $H_{stim}H_{att}$ ) enhanced the object response in the lower VF but inhibited it in upper VF.

For both stimulus types, the Attend condition modulated the  $M_{FG200}$  component in FG (Fig. 6B). A significantly higher response in the Attend condition than in the No-Attend condition was identified at 200 ms in the upper VF stimulus presentation for both Face and Hand stimuli. Fig. 6B, b1 shows responses for each stimulus in the Attend, No-Attend, and Passive conditions. The Attend condition enhanced object responses as compared with the No-Attend condition and the Passive condition. This

evidence suggests that the selectivity of the neural population in the FG is not exclusive to faces, because the FG showed responses to hands as shown in experiment 1, and the responses to hands were increased by attending to hand stimuli. Moreover, these results support our hypothesis that the FG areas showing a face preference would be differentially modulated depending on whether attention was drawn to the same category as that of the presented stimulus (see Discussion).

The modulation in the Attend condition could have been caused not by an attention effect but rather by the difficulty of the task, because subjects were required to discriminate the gender of objects only in the Attend condition. Generally, the more difficult a task is, the longer the reaction time tends to be. In contrast, attention tends to increase processing speed. In order to determine whether indeed  $M_{FG200}$  modulation reflects an attentional effect, we examined the latency of  $M_{FG200}$  in the Attend, No-Attend and Passive conditions for individual subjects (Fig. 5B, b2, b3). The RAC baseline test in individual subjects gave a significant FG activation corresponding to the  $M_{FG200}$  component. Fig. 6B, b2, b3 shows the relationship of the latency, corresponding to the  $M_{FG200}$  component, between the Attend and No-Attend/Passive conditions in individual subjects. The results show that the  $M_{FG200}$  latency in the Attend condition is earlier than in the No-Attend and Passive conditions in most subjects. We therefore conclude that attention accelerates information processing related to the  $M_{FG200}$ .

## DISCUSSION

We extracted accurate, millisecond-by-millisecond tomographic estimates of brain activity from MEG data and used them to clarify two important issues. First, we defined the timing of face specificity in the FG, and specifically, we established that the FG is preferentially activated by Face stimuli compared with other objects within 100 ms. Second, we documented the timing and nature of modulations in the FG and striate cortex responses to Face and Non-Face stimuli for various degrees of attention toward Face and Non-Face categories.

### Face-selective responses

Regions of the ventral occipito-temporal pathway in the brain, such as a lateral part of the FG, have been shown to respond more to faces than other stimuli by numerous fMRI and PET studies, but they have not provided information on the timing of the processing stages. This crucial timing information is available with ERPs. The potential weaknesses in these studies are that the ERPs for faces were rarely compared with the non-face objects, and critically, that the precise location and complexity of sources generating these components were not investigated. In this study we extracted tomographic estimates of activity from MEG data to directly compare the activity elicited by face and non-face stimuli in specific, well-circumscribed brain regions.

We identified activity in the FG at the latency reported in many other studies ( $M_{FG170}$ ), and also at earlier latencies within 100 ms (i.e.  $M_{FG100}$ ), and at later latencies after 200 ms ( $M_{FG200}$ ). Both early and late components were reported in a few other studies (Itier and Taylor, 2004a; Liu et al., 2002; Liu and Ioannides, 2006), but ours is the first study to demonstrate a clear bias toward face stimuli for activity localized in the FG.

Our results relating to  $M_{FG170}$  are consistent with other MEG studies which found face preferences for a localized M170 in the inferior temporal cortex including the FG, and in superior temporal cortex and the middle temporal gyrus (Halgren et al., 2000; Streit et al., 1999; Watanabe et al., 1999). The peaks at the latencies reported in these studies approximately correspond to N170 as identified in many ERP studies including very early studies on face perception (Botzel and Grusser, 1989; Jeffreys, 1989), and more recently in combined ERP and fMRI studies (Henson et al., 2003). Moreover the wide distribution along the ventral and middle occipitotemporal region reported for the face-specific N200 with intracranial recording (Allison et al., 1999) is also found in our face-induced responses (Fig. 2D). It is therefore likely that these two widely distributed activations identified by MEG and intracranial recordings reflect similar activity, despite the difference in timing. Plausible explanations for the longer latency of the N200 are that object processing was delayed in the patients, especially as the other ERP components, P150 and P290, can be considered as shifted  $M_{FG100}$  and  $M_{FG200}$  components respectively. The effect of anticonvulsant medications taken by patients with epilepsy, the accuracy of the time-lock, stimulus size, and other differences between experimental designs could also contribute to these timing differences.

