PERGAMON

Neuropsychologia 41 (2003) 156–170

NEUROPSYCHOLOGIA

www.elsevier.com/locate/neuropsychologia

Brain activation evoked by perception of gaze shifts: the influence of context

Kevin A. Pelphrey [a,c], Jeffrey D. Singerman [a], Truett Allison [d], Gregory McCarthy [a,b,*]

[a] *Duke-UNC Brain Imaging and Analysis Center, Duke University Medical Center, 163 Bell Building, Box 3918, Durham, NC 27710, USA*
[b] *Department of Veterans Affairs Medical Center, Durham, NC 27710, USA*
[c] *Neurodevelopmental Disorders Research Center, University of North Carolina at Chapel Hill School of Medicine, Chapel Hill, NC 27710, USA*
[d] *Department of Neurology, Yale University School of Medicine, New Haven, CT, USA*

Abstract

Prior studies from our laboratory [Journal of Neuroscience 18 (1998) 2188; Cognitive Neuropsychology 17 (2000) 221] have demonstrated that discrete regions of the superior temporal sulcus (STS) are activated when a subject views a face in which the eyes shift their gaze. Here we investigated the degree to which activity in the STS and other brain regions is modulated by the context of the perceived gaze shift; that is, when the shift correctly or incorrectly acquires a visual target. Fifteen subjects participated in an event-related functional magnetic resonance imaging experiment in which they viewed an animated face that remained present throughout each run. On each of 21 trials within each run, a small checkerboard appeared and flickered at one of six locations within the character's visual field. On "correct" trials, the character shifted its gaze towards the checkerboard after a delay of 1 or 3 s. On "incorrect" trials, the character shifted its gaze towards empty space after the same delays. On "no shift" trials, the character's eyes did not move. Significantly larger hemodynamic responses (HDR) were evoked by gaze shifts compared to no gaze shifts in primarily right hemisphere STS. The gaze-evoked HDR was significantly delayed in peak amplitude for 3 s compared to 1 s shifts. For 1 s shifts, a strong effect of context was observed in which errors evoked a HDR with extended duration. Although this study focused upon STS, similar effects were also observed in the intraparietal sulcus and fusiform gyrus.
Published by Elsevier Science Ltd.

Keywords: fMRI; Superior temporal sulcus; Social perception

1. Introduction

From the earliest stages of postnatal development, faces are salient to typically developing individuals [20,37]. Faces derive their significance, in part, from the wealth of social information they provide. This information includes the bearer's identity [9], emotional state [6,16], intentions [4,5], and focus of attention [41,42]. The capacity to extract socially relevant information from faces is fundamental to normal reciprocal social interactions and interpersonal communication. Of the core internal facial features (i.e. eyes, nose, and mouth), the eyes are thought to provide the most critical information and preferentially draw a viewer's attention [17,42]. Adult viewers devote 70% or more of their fixations to the eyes [43,49,65]. This pattern of face scanning emerges as early as the second month of postnatal life [25,45], and is disturbed in schizophrenia [53] and autism [50]. Information regarding direction of gaze is thought to be particularly important in guiding social interactions [4,40]. Gaze can provide information concerning the mental states of others, facilitate social control, regulate turn taking, direct attention, and communicate intimacy [3,17,40]. Sensitivity to gaze direction emerges early in ontogeny. For example, infants detect direction of perceived gaze, and modulate their own attention accordingly [18,32,60,63].

Recent neurofunctional models of the human face processing system distinguish cortical regions involved in processing invariant (i.e. those that carry information about identity) characteristics of faces from those regions involved in processing dynamic (i.e. those that facilitate communication) aspects of faces [27,28,46,55]. For example, McCarthy [46] identified four nodes of the human face processing system. Two of these nodes, the lateral posterior fusiform gyrus (FFG) and anterior ventral temporal cortex, are involved, respectively, in structural encoding and face memory. A third node, centered in the superior temporal sulcus (STS), is involved in the analysis of face motion, such as eye and mouth

* Corresponding author. Fax: +1-919-681-7033.
E-mail address: gregory.mccarthy@duke.edu (G. McCarthy).

