

# Occipitotemporal activation evoked by the perception of human bodies is modulated by the presence or absence of the face

James P. Morris<sup>a</sup>, Kevin A. Pelphrey<sup>a,b</sup>, Gregory McCarthy<sup>a,b,c,\*</sup>

<sup>a</sup> Duke-UNC Brain Imaging and Analysis Center, Duke University, Durham, North Carolina, USA

<sup>b</sup> Department of Psychological and Brain Sciences, Duke University, Durham, North Carolina, USA

<sup>c</sup> Department of Veterans Affairs Medical Center, Durham, North Carolina, USA

Received 16 June 2005; received in revised form 23 January 2006; accepted 30 January 2006

Available online 20 March 2006

## Abstract

We investigated the degree to which activation in regions of the brain known to participate in social perception is influenced by the presence or absence of the face and other body parts. Subjects continuously viewed a static image of a lecture hall in which actors appeared briefly in various poses. There were three conditions: *Body-Face*, in which the actor appeared with limbs, torso, and face clearly visible; *Body-Only*, in which the actor appeared with his or her face occluded by a book; and *Face-Only*, in which the actor appeared behind a podium with only face and shoulders visible. Using event-related functional MRI, we obtained strong activation in those regions previously identified as important for face and body perception. These included portions of the fusiform (FFG) and lingual gyri within ventral occipitotemporal cortex (VOTC), and portions of the middle occipital gyrus (corresponding to the previously defined extrastriate body area, or EBA) and posterior superior temporal sulcus (pSTS) within lateral occipitotemporal cortex (LOT). Activation of the EBA was strongest for the *Body-Only* condition; indeed, exposing the face *decreased* EBA activation evoked by the body. In marked contrast, activation in the pSTS was largest when the face was visible, regardless of whether the body was also visible. Activity within the lateral lingual gyrus and adjacent medial FFG was strongest for the *Body-Only* condition, while activation in the lateral FFG was greatest when both the face and body were visible. These results provide new information regarding the importance of a visible face in both the relative activation and deactivation of brain structures engaged in social perception.

© 2006 Elsevier Ltd. All rights reserved.

**Keywords:** Social perception; Face processing; Social cognitive neuroscience; fMRI

## 1. Introduction

The ability to effortlessly process the identity, intentions, and actions of others constitutes an advanced skill set that we refer to as “social perception” (Allison, Puce, & McCarthy, 2000). Prior work has identified two broad brain regions, ventral occipitotemporal cortex (VOTC) and lateral occipitotemporal cortex (LOT), that appear functionally specialized for processing different types of social information. In VOTC, subdural electrical recordings (Allison et al., 1994a; Allison, Puce, Spencer, & McCarthy, 1999; McCarthy, Puce, Belger, & Allison, 1999; McCarthy, Puce, Gore, & Allison, 1997; Puce, Allison,

& McCarthy, 1999) and functional magnetic resonance imaging (fMRI) studies (Grill-Spector, Knouf, & Kanwisher, 2004; Kanwisher, McDermott, & Chun, 1997; Puce, Allison, Asgari, Gore, & McCarthy, 1996; Puce, Allison, Gore, & McCarthy, 1995) have identified a discrete region within the lateral fusiform gyrus (FFG) that responds selectively for faces or “face-like” stimuli.

Similarly, subdural electrical recordings (Allison et al., 1999) and fMRI studies (Downing, Jiang, Shuman, & Kanwisher, 2001; Puce et al., 1995, 1996) have identified at least two regions in LOTC that also respond to faces and/or body parts. The subdural recordings reported by Allison et al. (Allison et al., 1999), consisted of a cluster of sites centered upon the middle occipital gyrus (MOG) from which face-specific field potentials were obtained that were similar in their response properties to the aforementioned fusiform sites. The fMRI studies of Puce and colleagues (Puce et al., 1995, 1996) identified face activation in this same region of the MOG and a second region of activation in

\* Corresponding author at: Brain Imaging and Analysis Center Box 3918, Duke University Medical Center, Durham, NC 27710, Tel.: +919 681 9337; fax: +919 681 7033.

E-mail address: [gregory.mccarthy@duke.edu](mailto:gregory.mccarthy@duke.edu) (G. McCarthy).

the posterior superior temporal sulcus (pSTS). The subsequent demonstration that this pSTS region was differentially sensitive to the perception of eye and mouth movements (Hoffman & Haxby, 2000; Puce, Allison, Bentin, Gore, & McCarthy, 1998) led to formulations that emphasized a ventral fusiform system for face identity and a more dorsal pSTS system for processing facial and biological motion (see review by McCarthy, 1999).

This formulation, however, does not adequately explain all face and body part activity obtained within either LOTC or VOTC. For example, Downing and colleagues (Downing et al., 2001), identified activation within and near the MOG that was responsive to images of body forms and body parts, but showed little activation to images of isolated faces. Given these properties, they named this region the “extrastriate body area” or EBA. The specific role of the FFG in face processing has also been challenged. For example, a recent report indicated that portions of the FFG demonstrate selectivity for images of body parts and human bodies without faces (Peelen & Downing, 2004). Two other studies have demonstrated that regions of the FFG are sensitive to images of bodies conveying emotions even when facial expression is blurred (de Gelder, Snyder, Greve, Gerard, & Hadjikhani, 2004; Hadjikhani & de Gelder, 2003). These latter results raise two, non-exclusive possibilities: First, the FFG may contain more than one category selective region. Second,

face regions may be activated by contextual factors that imply the presence of an unseen face. Indeed, Cox and colleagues have demonstrated “face-like” responses in the FFG even when the face was not viewable, suggesting that the context in which a face is usually present may be capable of driving activation in the face specific region of the FFG (Cox, Meyers, & Sinha, 2004). Similarly, the face-specific N170 component, extracted from scalp-recorded event-related potentials, has also been shown to be sensitive to contextual factors (Bentin & Golland, 2002; Bentin, Sagiv, Mecklinger, Friederici, & von, 2002).