Compared with other objects, the amplitude for Face stimuli was significantly higher for central and peripheral presentations in an early component within 100 ms ( $M_{FG100}$ ), and in a late component after 200 ms ( $M_{FG200}$ ; see Fig. 4). These components could partly correspond to the P1 and P2. The inversion effect was used to make claims for specificity for processing faces for P1 and P2 have been made (Boutsen et al., 2006; Itier and Taylor, 2002; Latinus and Taylor, 2006; Rossion et al., 1999), but this has been hotly disputed because the inversion effect is much weaker for P2 and especially P1 than for the N170/M170 (Linkenkaer-Hansen et al., 1998; Rossion et al., 1999). However, a direct comparison between localized responses to face and non-face objects has yet to be performed in order to clarify the face selectivity of circumscribed brain activity at the corresponding latencies. The only direct comparison between face and non-face objects (Boutsen et al., 2006) showed no difference in EEG signal strength for P2. This apparent discrepancy likely arises because extracranial measurements of EEG or MEG reflect neural activity from a relatively large area of the cortex as Fig. 2 demonstrates. If no localization is attempted, selective responses from specific areas can be easily lost in the activity from other areas.



The demonstration of higher activity for face than non-face stimuli for  $M_{FG}100$  is particularly relevant to models of early object encoding. The absence of preference in V1/V2 (Fig. 5) makes it less likely that the effect we observe in the FG is entirely due to physical features. Our selection of hands and shoes as non-face stimuli has advantages and limitations. In our selection of stimuli, we placed more emphasis on using objects which were approximately symmetrical about the vertical axis, of similar size, three-dimensional, and dynamic in the sense that they implied biological action. We note two limitations of our design that should be addressed in future experiments to quantify any residual effects that may be attributable to differences in the physical properties of the stimuli. First, the physical properties of the stimuli could not be precisely controlled. Ideally, more controlled conditions for low-level visual parameters such as phase-scrambled faces should have also been used, but this would have unreasonably prolonged MEG recording time. Second, one complex object (a face) was compared with objects with two distinct elements (hands and shoes). We also acknowledge the possibility that face preference in  $M_{V1}100$  did not reach significance because of the small number of subjects. However, even if such an effect is identified with studies with more subjects, it is likely to be a weak one, since across our seven subjects the  $M_{V1}100$  amplitude showed no systematic behavior; it was stronger for the face stimuli in four and for hand stimuli in three subjects.

In only central presentation we find strong biased response for biological stimuli (faces and hands) in  $M_{V1}100$  and  $M_{V1}130$ . This result suggests that centrally and peripherally presented stimuli may be processed using different mechanisms (see Fig. 5c), a suggestion that echoes a similar conclusion reached recently by a study of illusory contours in our group (Bakar et al., *in press*). If the separation of objects start at higher levels in the visual hierarchy as early as our findings suggest, they must proceed in parallel with the ongoing striate cortex processing reflected in the  $M_{V1}100$  component of striate cortex, particularly for peripheral stimuli. The pattern of  $M_{V1}130$  activation for central presentation suggests that around that latency V1/V2 is the recipient of strong face-biased feedback signals, from extrastriate areas including direct or indirect input from the FG. The segregation of object representation rather early at higher levels of the visual hierarchy is supported by several human and animal studies, showing that activity in temporal cortex initially conveys coarse information such as stimulus category, and later, finer information such as identity (Liu et al., 2002; Sugase et al., 1999). In most reports, early face-induced components (before M170) have been localized in the occipital visual area but not in the FG (Linkenkaer-Hansen et al., 1998; Watanabe et al., 1999). This may be because the dipole analysis, which most studies use, fails to discriminate mixtures of sources. In a dipole analysis, “minor activity” such as the small  $M_{FG}100$  could easily be blurred by the adjacent “major activity.”