0028-3932/02/$ – see front matter. Published by Elsevier Science Ltd.
PII: S0028-3932(02)00146-X

movements. The remaining node, located in the amygdala, is involved in the analysis of facial expression.

Allison et al. [1] used the term "STS region" to refer to cortex within the STS, to adjacent cortex on the surface of the superior temporal gyrus and middle temporal gyrus (near the straight segment of the STS), and to adjoining cortex on the surface of the angular gyrus (near the ascending limb of the STS). Several sources of evidence have converged to indicate that the STS region is involved in the perception of gaze direction. This role was suggested initially by experimental studies of nonhuman primates [11,26,30,51,52,67] and neuropsychological studies of human lesion patients [11]. More recently, functional neuroimaging and electrophysiology studies have started to enhance our knowledge of the STS region's involvement in processing gaze direction.

Using functional magnetic resonance imaging (fMRI), Puce et al. [55] first identified a bilateral region of activation centered in the posterior STS in response to observed eye or mouth movements, but not in response to an inwardly moving radial pattern presented to control for effects related to movement per se. With event-related potential (ERP) recordings, Bentin et al. [7] demonstrated that an N170 ERP recorded from scalp electrodes overlying the STS was larger when evoked by isolated eyes than by whole faces or other face components, and Puce et al. [56] demonstrated that the N170 ERP was larger in response to the movement of eyes averting their gaze away from the viewer than to eyes returning to gaze at the observer. In a positron emission tomography (PET) study, Wicker et al. [66] identified several regions of activation in response to mutual and averted gaze including portions of the STS. Finally, using fMRI, Hoffman and Haxby [31] demonstrated that attention to gaze elicited a stronger response in STS than did attention to identity. Note that some of these studies used static stimuli that varied in direction of gaze [10,31,39,66] while others used dynamic stimuli in which the eyes moved [55,56].

Research concerning the role of the STS region in processing eye movements is fundamental to our understanding of the neurofunctional organization of the human face processing system. However, this line of inquiry is equally significant for its potential to provide information about the neuroanatomical systems underlying social perception and social cognition [8,19]. Allison et al. [1] defined social perception as the initial stages of evaluating the intentions of others by analysis of gaze direction, body movement, and other types of biological motion, and stressed the role of the STS region in a larger social perception system involved in processing the emotional value and social significance of biological stimuli.

Baron-Cohen [4] has defined a four-component neuropsychological model of a "mindreading" system, whereby we acquire information from the face of another person during shared attention, use this information to attribute a mental state to the person, and then predict that individual's behavior from his or her inferred mental state. In Baron-Cohen's model, the "intentionality detector" (ID) detects self-propelled moving stimuli and allows us to interpret their movement in terms of simple volitional mental states (e.g. goals and desires). An "eye-direction detector" (EDD) perceives the presence of eyes and the direction of their gaze and attributes the mental state of seeing to the owner of those eyes. These two components are linked together by a "shared attention module" (SAM), which supports the identification of occasions when the self and another agent are attending to the same stimulus. Lastly, by integrating data from the three previous components, the "theory-of-mind mechanism" (ToMM) provides the ability to examine information gathered from another individual during shared attention, allows us to ascribe a mental state to that individual, and then permits us to explain or predict that individual's behavior in terms of the inferred mental state. Information concerning the role of the STS region in processing eye movements is particularly significant for Baron-Cohen's model, because two components of the mindreading system, the EDD and the SAM rely on normal gaze shift perception, and the ToMM, in turn, relies upon data from these two components.