These recent studies also question the functional distinctions previously drawn between activations obtained in the FFG and the EBA. Our own experience has been that differentiating the functional properties of the VOTC and LOTC is challenging. For example, in a recent fMRI study (Morris, Pelphrey, & McCarthy, 2005), we observed similar profiles of activity in the FFG region of the VOTC and an area of LOTC encompassing the EBA when subjects experienced the perception of approaching either a three-dimensional character in a virtual reality environment or an inanimate portrait of a face without a body. Both the FFG and EBA regions were more strongly activated by the experience of approaching the three-dimensional character compared to the face portrait, leading us to question whether activity in both regions was simply related to the number of visible body parts.

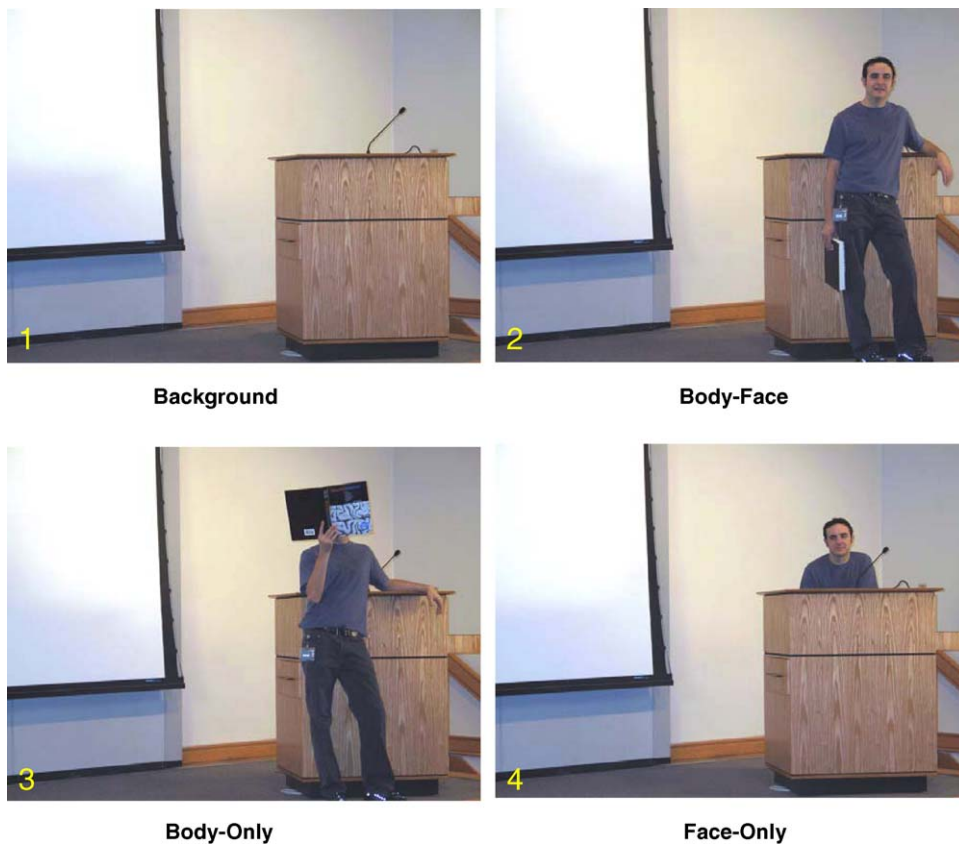


Fig. 1. Experimental conditions. In each experimental run, subjects viewed an empty stage containing a podium into which one of three different actors appeared briefly in one of three poses consistent with a lecture scene. In the *Body-Face* condition, the actor appeared in the scene with both the face and body clearly visible. In the *Body-Only* condition, the actors appeared with a book occluding his or her face. In the *Face-Only* condition, the actor stood behind a podium with only the face, neck and shoulders visible.

These findings have thus led us here to investigate the degree to which activations within the EBA, pSTS, and FFG are influenced by the visibility of body parts of a human appearing in a familiar environment. We frequently encounter other humans in situations in which body parts are occluded from direct view, but where we can reasonably assume that those occluded body parts exist and are only hidden. We wished to determine whether the EBA, pSTS, and/or FFG responded equally to the presence of a human in the scene, regardless of partial occlusion, or whether the activations in these regions were strongly modulated by the number and nature of the body parts that were visible.

In our experiment, subjects viewed a static lecture hall (see Fig. 1, Panel 1) where actors appeared briefly in different poses. On *Body-Face* trials (Fig. 1, Panel 2), the actor's face, torso, and limbs were fully visible. On *Body-Only* trials (Fig. 1, Panel 3), the actor's face was occluded by a book from which he or she appeared to be reading. On *Face-Only* trials (Fig. 1, Panel 4), the actor's face and shoulders were visible, but his or her body and limbs were occluded by a podium.

Based upon our recent work, and upon the initial study by Downing and colleagues (Downing et al., 2001), we hypothesized that the EBA is sensitive to the amount of the body visible, or to the number of body parts visible excluding the face. Thus, we predicted the strongest EBA activation would occur for *Body-Only* and *Body-Face* trials, the full body was visible in each. As no explicit manipulation of biological motion was included in this study, we did not provide an optimal stimulus for evaluating the response of the pSTS. Nevertheless, given its previously demonstrated sensitivity to face and mouth movements, we expected that the pSTS would respond strongest when the face was visible. Finally, we hypothesized that activation within the lateral FFG, would be potentiated by the presence of a viewable face.

## 2. Methods

### 2.1. Participants

Twelve healthy volunteers (age 19–37; 8 female, 4 male) participated in this study. All volunteers had normal or corrected-to-normal vision, and all were screened against neurological and psychiatric illnesses. Volunteers gave informed consent and were paid for their participation. The Duke University Medical Center Institutional Review Board approved this project.

### 2.2. Experimental design

Subjects viewed a picture of an empty lecture stage through scanner-compatible LCD goggles. Although the lecture stage was continuously displayed throughout the run, different male and female actors briefly appeared within this scene and assumed three different poses consistent with a lecture scenario (Fig. 1) that comprised the experimental conditions. In the *Body-Face* pose, the face, torso, and limbs of the actor were clearly visible. In the *Body-Only* pose, the actor's face was occluded by a text book. In the *Face-Only* pose, the actor appeared behind a podium with only his or her head and shoulders visible. The actor appeared instantly within the scene in the assumed pose; i.e., there was no explicit motion. There were three different actors and poses were presented in random order. Each pose was presented for three seconds and there was a twelve second interval between successive poses. Subjects were instructed to simply view the stimuli. The experiment contained 8 runs and each run contained 18 trials providing a total of 48 trials for each of the three conditions.