SPM results in the current study concur with other reports that the representations of faces are widely distrib-

uted in occipito-temporal visual areas. By focusing on the properties of object-induced activity in FG, we confirmed the face specificity of bilateral FG in three components. We note that hands and shoes also elicited activity in the same areas (see Fig. 3a, b). Analysis of successive peaks in the MEG signal of “sensors of interest” has been used to make claims for stages of processing (Liu et al., 2002). The existence of stages of processing has also been supported by a detailed analysis of pairwise measures of connectivity between brain areas derived from single-trial tomographic estimates of activity (Ioannides et al., 2004). We have recently proposed that stages of processing correspond to network excitations, with processing within each stage coordinated by activity in one or more “hub” areas (Ioannides, 2007). Our results show that one of the hub areas, the FG, is part of a network dedicated to the visual processing of objects in general (since it is activated by all objects), but with a relative specialization for faces. The next logical step in advancing our understanding would be to estimate the functional connectivity between each pair of areas as a function of time, and thus reconstruct the dynamics of network activity excited by different visual object categories.

### Attention and working memory

In general, it is difficult to separate one specific form of attention from others and from working memory effects. The confounding role of working memory is acknowledged at two levels; first as part of an attention task that relies on the display of a priming cue and second in the task requirement of holding the number of “hits” in memory. Avoiding working memory confounds altogether is difficult, some would say impossible because working memory and attention are inseparable, see for example (Awh et al., 2006) for a measured discussion of the relevant arguments. In our experiment the target object for a given block was indicated to the subject at the start of the block, i.e. the object bias response according to the “cue” involved typically top-down selection for the object. Presumably it required working memory, because relevant object information must be stored in working memory to compare with incoming perceptual representation of objects. This storage of target representation has been called the attentional template (Duncan and Humphreys, 1989; Desimone and Duncan, 1995), and it may be seen as one aspect of working memory (Baddeley, 1986). Furthermore, the top down selection template for objects probably shares common neural circuits with working memory, perhaps especially in prefrontal cortex (Courtney et al., 1997; Curtis et al., 2004; LaBar et al., 1999). This kind of working memory influence is present in most experiments on attention.

The second role of working memory in our task is the requirement to count the occurrences of targets. The working memory load involved is rather low and it is present in all conditions. We therefore accepted this slight confound in preference to a response in each trial to avoid possible motor contamination, especially for late latencies. The identification of attentional modulation in the third FG component justified this experimental design decision.

## Object-selective attention

The second aim of our study was to examine the role of attention for the responses elicited by stimuli on the periphery of the VFs. Many studies have found attentional modulation of neuronal responses in extrastriate cortex (Hayden and Gallant, 2005; Murray and Wojciulik, 2004; O'Craven et al., 1999), but there are few reports of attentional modulation in striate cortex (Luck et al., 1997; Motter, 1993; Poghosyan et al., 2005). Dipole modeling studies have suggested a mechanism whereby spatial attentional modulation first occurs in extrastriate cortical areas at a latency of 80–130 ms, and delayed feedback from these higher areas then modulates neural activity in V1 at 150–200 ms (Di Russo et al., 2003; Martinez et al., 2001). In our study we found instead that attentional modulation in V1/V2 precedes that of FG, which may simply have been missed by the simpler analyses used in earlier studies. Alternatively, the discrepancy may be related to the attentional control mechanisms involved in the type of attentional selection, i.e. spatial vs. non-spatial (Giesbrecht et al., 2003).

We first identified attentional modulation in V1/V2 at three distinct latency ranges, followed by an attentional effect in FG. Our results showed that attentional modulation in V1/V2 varies depending on conditions such as timing, attention, and the location of stimulus presentations. The V1/V2 attentional effect was confined to the left hemisphere. The earlier two attentional modulations appeared as a greater activity while attending to faces compared with attending to hands, and it was found for both face and hand stimuli (Fig. 6A, a1, a2). The difference between Attend and No-Attend condition in Fig. 6A, a1 is rather small and it also corresponds to a decrease in activity when attending to hand compared with the response in the Passive condition. The next significant attentional effect (Fig. 6A, a2) also showed that attending to face induced higher activity than attending to hand. Our results provide further evidence for biased-modulation in favor of particular objects in V1/V2, but the underlying mechanism remains unresolved.