In prior functional neuroimaging studies concerned with gaze perception, the stimulus face gazed towards empty space [31,55,66]. Thus, it is not clear whether the identified brain regions participated merely in simple gaze detection or in a more complex analysis related to the context in which the gaze shift occurred. Here, by providing a target for the gaze shift, we investigated whether regions activated by the perception of gaze are modulated by the context of the observed gaze shift. Participants observed an animated female character as visual target appeared within the character's visual field at regular intervals. The character either made no gaze shift to the target, shifted gaze to the target with a 1 or 3 s delay, or shifted gaze to an empty location of space with the same delay. This allowed us to determine whether activity within the face processing system is influenced by the perceived intention or goal of the action, and whether a gaze shift toward an object produces a different pattern of activity than that of an identical gaze toward empty space. That is, we were interested in determining if elements of the face processing system are sensitive to the social relevance of a biological motion—whether the action is intentional and goal-directed within the established context.

In addition to investigating context-dependent activity within the STS, we also wished to examine the effects of perceived gaze in other brain regions. In a similarly designed pilot study with eight subjects, we observed, in addition to STS activity, activations related to gaze perception in the intraparietal sulcus (IPS) and FFG. However, we employed a constant 1 s delay between the appearance of the target and the gaze shift, and so it was uncertain whether the activity was related to processing the visual target or the gaze shift. We predicted that varying the delay interval between target and gaze would result in a systematic change in the latency to peak amplitude of the gaze-related hemodynamic response (HDR).

Finally, we were concerned that any observed differences in activity might not result from true differences in gaze processing, but rather from differences in the way that participants viewed the stimuli. For example, participants might move their own eyes more in one condition than another, and this differential eye movement might be related to a participant's experience with the task. We therefore conducted a parallel study outside of the scanner in which we recorded the visual scanpaths of naïve and experienced volunteers in response to the stimuli used in the current fMRI study. The point-of-regard (POR) recordings allowed us to address these two potential confounds.

2. Methods

2.1. Participants

Fifteen healthy right-handed volunteers participated in the fMRI experiment (8 male, 7 female; age range 19–30 years; mean age 24.2 years). Eight healthy right-handed volunteers (4 females; age range 22–33; average age 27.1 years) participated in the eye movement monitoring study. Of the eight subjects in this latter study, four also participated in the fMRI experiment (2 females). All participants had normal or corrected to normal visual acuity. The studies were approved by the Institutional Review Boards of the Duke University Medical Center (fMRI) and the University of North Carolina at Chapel Hill School of Medicine (eye movement monitoring). Subjects provided written informed consent, and they were paid for participating.

2.2. Experimental stimuli

2.2.1. fMRI stimuli

An animated female character was created using the Poser 4.0® software program (Curious Labs Inc., Santa Cruz, CA). This character was presented throughout the duration of each run, displayed from the shoulders up with eyes forward (see Fig. 1). On each of the 21 trials within each run, a checkerboard-patterned box appeared (S1) and flickered for 5 s in one of six positions within the character's field of view (i.e. above each shoulder at eye level, above eye level, or below eye level). The character's response (S2) to the checkerboard stimulus distinguished the five stimulus conditions. In two conditions, the character shifted her gaze directly toward the checkerboard with a 1- or 3-s stimulus onset asynchrony (SOA) (correct 1 s, correct 3 s). In two other conditions, the character moved her eyes towards one of six empty locations in space with a 1- or 3-s SOA (incorrect 1 s, incorrect 3 s). In the final condition, the character's eyes remained motionless (no shift). When the eyes moved, they returned to the original forward position within 500 ms of S1 offset. Trials were separated by a 21-s onset-to-onset interval during which the character alone was presented with eyes forward.