### 2.3. Imaging

Scanning was performed on a General Electric 4T LX NVi MRI scanner system equipped with 41 mT/m gradients (General Electric, Waukesha, Wisconsin, USA). A quadrature birdcage radio frequency (RF) head coil was used for transmit and receive. The subject's head was immobilized using a vacuum cushion and tape. Sixty-eight high resolution images were acquired using a 3D fast SPGR pulse sequence (TR = 500 ms; TE = 20 ms; FOV = 24 cm; image matrix =  $256^2$ ; voxel size =  $0.9375 \text{ mm} \times 0.9375 \text{ mm} \times 1.9 \text{ mm}$ ) and used for coregistration with the functional data. These structural images were aligned in the near axial plane defined by the anterior and posterior commissures. Whole brain functional images were acquired using a gradient-recalled inward spiral pulse sequence (Glover & Law, 2001; Guo & Song, 2003) sensitive to blood oxygenation level dependent (BOLD) contrast (TR, 1500 ms; TE, 35 ms; FOV, 24 cm; image matrix,  $64^2$ ;  $\alpha = 62^\circ$ ; voxel size,  $3.75 \times 3.75 \times 3.8 \text{ mm}$ ; 34 axial slices). These functional images were similarly aligned as the structural images. A semi-automated high-order shimming program ensured global field homogeneity. Runs consisted of the acquisition of 206 successive brain volumes and began with 4 discarded RF excitations to allow for steady state equilibrium.

### 2.4. Data analysis

Image preprocessing was performed with custom programs and SPM modules (Wellcome Department of Cognitive Neurology, UK). Head motion was detected by center of mass measurements. No subject had greater than a 3-mm deviation in the center of mass in any dimension. Images were time-adjusted to compensate for the interleaved slice acquisition and realigned to the tenth image to correct for head movements between scans. The realigned scans were then normalized to the Montréal Neurologic Institute (MNI) template found in SPM 99. The functional data were spatially smoothed with an 8 mm isotropic Gaussian kernel prior to statistical analysis. These normalized and smoothed data were used in the analysis procedures described below.

We used an internal localizer analysis (Lerner, Hendler, & Malach, 2002) in which half of the data was used to define functional regions of interest (ROIs), while the other half of the data was used to evaluate differences in those functional ROIs. This technique allows for two independent statistical assessments of differences between conditions.

To identify functional ROIs, we used half of the data to conduct a random-effects assessment of the differences among the three conditions at the peak of the hemodynamic response (HDR). This analysis consisted of the following steps: (1) The epoch of image volumes beginning 2 images before ( $-3.0 \text{ s}$ ) and 8 images after ( $13.5 \text{ s}$ ) the onset of each pose was excised from the continuous time series of volumes. (2) The average intensity of the HDR waveform was computed from 4.5–6 s post-stimulus all three conditions. A  $t$ -statistic was then computed at each voxel within the brain to quantify the HDR differences among the poses. This process was performed separately for each subject. (3) The individual  $t$ -maps created in the preceding step were then subjected to a random-effects analysis that assessed the significance of differences between conditions across-subjects.

To reduce the number of statistical comparisons and thus the false positive rate, the results of the random-effects analyses computed above were then restricted to only those voxels in which a significant HDR was evoked by any of the three different poses. For this analysis, we thresholded our activation at a False Discovery Rate (FDR) (Genovese, Lazar, & Nichols, 2002) of 0.01 ( $t_{(11)} > 4.37$ ). The voxels with significant HDRs were identified in the following steps: (4) For each subject, the average evoked hemodynamic response was computed for each of the three trial types and these average BOLD-intensity signal values for each voxel within the averaged epochs were converted to percent signal change relative to the pre-stimulus baseline. (5) The time waveforms for each voxel were correlated with an empirical reference waveform and  $t$ -statistics were calculated for the correlation coefficients. This procedure provided a whole-brain  $t$ -map in MNI space for each of the three trial types. (6) The  $t$ -maps for each subject and for each trial type were used to calculate an average  $t$ -map for the union of all three trial types across subjects. We then identified active voxels as those that surpassed the FDR threshold. (7) The difference  $t$ -map computed in step 3 above was then masked by the results of step 6. Thus, the differences in HDR amplitude between poses were only evaluated for those voxels in which at least one pose evoked a significant HDR as defined above. The threshold for

significance in the HDR peak was set at  $p < .01$  (two-tailed, uncorrected) and a minimal spatial extent of 4 uninterpolated voxels.

For each functional ROI showing a significant difference between any two trial types, we computed the average time course for each condition on the second subset of epochs that were withheld from the analysis used for identification of the functional ROIs. We then conducted a repeated-measures analysis of variance for each region of interest in order to statistically evaluate differences in peak amplitude as a function of stimulus condition. Greenhouse–Geisser epsilons were used to adjust the degrees of freedom in order to guard against violations of the sphericity assumption.

### 3. Results

#### 3.1. Regions in which poses evoked a significant HDR

As expected, each of the three poses evoked significant activation in the LOTC and VOTC. The overlay in Fig. 2 depicts regions in which, across subjects, a significant HDR was evoked by at least one of the three different poses (see analysis step 6 in Methods). The LOTC activation included both the MOG, encompassing the expected location of the functionally defined EBA, and extended anterior and superior into the pSTS. The VOTC activation included the FFG but also extended medially to encompass the lingual gyrus. Strong activation was also obtained in two regions for which we offered no predictions: the intraparietal sulcus (IPS) bilaterally and the right

frontal lobe within and near the precentral gyrus (PCG). Differences in activation between poses were not observed in either the IPS or PCG and we will not consider these regions further.