Why would V1/V2, where there is no object selectivity as indicated in experiment 1, show face-biased modulation? For hand stimuli, attending to faces resulted in a larger response than attending to hands in left V1/V2, even in the No-Attend condition for Hands. This can be understood as a more effective modulation of V1/V2 when a face is a relevant stimulus, but with no specificity for faces at the level of V1/V2, as demonstrated in experiment 1. The response of V1/V2 would be enhanced by the top-down effect of expecting a target stimulus. If there is bias in subject expectation or image salience, then a face could affect V1/V2 more than the other stimuli. The third modulation in V1/V2 was around 170 ms. This was the first object-selective attentional effect, demonstrating that stimulus information is selectively modulated by attention to the same categorical object as the target (Fig. 6A, a3). It is assumed to be caused by a selective top-down effect following the encoding of an object. As discussed in many

studies of face encoding (Eimer, 2000; Mouchetant-Rostaing et al., 2000), object encoding can be completed by 170 ms after stimulus onset.

The laterality of the V1/V2 attentional effect may be related to the more conscious control of the dominant hemisphere. This explanation is plausible since we employed right-handed subjects (all with right-dominant vision), given earlier studies showing that the left inferior temporal cortex plays a dominant role in the discrimination of visual patterns in right-handed subjects (Kawashima et al., 1998), and that monocular viewing is associated with attentional systems in the contralateral hemisphere (Roth et al., 2002). More studies are needed to fully clarify this point.

If the three components identified in experiment 1 ( $M_{FG100}$ ,  $M_{FG170}$ , and  $M_{FG200}$ ), are face-specific, it is expected that these would be modulated by attending to faces, because selective attention to visual features or whole objects enhanced activity in the regions of extrastriate visual cortex that selectively process these same attributes. We hypothesized that attending to a face results in a top-down process which enhances responses in the extrastriate areas that encode the face. In face-target blocks, because the task consists of counting the number of target faces, subjects are likely to treat faces with higher priority than hands.

However,  $M_{FG100}$  and  $M_{FG170}$  were not affected, while  $M_{FG200}$  was significantly enhanced for face stimuli presented to the upper VF (Fig. 6B). Similar results were obtained for the Hand stimuli. These findings are not in line with a report on modulated visual processing of faces and scrambled faces by spatial attention (Jacques and Rossion, 2007). This report provided evidence that the P1 and N170 responses were modulated by spatial attention, but it did not take into account the sources of components or consider the possibility that components were not face-specific. As many studies with spatial attention observed early attentional modulation (Di Russo et al., 2003; Martinez et al., 2001), spatial attention may modulate visual processing earlier than object attention. Furthermore, it has not been clarified where the modulation in P1 and N170 occurred, not even distinguished between generators in striate or extrastriate cortex.

In our study, the possibility exists that the modulation in the Attend condition could be the result of the difficulty of the task rather than attention per se, because of the required discrimination of only the gender of objects in the Attend condition. Traditional reaction time studies have shown earlier responses in cases where prior expectations were fulfilled. We examined conditions where similar evidence could be seen in our Attention condition, such as the shortening of the  $M_{FG200}$  latency in the Attention condition. The results showed that the  $M_{FG200}$  latency in the Attend condition was significantly earlier than in the No-Attend conditions. Thus, we concluded that the object bias response in  $M_{FG200}$  according to the “cue” involved indeed top-down attentional selection for the target object.