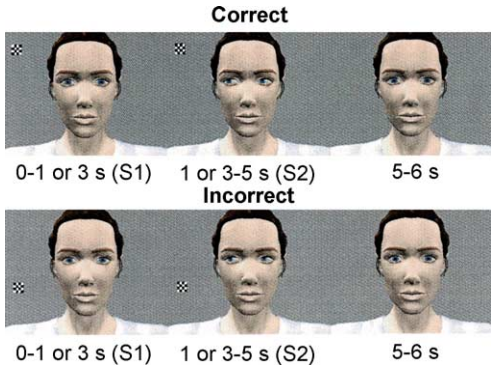

Fig. 1. Trials consisted of five stimulus types and began when a small checkerboard appeared and flickered for 5 s in the character's FOV. In two conditions, the character gazed to the checkerboard with a 1- or 3-s SOA. In two other conditions, the character gazed towards one of five empty locations in space with a 1- or 3-s SOA. In the final condition, the character's eyes did not move.

The CIGAL [64] program was used to control stimulus presentation. Stimuli were presented using an LCD projector (XGA resolution, 900 lumens) that projected images upon a translucent rear projection screen ($\sim$ 56 cm $\times$ 66 cm) placed at the subject's feet. Participants viewed the stimuli through custom glasses with angled mirrors. The stimuli subtended a horizontal visual angle of approximately 17° and a vertical visual angle of approximately 20°. Participants were instructed to attend to the screen at all times, but otherwise were allowed to visually scan the stimulus presentation in any way they wished. Trials were randomized within runs lasting 7.5 min ($\sim$21 trials per run). Participants completed an average of 9.4 presentation runs (198 trials) per subject.

2.2.2. Eye movement monitoring stimuli

The stimulus presentation for the eye movement monitoring study was an abridged version of the presentation used in the fMRI study. Three trials of each stimulus condition (15 trials total) were presented in random order over a single 3-min run, which all participants completed. Trials were separated by a 6-s interval, during which the animated character was presented with her eyes facing forward. The stimulus presentation software and participant instructions were identical to those used in the fMRI study.

2.3. Eye movement data acquisition and analysis

An ISCAN remote infrared pupil–corneal reflection eye movement monitoring system (ISCAN Inc., Cambridge, Massachusetts, USA) was used to record POR data at 60 Hz. Each 6-s data epoch was checked for tracking integrity. Eye blinks and off-screen gazes were excluded. The epochs were averaged across subjects and grouped by stimulus condition. For each subject, the mean length of the total scanpath was calculated for each experimental condition.

2.4. fMRI data acquisition

MRI scanning was performed on a General Electric 4T LX NVi scanner system equipped with 41 mT/m gradients, and using a birdcage radio frequency (RF) head coil for transmit and receive (General Electric, Milwaukee, Wisconsin, USA). Sagittal T1-weighted localizer images were first acquired and used to define a target volume for a semi-automated high-order shimming program. After shimming, the anterior commissure (AC) and posterior commissure (PC) were identified in the mid-sagittal slice and used as landmarks for the prescription of blood oxygen level dependent (BOLD) contrast images. A series of 60 high-resolution coronal T1-weighted images (repetition time (TR) = 450 ms; echo time (TE) = 20 ms; field of view (FOV) = 24 cm; image matrix = 256^2; slice thickness = 5 mm, in-plane resolution = 0.9375 mm^2) was acquired along the AC–PC line. The T1-weighted images were used to select 20 contiguous 5 mm coronal slices for functional imaging. These slices were acquired from posterior to anterior along the intercommissural line such that the 20th slice was anchored at the AC. Functional images were collected using a spiral imaging sequence sensitive to BOLD contrast (TR = 1.5 s; TE = 30 ms; FOV = 24 cm; image matrix = 64^2; flip angle = 62°; slice thickness = 5 mm; in-plane resolution = 3.75 mm). Each imaging run began with five discarded RF excitations to allow for steady-state equilibrium.