#### 3.2. Lateral occipitotemporal cortex (LOTC)

##### 3.2.1. Extrastriate body area (EBA)

Regions of the bilateral LOTC showed significant amplitude differences as a function of condition,  $F_{(2,14)} = 12.60$ ,  $p < .01$ . Within this broad region of activation, a large subset of voxels centered on the MOG evoked greater activity for *Body-Only* trials compared to both *Body-Face* and *Face-Only* trials (Fig. 3A). An analysis of the time course data, sampled independently from the data represented in the contrasts, confirmed that these voxels evoked a significantly greater response at peak for *Body-Only* trials compared to *Face-Only* trials,  $t_{(10)} = 5.39$ ,  $p < .001$  and *Body-Face* trials,  $t_{(10)} = 2.65$ ,  $p = .05$ . Furthermore, the average evoked activity at peak for *Body-Face* trials was significantly larger than the average evoked activity at peak for *Face-Only* trials,  $t_{(10)} = 3.00$ ,  $p < .05$ . The average HDR waveforms from the time course analysis are plotted in Fig. 3B. The MNI coordinates for the centroids of these regions ( $x = 53$ ,  $y = -81$ ,  $z = 9$ ;  $x = -42$ ,  $y = -82$ ,  $z = 9$ ) were within 12.9 and 11.7 mm, respectively of the coordinates ( $x = 51$ ,  $y = -71$ ,  $z = 1$ ;  $x = -51$ ,  $y = -72$ ,

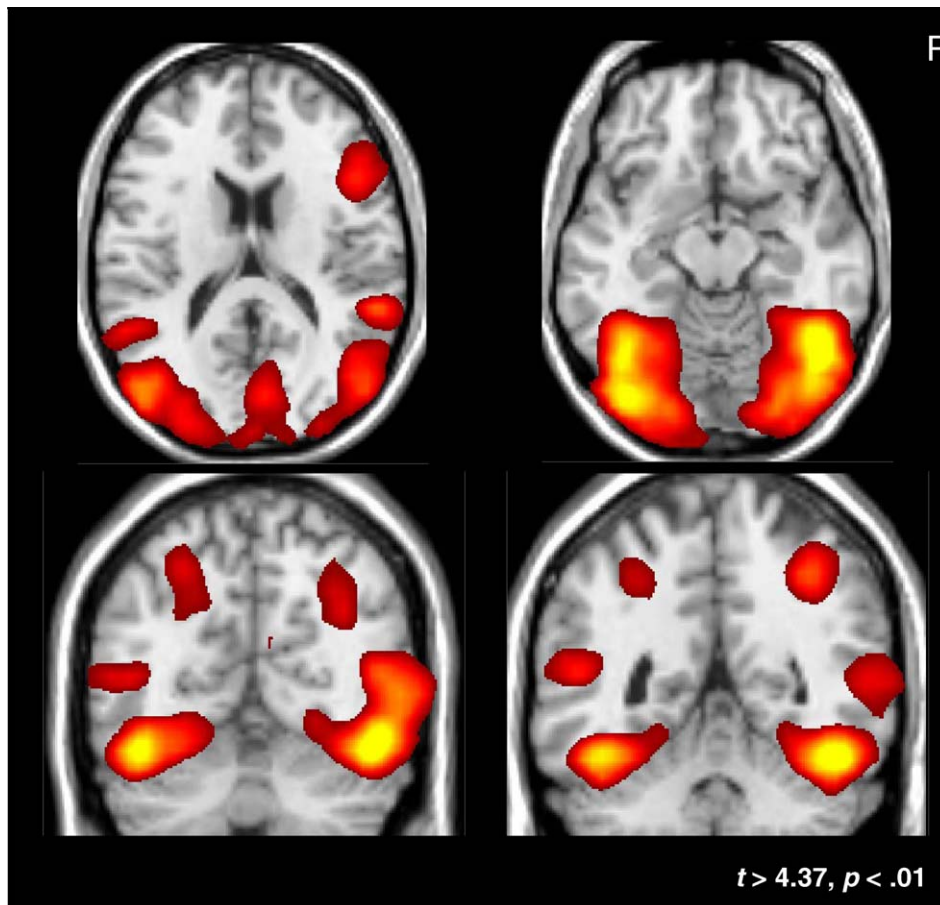


Fig. 2. Active regions. Grand-average  $t$ -maps plotted on a template brain. The  $t$ -maps were thresholded at an FDR of  $p < .01$  ( $t < 4.37$ ) and represent regions of the brain that are active to any of the three experimental conditions. These regions served as the basis for all further statistical analyses.

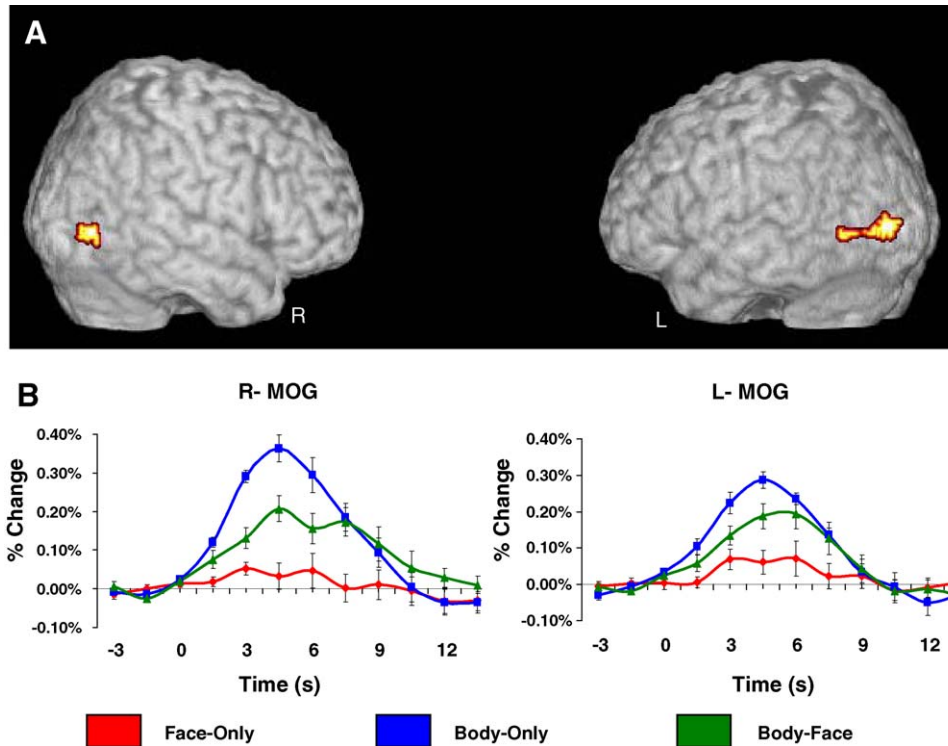


Fig. 3. Lateral occipitotemporal cortex. Panel A reflects the results of a random-effects analysis where peak amplitude was dependent upon stimulus condition. Panel B displays the average HDR plotted as a function of condition for the regions of interest highlighted in the colormap. The HDR was sampled from a different subset of trials that were used to define the functional regions of interest. Regions of the right and left LOTC evoked a significantly greater response for *Body-Only* trials when compared to both *Face-Only* and *Body-Face* trials ( $p < .01$ ). *Body-Face* trials evoked significantly more activity at peak than *Face-Only* trials ( $p < .05$ ).