It is generally agreed that components after N170, e.g. P2 deal with individual face recognition and other pro-

cesses that can only follow the structural encoding that faces achieve within N170. Decision making among competing percepts is helped by anticipating the forthcoming sensory environment, increasing in top-down connectivity from the frontal cortex to object-sensitive visual areas (Summerfield et al., 2006). The use of averaging limits our ability to recover activity in the frontal lobes which are expected to be determined more by endogenous factors, and hence to be rather loosely time-locked to the onset of the stimulus. The sequence of activations and the identification of the first FG attentional modulation for the  $M_{FG}200$  are consistent with the following scenario for object attention. Visual objects are first encoded during 170 ms and then matched to the expected target objects. When the expected and encoded objects match, the  $M_{FG}200$  response is enhanced, probably by top-down influences to object-sensitive visual areas via attention-control areas, such as the frontal cortex.

*Acknowledgments*—We thank Kenji Haruhana for his help in running the experiments, Vahe Poghosyan for his help in data analysis and Masaki Maruyama for his valuable discussion and advice in data analysis. We also thank Daniel Palomo for his help in redaction of the manuscript and valuable comments. Research at the University of Western Sydney (UWS) was supported by the UWS International Research Initiative Scheme (80461) and a Research Agreement with RIKEN Brain Science Institute.

## REFERENCES

- Abrahamyan A, Stevens C, Ioannides AA (2005) The effect of emblematic hand gestures on visual attention. *Aust J Psychol* 57:42.
- Allison T, Puce A, Spencer DD, McCarthy G (1999) Electrophysiological studies of human face perception. I: Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cereb Cortex* 9:415–430.
- Awh E, Vogel EK, Oh SH (2006) Interactions between attention and working memory. *Neuroscience* 139:201–208.
- Baddeley AD (1986) Working memory. Oxford: Oxford University Press.
- Bakar AA, Liu L, Conci M, Elliott MA, Ioannides AA (2007) Visual field and task influence illusory figure responses. *Hum Brain Mapp*, in press.
- Botzel K, Grusser OJ (1989) Electric brain potentials evoked by pictures of faces and non-faces: a search for “face-specific” EEG-potentials. *Exp Brain Res* 77:349–360.
- Boutsen L, Humphreys GW, Praamstra P, Warbrick T (2006) Comparing neural correlates of configural processing in faces and objects: an ERP study of the Thatcher illusion. *Neuroimage* 32:352–367.
- Corbetta M, Miezin FM, Dobmeyer S, Shulman GL, Petersen SE (1990) Attentional modulation of neural processing of shape, color, and velocity in humans. *Science* 248:1556–1559.
- Courtney SM, Ungerleider LG, Keil K, Haxby JV (1997) Transient and sustained activity in a distributed neural system for human working memory. *Nature* 386:608–611.
- Curtis CE, Rao VY, D’Esposito M (2004) Maintenance of spatial and motor codes during oculomotor delayed response tasks. *J Neurosci* 24:3944–3952.
- Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. *Annu Rev Neurosci* 18:193–222.
- Di Russo F, Martinez A, Hillyard SA (2003) Source analysis of event-related cortical activity during visuo-spatial attention. *Cereb Cortex* 13:486–499.
- Duncan J, Humphreys GW (1989) Visual search and stimulus similarity. *Psychol Rev* 96:433–458.