2.5. fMRI data analysis

The center of mass for each functional image volume within each time series was computed and plotted for each participant and imaging run. No volunteer had greater than a 3-mm deviation in the center of mass in the *x*-, *y*-, or *z*-dimensions. The center of mass procedure was supplemented by visual inspection of each functional image series in a cine loop. The MR signal for each voxel was temporally aligned to correct for the interleaving of slice acquisition within each 1.5 s TR by fitting the time series of each voxel with a cubic spline and then resampling this function for all voxels at the onset of each TR. Epochs, time-locked to the onset of S1, were extracted from the continuous time series and averaged according to trial type. The temporal order relative to S1 was maintained throughout averaging. Averaged epochs consisted of the two-image volumes before (−3 to −1.5 s) and the 11-image volumes after (1.5–16.5 s) the onset of each S1 event, for a total of 14 image volumes or 21 s. The baseline was subtracted from the fMRI data on a voxel-by-voxel basis. The averaged MR signal time-epochs were used in the analytic procedures described further.

2.5.1. Regions of interest (ROI)

A priori ROI selection was based on pertinent literature [31,36,55,66] and results from pilot research involving a similar gaze shift paradigm. ROIs were drawn on each participant's high-resolution anatomical images by two trained research assistants who were blind to the other data analytic procedures. ROI were traced on the left and right fusiform gyri (FFG), intraparietal sulci (IPS), superior temporal sulci (STS), and Sylvian fissures and surrounding insular cortex (SYF/INS). The STS was further subdivided into four anatomical components: the straight segment divided into the anterior half of the straight segment (STSant) and the posterior half of the straight segment (STSpost), the ascending limb of the straight segment (STSal), and the posterior continuing branch (STSpc). Identification of anatomical landmarks and ROI was guided by human brain atlases [15,44,59,62]. For each of the 14 (seven anatomical areas by two hemispheres) ROI, labels indicated the distance (in mm) from the AC, facilitating the creation of summary activation waveforms across subjects. ROI for the STSant were drawn on five slices ranging from 0 to 20 mm posterior from the AC. The STSpost was outlined on five slices ranging from 25 to 45 mm posterior from the AC. The STSal was drawn on four slices from 45 to 60 mm from the AC. The STSpc was outlined on five slices ranging from 45 to 65 mm posterior from the AC. The IPS was outlined on 12 slices ranging 3–90 mm posterior from the AC. The SF/INS was outlined on 12 slices 0–55 mm from the AC. The FFG was outlined on eight slices 40–75 mm posterior from the AC.

The average signal for all voxels within each ROI was computed for each of the 14 time points and plotted to visualize the time course of the hemodynamic response (HDR) for each ROI during each stimulus condition. This analysis was not dependent upon choice of waveform or the results of any correlation analysis. The HDR time course was examined separately for each slice and hemisphere within each ROI. ROI were also used to group and count activated voxels that were identified in a correlation analysis as described further.

2.5.2. Correlation with a reference waveform

Many voxels within an anatomically defined ROI may not be activated by the stimulus presentation. Thus, the mean activity measured within the ROI may represent a dilution of the activity in a subset of voxels. To address this issue, we conducted a correlation analysis with an empirically defined reference waveform to identify subsets of gaze shift activated voxels. The template waveform was the mean waveform representing the average HDR time course within the right hemisphere of the STSpost and STSal (30–45 mm from the AC) across the four experimental conditions involving observed eye movements (all shifts). This reference waveform is shown in the inset of Fig. 2B. The *t*-statistics were calculated from the correlation coefficients, and activated voxels were defined as those that significantly cross-correlated with the reference waveform, with the threshold for activation set at $t \geq 1.96$ ($P < 0.05$, uncorrected). By calculating the cross-correlation of voxels with the reference waveform, we identified sets of voxels for each subject that showed significant activation ($t \geq 1.96$) to eye–gaze shifts. These voxel clusters were superimposed on the subjects' anatomical images for visual inspection.