$z = 8$ ) originally reported by Downing and colleagues for EBA (Downing et al., 2001) which were well within the spatial extent of the current activation.

3.2.2. *Posterior superior temporal sulcus (pSTS)*

A random-effects analysis revealed a focal cluster of activation in the pSTS (Fig. 4A), where the average evoked response for *Body-Only* trials was significantly smaller in amplitude than the average response evoked by *Face-Only* and *Body-Face* trials.

A repeated measures ANOVA conducted on the independently sampled time course data confirmed a difference in amplitude at peak as a function of condition,  $F_{(2,18)} = 5.27, p < .05$ . The average HDR evoked for *Body-Only* trials was significantly smaller in amplitude at peak compared to both *Body-Face*  $t_{(10)} = 2.32, p < .05$ , and *Face-Only* trials,  $t_{(10)} = 3.69, p < .01$ . The average evoked HDRs for *Body-Only* and *Body-Face* trials did not differ from each other. The average evoked HDR from the time course analysis are plotted in Fig. 4B.

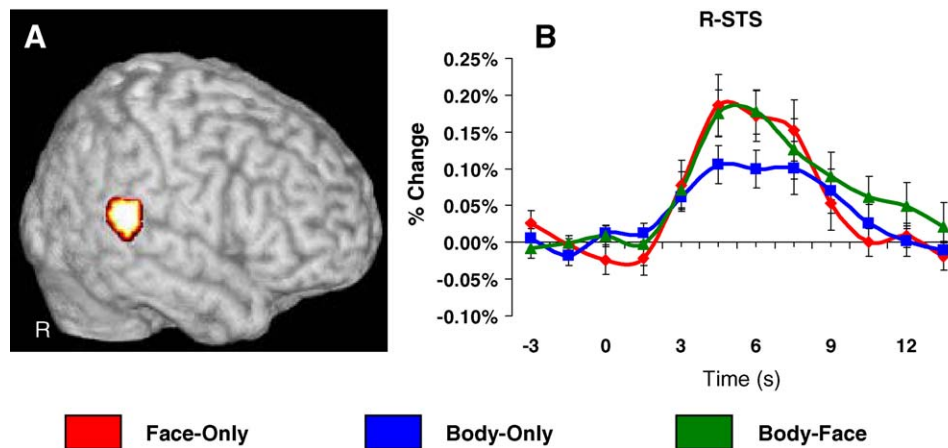


Fig. 4. Right superior temporal sulcus. Panel A reflects the results of a random-effects analysis where peak amplitude was dependent upon stimulus condition. Panel B displays the average HDR plotted as a function of condition for the regions of interest highlighted in the colormap. The HDR was sampled from a different subset of trials that were used to define the functional region of interest. The right pSTS evoked significantly greater activity to *Face-Only* and *Body-Face* trials when compared to *Body-Only* trials ( $p < .05$ ). There was no significant difference between *Face-Only* and *Body-Face* trials.