- Eimer M (2000) The face-specific N170 component reflects late stages in the structural encoding of faces. *Neuroreport* 11:2319–2324.
- Ekman P, Friesen WV (1976) Pictures of facial affect. Palo Alto, CA: Consulting Psychologists Press.
- Fisher NI (1993) Statistical analysis of circular data. New York, NY: Cambridge University Press.
- Gauthier I, Skudlarski P, Gore JC, Anderson AW (2000) Expertise for cars and birds recruits brain areas involved in face recognition. *Nat Neurosci* 3:191–197.
- Giesbrecht B, Woldorff MG, Song AW, Mangun GR (2003) Neural mechanisms of top-down control during spatial and feature attention. *Neuroimage* 19:496–512.
- Hadjikhani N, de Gelder B (2002) Neural basis of prosopagnosia: an fMRI study. *Hum Brain Mapp* 16:176–182.
- Halgren E, Raji T, Marinkovic K, Jousmaki V, Hari R (2000) Cognitive response profile of the human fusiform face area as determined by MEG. *Cereb Cortex* 10:69–81.
- Hayden BY, Gallant JL (2005) Time course of attention reveals different mechanisms for spatial and feature-based attention in area V4. *Neuron* 47:637–643.
- Henson RN, Goshen-Gottstein Y, Ganel T, Otten LJ, Quayle A, Rugg MD (2003) Electrophysiological and haemodynamic correlates of face perception, recognition and priming. *Cereb Cortex* 13:793–805.
- Hironaga N, Ioannides AA (2002) Accurate co-registration for MEG reconstructions. In: Proceedings of the 13th international conference on biomagnetism (Nowak H, Hauelsen J, Giebler F, Huonker R, eds), pp 931–933. Berlin: Verlag.
- Ioannides AA (2007) Dynamic functional connectivity. *Curr Opin Neurobiol* 17:161–170.
- Ioannides AA, Bolton JPR, Clarke CJS (1990) Continuous probabilistic solutions to the biomagnetic inverse problem. *Inverse Problems* 6:523–542.
- Ioannides AA, Fenwick PB, Liu L (2005) Widely distributed magnetoencephalography spikes related to the planning and execution of human saccades. *J Neurosci* 25:7950–7967.
- Ioannides AA, Poghosyan V, Dammers J, Streit M (2004) Real-time neural activity and connectivity in healthy individuals and schizophrenia patients. *Neuroimage* 23:473–482.
- Ishai A, Ungerleider LG, Martin A, Schouten JL, Haxby JV (1999) Distributed representation of objects in the human ventral visual pathway. *Proc Natl Acad Sci U S A* 96:9379–9384.
- Itier RJ, Herdman AT, George N, Cheyne D, Taylor MJ (2006) Inversion and contrast-reversal effects on face processing assessed by MEG. *Brain Res* 1115:108–120.
- Itier RJ, Taylor MJ (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 RJ, Taylor MJ (2004a) Face inversion and contrast-reversal effects across development: in contrast to the expertise theory. *Dev Sci* 7:246–260.
- Itier RJ, Taylor MJ (2004b) N170 or N1? Spatiotemporal differences between object and face processing using ERPs. *Cereb Cortex* 14:132–142.
- Jacques C, Rossion B (2007) Electrophysiological evidence for temporal dissociation between spatial attention and sensory competition during human face processing. *Cereb Cortex* 17:1055–1065.
- Jahn O, Cichocki A, Ioannides AA, Amari S (1999) Identification and elimination of artifacts from MEG signals using extended independent component analysis. In: Recent advances in biomagnetism (Yoshimoto T, Kotani M, Kuriki S, Karibe H, Nakasato N, eds), pp 224–228. Sendai: Tohoku University Press.
- Jeffreys DA (1989) A face-responsive potential recorded from the human scalp. *Exp Brain Res* 78:193–202.
- Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci* 17:4302–4311.