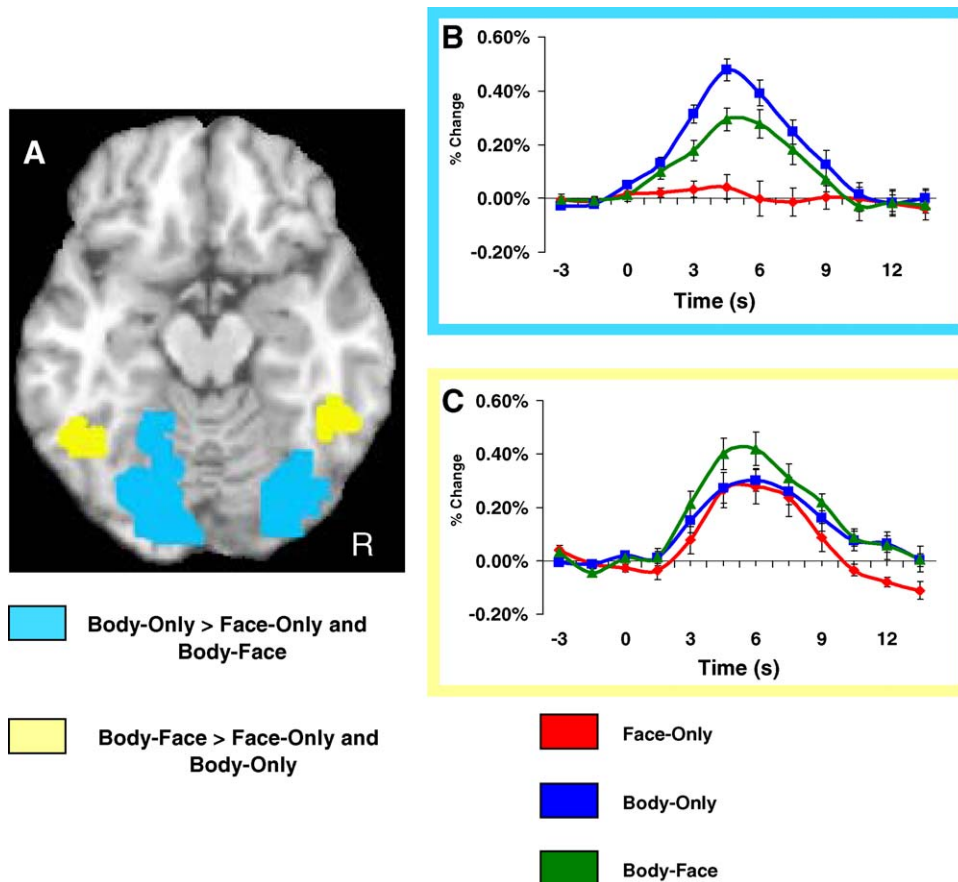


Fig. 5. Ventral occipitotemporal activation. Panel A shows two different regions of interest in the VOTC. The light-blue color map displays voxels that responded significantly stronger for *Face-Only* trials compared to both *Body-Only* and *Body-Face* trials in a random-effects analysis that was employed to identify regions of interest. The yellow color map displays voxels that responded significantly stronger for *Body-Face* trials compared to both *Body-Only* and *Face-Only* trials. Panel B displays the average HDR from the voxels displayed in the light blue color map. The HDR was independently sampled from a separate subset of trials that were used to define the functional region of interest. The average peak amplitude varied as a function of stimulus condition. Specifically, *Body-Only* trials evoked a larger peak amplitude response than that evoked by both *Face-Only* and *Body-Face* trials ( $p < .05$ ). *Body-Face* trials evoked a significantly larger peak response compared to *Face-Only* trials ( $p < .01$ ). Panel C displays the average HDR from the voxels displayed in the yellow color map. The HDR was independently sampled from a separate subset of trials that were used to define the functional region of interest. The average peak amplitude varied as a function of stimulus condition. Specifically, *Body-Face* trials evoked a larger peak amplitude response than both *Body-Only* and *Face-Only* conditions ( $p < .05$ ). There was no difference between *Body-Only* and *Face-Only* trials.

### 3.3. Ventral occipitotemporal cortex (VOTC)

A random-effects analysis revealed two significant contrasts in the VOTC. A medial ROI centered on the lingual gyri and depicted in light blue on Fig. 5 consists of voxels in which *Body-Only* trials evoked a higher amplitude response than either *Face-Only* or *Body-Face* trials. Thus, within this medial region, the largest response was evoked by bodies without faces. A repeated measures ANOVA conducted on the independently sampled time course data confirmed a difference in peak amplitude as a function of occlusion type,  $F_{(2,16)} = 24.15$ ,  $p < .001$ , and post hoc analyses showed that the average HDR evoked for *Body-Only* trials had a significantly larger peak amplitude compared to both *Body-Face*  $t_{(10)} = 3.49$ ,  $p < .01$ , and *Face-Only* trials,  $t_{(10)} = 6.49$ ,  $p < .001$ . *Body-Face* trials evoked a significantly larger peak amplitude when compared independently to the response evoked by *Face-Only* trials,  $t_{(10)} = 3.74$ ,  $p < .01$ .

The yellow color map plotted on VOTC in Fig. 5 shows voxels for which a higher peak amplitude response was evoked

for *Body-Face* trials when compared to both *Body-Only* and *Face-Only* trials. A repeated measures ANOVA conducted on the independently sampled time course data confirmed a difference in peak amplitude as a function of condition,  $F_{(2,15)} = 4.86$ ,  $p < .05$ . Specifically, the average HDR evoked for *Body-Face* trials was significantly larger in peak amplitude compared to both *Body-Only*  $t_{(10)} = 3.87$ ,  $p < .01$ , and *Face-Only* trials,  $t_{(10)} = 2.54$ ,  $p < .05$ . There was no difference between *Body-Only* and *Face-Only* trials. The average evoked HDRs from the time course analysis are plotted in Fig. 5 for both contrast types identified in the VOTC.

## 4. Discussion

Our results demonstrate differences in activation evoked by the appearance of partially occluded or fully visible human actors in both the LOTC and VOTC – regions known to be critically involved in social perception. Within each of these broad regions, two closely adjacent ROIs were identified that

responded differently to face and body occlusion. Within the LOTC, these ROIs encompassed the EBA and the pSTS. Within the VOTC, these ROIs encompassed the lingual gyri and the fusiform gyri.

To briefly recapitulate our results, consistent with prior literature, poses in which the face was clearly visible (the *Body-Face* and *Face-Only* conditions) evoked strong activation in the lateral FFG of the VOTC and the pSTS of the LOTC. In the lateral FFG, activation was greater for the *Body-Face* than *Face-Only* condition, while no difference between *Body-Face* and *Face-Only* conditions were observed in the pSTS. The *Body-Only* condition strongly activated the lingual gyrus and medial FFG (the medial VOTC) and the EBA region of the LOTC. In both of these regions, the *Face-Only* condition evoked little or no activation. Of particular note, exposing the hidden face of the strongly activating *Body-Only* condition caused a *decrease* in activation in both the medial VOTC and the EBA. We will now review these results in more detail by region and then offer two explanations for what processes may underlie these results.

#### 4.1. Lateral occipitotemporal cortex (LOTC)

Activation in LOTC was strongly modulated by the presence or absence of a face in the visual scene. The MOG and surrounding regions that encompassed the reported locations of EBA were most strongly activated when the actor's torso and limbs were visible and weakest when only the face was visible, replicating the results of prior studies (Downing et al., 2001; Peelen & Downing, 2004). Of most interest, adding a face to the same strongly activating torso and limbs *decreased* the activation that was evoked by the torso and limbs alone suggesting that the face inhibited the response to the body parts.

In strong contrast to the EBA, the pSTS showed a significant preference for face stimuli, regardless of whether the face sat atop a body that was either occluded or visible. Neuroimaging studies of pSTS function have previously implicated this region in the processing of biological motion. Recent investigations have shown that this region might also be a critical contributor to the visual analysis of the intentions and actions of others (Blakemore, Sarfati, Bazin, & Decety, 2003; Castelli, Happe, Frith, & Frith, 2000; Decety & Grezes, 1999; Grezes, Frith, & Passingham, 2004; Morris et al., 2005; Pelphrey, Morris, & McCarthy, 2004a; Pelphrey, Singerman, Allison, & McCarthy, 2003; Pelphrey, Viola, & McCarthy, 2004b; Saxe & Kanwisher, 2003; Saxe, Xiao, Kovacs, Perrett, & Kanwisher, 2004), although biological motion may be necessary in conveying such intentions (Morris et al., 2005). Nevertheless, many reports of activity evoked by faces have included significant activation in the pSTS (Haxby, Hoffman, & Gobbini, 2000; Hoffman & Haxby, 2000; Kanwisher et al., 1997; McCarthy et al., 1997; Puce et al., 1995), which has suggested that this area is sensitive to dynamic aspects of the face, and in particular the eyes (Hoffman & Haxby, 2000; McCarthy, 1999; Pelphrey et al., 2003; Puce et al., 1998).

Our recent study of the somatotopic organization of the STS for the perception of biological motion is consistent with the results reported here (Pelphrey, Morris, Michelich, Allison, &

McCarthy, 2005). In that study, we compared activity elicited by movement of the eyes, mouth, or hand. While each motion evoked robust activation of the right LOTC, the spatial distribution of the hemodynamic responses differentiated the movements in an interesting manner. Movements of the face (i.e., eyes and mouth) activated discrete regions in the crux of the pSTS, while hand movements evoked activity in more inferior and posterior portions of the pSTS and, notably, a region of the LOTC consistent with the reported location of the EBA. Thus, consistent with our present findings, we found a preference for face movement in the pSTS and a preference for the movement of at least one body part in the more posterior and inferior regions of LOTC overlapping with the location of the EBA.

#### 4.2. Ventral occipitotemporal cortex (VOTC)

Like LOTC, activation in VOTC was strongly modulated by the presence or absence of a face in the visual scene. Indeed, the two foci of activation within VOTC showed differential activation patterns that resembled that of LOTC. The medial and posterior portions of the VOTC (encompassing the lingual gyrus and parts of medial FFG) responded most strongly for bodies without faces, while virtually no activation was evoked above prestimulus baseline for faces without bodies. Exposing the face of the strongly activating body stimulus decreased activation in the medial VOTC. This pattern was remarkably similar to that discussed above for the EBA, and the striking similarities of the responses can be appreciated by comparing the HDRs in Fig. 3B (EBA) and Fig. 5B (medial VOTC).

However, while the pattern was similar for the poses used in the present study, other studies from our lab (Allison, McCarthy, Nobre, Puce, & Belger, 1994b; Puce et al., 1995, 1996) have found that scrambled faces and complex textures also activate medial VOTC regions while faces activated more lateral portions of FFG. Kanwisher and colleagues (Kanwisher et al., 1997) observed that objects activated medial/posterior portions of VOTC, while face activations were localized more anterior and lateral in the FFG. That the human body with an occluded face activated medial VOTC areas previously identified with differential processing of objects and textures suggests that domain non-specific processes may have been evoked – i.e., bodies without faces activated similarly as objects and textures. This conclusion raises the question of why adding a face to the strongly activating faceless body resulted in a decrease in activation – the same result obtained for EBA above. The presence of the occluding textbook may have enhanced this medial activation, although the book was also present and visible (albeit, not as prominently) in the *Body-Face* poses.

At more lateral sites in the FFG, the activation pattern by both poses that included faces was greatly enhanced while activity evoked by the *Body-Only* pose decreased at this same region. This shift can be appreciated by comparing Fig. 5B and C. The HDR evoked by the *Body-Only* pose decreased by a factor of two moving from medial to lateral, while the HDR evoked by the face with the body occluded went from no response at medial locations to a robust response at lateral locations. That is, the response to the *Body-Only* face-occluded pose declined contin-

uously over the medial to lateral VOTC axis, while the response to the face increased continuously over the same axis.

This change in the response pattern along the axis from the medial to lateral VOTC is similar to the change in response pattern along the axis from EBA to pSTS in the LOTC with one exception. At pSTS, the two poses with visible faces evoked a HDR with equal amplitudes. At the lateral FFG, the response to the body and face together was larger than the response to the face alone.

These results could be explained if there were limited neural resources devoted to processing the complex social scene and the presence of a face drew resources away from the processing of other objects in the scene. Activity in regions specialized for processing non-face body parts might then be passively reduced. This outcome would be further potentiated if the viewer's scan paths were systematically biased towards the face. Hershler and Hochstein have recently demonstrated a "pop-out" effect for faces, which suggests that faces might be particularly salient stimuli and thus likely to attract processing resources (Hershler & Hochstein, 2005).

This explanation posits passive reduction of activity in brain regions that process the less salient stimuli in the scene. However, it is also possible that an active inhibitory process is engaged such that the region processing the more salient stimulus (e.g., the face) inhibits regions associated with processing the less salient stimulus (e.g., the body). There is precedence for this explanation. Pelphrey et al. (2003) suggested such a mechanism in their study that demonstrated that faces evoked a relative deactivation in brain regions located medially to face-specific regions of lateral FFG. Likewise, Allison et al. (2002) presented electrophysiological evidence for category-specific inhibitory interactions. They identified focal N200 potentials indicative of word form and face specific sites using subdural recordings made from the FFG and adjacent cortex. At half of these sites, the non-preferred stimulus category evoked a P200 potential which Allison et al. interpreted as evidence for inhibition of category-specific neurons, and for which they provided a model of synaptic connectivity between neurons selectively activated by faces and letterstrings to account for their results. Finally, Bentin, Allison, Puce, Perez, and McCarthy (1996) recorded event-related potentials from the scalp while subjects viewed whole faces and isolated face parts. They observed a large N170 component evoked by eyes presented in isolation that was diminished in amplitude when those same eyes were embedded within whole faces (Bentin et al., 1996).

Explanations based upon passive reduction or active inhibition of processing make similar predictions that cannot be discriminated on the basis of the relative measure of activation provided by fMRI alone. However, both explanations raise a general issue about domain specificity. If there are regions, such as EBA, for specialized processing of non-face body parts, why should processing in this region be either resource limited or actively inhibited by the presence of a face when a separate face specific region is available for processing? We speculate that initial processing of domain specific stimuli proceeds in parallel but that higher order processing of the scene requires selection among stimuli relevant to the subject's dispositions and goals.

In conclusion, we demonstrated that a visible face strongly modulated activation evoked by the appearance of a partially occluded human body during free viewing of a natural social scene. Within MOG/EBA and pSTS we observed an inferior/posterior to superior/anterior change in activity such that bodies without faces more strongly activated the MOG/EBA while faces more strongly activated the pSTS. Our hypothesis that the EBA would show a greater response when more body parts were visible was not supported by the current data. Indeed the presence of a complete body with a face evoked less activity than the presence of the same body without a face. Our initial hypothesis that activation in the lateral FFG of the VOTC would be potentiated by the presence of a viewable face was partially upheld, but we noted that the strongest response in this region was evoked when both the face and body were visible.