- Kawashima R, Satoh K, Goto R, Inoue K, Itoh M, Fukuda H (1998) The role of the left inferior temporal cortex for visual pattern discrimination—a PET study. *Neuroreport* 9:1581–1586.
- LaBar KS, Gitelman DR, Parrish TB, Mesulam M (1999) Neuroanatomic overlap of working memory and spatial attention networks: a functional MRI comparison within subjects. *Neuroimage* 10:695–704.
- Latinus M, Taylor MJ (2006) Face processing stages: Impact of difficulty and the separation of effects. *Brain Res* 1123:179–187.
- Lavie N, Tsal Y (1994) Perceptual load as a major determinant of the locus of selection in visual attention. *Percept Psychophys* 56:183–197.
- Linkenkaer-Hansen K, Palva JM, Sams M, Hietanen JK, Aronen HJ, Ilmoniemi RJ (1998) Face-selective processing in human extrastriate cortex around 120 ms after stimulus onset revealed by magneto- and electroencephalography. *Neurosci Lett* 253:147–150.
- Liu J, Harris A, Kanwisher N (2002) Stages of processing in face perception: an MEG study. *Nat Neurosci* 5:910–916.
- Liu L, Ioannides AA (2006) Spatiotemporal dynamics and connectivity pattern differences between centrally and peripherally presented faces. *Neuroimage* 31:1726–1740.
- Liu L, Ioannides AA, Streit M (1999) Single trial analysis of neurophysiological correlates of the recognition of complex objects and facial expressions of emotion. *Brain Topogr* 11:291–303.
- Luck SJ, Chelazzi L, Hillyard SA, Desimone R (1997) Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J Neurophysiol* 77:24–42.
- Martinez A, DiRusso F, Anillo-Vento L, Sereno MI, Buxton RB, Hillyard SA (2001) Putting spatial attention on the map: timing and localization of stimulus selection processes in striate and extrastriate visual areas. *Vision Res* 41:1437–1457.
- Mishkin M, Ungerleider LG, Macko KA (1983) Object vision and spatial vision: two cortical pathways. *Trends Neurosci* 6:414–417.
- Motter BC (1993) Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *J Neurophysiol* 70:909–919.
- Mouchetant-Rostaing Y, Giard MH, Bentin S, Aguera PE, Pernier J (2000) Neurophysiological correlates of face gender processing in humans. *Eur J Neurosci* 12:303–310.
- Murray SO, Wojciulik E (2004) Attention increases neural selectivity in the human lateral occipital complex. *Nat Neurosci* 7:70–74.
- O'Craven KM, Downing PE, Kanwisher N (1999) fMRI evidence for objects as the units of attentional selection. *Nature* 401:584–587.
- O'Craven KM, Rosen BR, Kwong KK, Treisman A, Savoy RL (1997) Voluntary attention modulates fMRI activity in human MT-MST. *Neuron* 18:591–598.
- Poghosyan V, Shibata T, Ioannides AA (2005) Effects of attention and arousal on early responses in striate cortex. *Eur J Neurosci* 22:225–234.
- Rees G, Frith CD, Lavie N (1997) Modulating irrelevant motion perception by varying attentional load in an unrelated task. *Science* 278:1616–1619.
- Rossion B, Delvenne JF, Debatisse D, Goffaux V, Bruyer R, Crommelinck M, Guerit JM (1999) Spatio-temporal localization of the face inversion effect: an event-related potentials study. *Biol Psychol* 50:173–189.
- Roth HL, Lora AN, Heilman KM (2002) Effects of monocular viewing and eye dominance on spatial attention. *Brain* 125:2023–2035.
- Schwarzlose RF, Baker CI, Kanwisher N (2005) Separate face and body selectivity on the fusiform gyrus. *J Neurosci* 25:11055–11059.
- Seeck M, Michel CM, Mainwaring N, Cosgrove R, Blume H, Ives J, Landis T, Schomer DL (1997) Evidence for rapid face recognition from human scalp and intracranial electrodes. *Neuroreport* 8:2749–2754.
- Streit M, Ioannides AA, Liu L, Wolwer W, Dammers J, Gross J, Gaebel W, Muller-Gartner HW (1999) Neurophysiological correlates of the recognition of facial expressions of emotion as revealed by magnetoencephalography. *Brain Res Cogn Brain Res* 7:481–491.
- Sugase Y, Yamane S, Ueno S, Kawano K (1999) Global and fine information coded by single neurons in the temporal visual cortex. *Nature* 400:869–873.
- Summerfield C, Egner T, Greene M, Koechlin E, Mangels J, Hirsch J (2006) Predictive codes for forthcoming perception in the frontal cortex. *Science* 314:1311–1314.
- Tanner WP Jr, Swets JA (1954) A decision-making theory of visual detection. *Psychol Rev* 61:401–409.
- Taylor JG, Ioannides AA, Muller-Gartner HW (1999) Mathematical analysis of lead field expansions. *IEEE Trans Med Imaging* 18:151–163.
- VanRullen R, Koch C (2003) Competition and selection during visual processing of natural scenes and objects. *J Vis* 3:75–85.
- Watanabe S, Kakigi R, Koyama S, Kirino E (1999) Human face perception traced by magneto- and electro-encephalography. *Brain Res Cogn Brain Res* 8:125–142.

(Accepted 5 February 2008)  
(Available online 10 February 2008)