### Acknowledgements

We wish to thank M.B. Nebel and C. Michelich for assistance with stimulus development and data analysis, and E.J. Carter and Y. Peters for comments on the manuscript. We are also grateful for the comments and suggestions from two anonymous reviewers. This research was supported by NIH grant MH-05286. Dr. Morris was supported by F32 MH073367. Dr. Pelphrey was supported by K01 MH071284. Dr. McCarthy is a DVA Senior Research Career Scientist and was partially supported by the Department of Veterans Affairs Mental Illness Research, Education, and Clinical Center (MIRECC).

### References

- Allison, T., Ginter, H., McCarthy, G., Nobre, A. C., Puce, A., Luby, M., et al. (1994). Face recognition in human extrastriate cortex. *Journal of Neurophysiology*, *71*(2), 821–825.
- Allison, T., McCarthy, G., Nobre, A., Puce, A., & Belger, A. (1994). Human extrastriate visual cortex and the perception of faces, words, numbers, and colors. *Cerebral Cortex*, *4*(5), 544–554.
- Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the sts region. *Trends in Cognitive Sciences*, *4*(7), 267–278.
- 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*(6), 551–565.
- Bentin, S., & Golland, Y. (2002). Meaningful processing of meaningless stimuli: the influence of perceptual experience on early visual processing of faces. *Cognition*, *86*(1), B1–B14.
- Bentin, S., Sagiv, N., Mecklinger, A., Friederici, A., & von, C. Y. (2002). Priming visual face-processing mechanisms: Electrophysiological evidence. *Psychological Science*, *13*(2), 190–193.
- Blakemore, S. J., Sarfati, Y., Bazin, N., & Decety, J. (2003). The detection of intentional contingencies in simple animations in patients with delusions of persecution. *Psychological Medicine*, *33*(8), 1433–1441.
- Castelli, F., Happe, F., Frith, U., & Frith, C. (2000). Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *Neuroimage*, *12*, 314–325.
- Cox, D., Meyers, E., & Sinha, P. (2004). Contextually evoked object-specific responses in human visual cortex. *Science*, *304*(5667), 115–117.
- de Gelder, B., Snyder, J., Greve, D., Gerard, G., & Hadjikhani, N. (2004). Fear fosters flight: a mechanism for fear contagion when perceiving emo-



- tion expressed by a whole body. *Proceedings of the National Academy of Sciences of the USA*, 101(47), 16701–16706.
- Decety, J., & Grezes, J. (1999). Neural mechanisms subserving the perception of human actions. *Trends in Cognitive Science*, 3(5), 172–178.
- Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, 293(5539), 2470–2473.
- Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage*, 15(4), 870–878.
- Glover, G. H., & Law, C. S. (2001). Spiral-in/out bold fMRI for increased snr and reduced susceptibility artifacts. *Magnetic Resonance in Medicine*, 46(3), 515–522.
- Grezes, J., Frith, C. D., & Passingham, R. E. (2004). Inferring false beliefs from the actions of oneself and others: an fMRI study. *Neuroimage*, 21(2), 744–750.
- Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience*, 7(5), 555–562.
- Guo, H., & Song, A. W. (2003). Single-shot spiral image acquisition with embedded z-shimming for susceptibility signal recovery. *Journal of Magnetic Resonance Imaging*, 18(3), 389–395.
- Hadjikhani, N., & de Gelder, B. (2003). Seeing fearful body expressions activates the fusiform cortex and amygdala. *Current Biology*, 13(24), 2201–2205.
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Science*, 4(6), 223–233.
- Hershler, O., & Hochstein, S. (2005). At first sight: a high-level pop out effect for faces. *Vision Research*, 45(13), 1707–1724.
- Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience*, 3(1), 80–84.
- 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(11), 4302–4311.
- Lerner, Y., Hendler, T., & Malach, R. (2002). Object-completion effects in the human lateral occipital complex. *Cereb Cortex*, 12(2), 163–177.
- McCarthy, G. (1999). Physiological studies of face processing in humans. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (2nd ed., pp. 393–410). Cambridge: MIT Press.
- 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.
- McCarthy, G., Puce, A., Gore, J. C., & Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience*, 9(5), 605–610.
- Morris, J. P., Pelphrey, K. A., & McCarthy, G. (2005). Regional brain activation evoked when approaching a virtual human on a virtual walk. *Journal of Cognitive Neuroscience*, 17(11), 1744–1752.
- Peelen, M. V., & Downing, P. E. (2004). Selectivity for the human body in the fusiform gyrus. *Journal of Neurophysiology*.
- Pelphrey, K. A., Morris, J. P., & McCarthy, G. (2004). Grasping the intentions of others: The perceived intentionality of an action influences activity in the superior temporal sulcus during social perception. *Journal of Cognitive Neuroscience*, 16(10), 1706–1716.
- Pelphrey, K. A., Morris, J. P., Michelich, C. R., Allison, T., & McCarthy, G. (2005). Functional anatomy of biological motion perception in posterior temporal cortex: an fMRI study of eye, mouth and hand movements. *Cerebral Cortex*, 15(12), 1866–1876.
- Pelphrey, K. A., Singerman, J. D., Allison, T., & McCarthy, G. (2003). Brain activation evoked by perception of gaze shifts: the influence of context. *Neuropsychologia*, 41(2), 156–170.
- Pelphrey, K. A., Viola, R. J., & McCarthy, G. (2004). When strangers pass: processing of mutual and averted social gaze in the superior temporal sulcus. *Psychological Science*, 15(9), 598–603.
- Puce, A., Allison, T., Asgari, M., Gore, J. C., & McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letter strings, and textures: A functional magnetic resonance imaging study. *Journal of Neuroscience*, 16(16), 5205–5215.
- Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *Journal of Neuroscience*, 18(6), 2188–2199.
- Puce, A., Allison, T., Gore, J. C., & McCarthy, G. (1995). Face-sensitive regions in human extrastriate cortex studied by functional MRI. *Journal of Neurophysiology*, 74(3), 1192–1199.
- 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, 445–448.
- Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in “theory of mind”. *Neuroimage*, 19(4), 1835–1842.
- Saxe, R., Xiao, D. K., Kovacs, G., Perrett, D. I., & Kanwisher, N. (2004). A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia*, 42(11), 1435–1446.