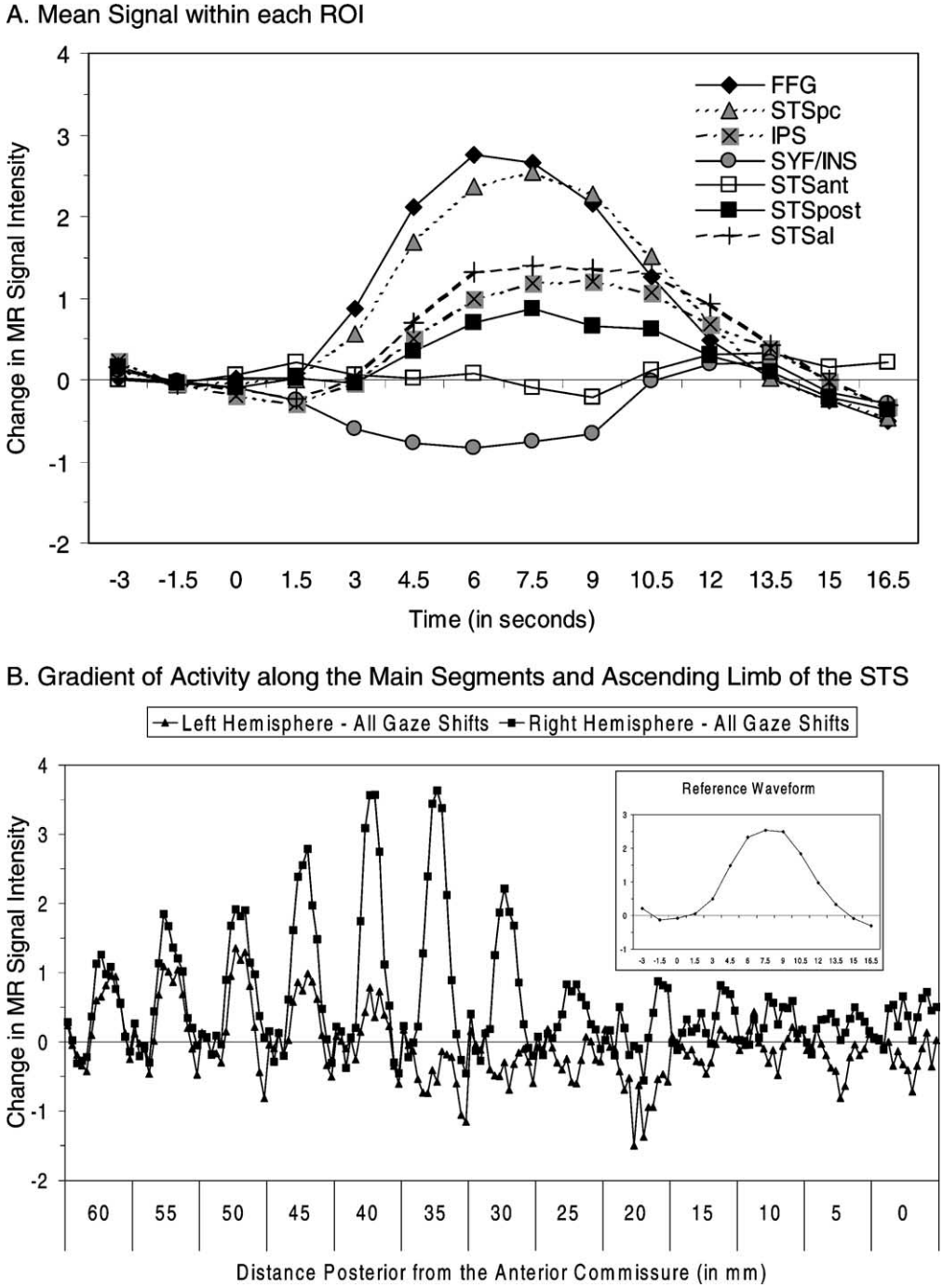

Fig. 2. (A) Mean HDR waveform from each ROI. Waveforms represent the contribution of all voxels within the ROI averaged across the four conditions involving an observed eye–gaze. Significant positive HDRs were observed 6–9 s following stimulus onset in all but STSant and SYF/INS. A negative HDR waveform was observed in SYF/INS. (B) Activity-time waveforms for all voxels within each slice of the main segments and ascending limb of the STS are plotted on a slice-by-slice and hemisphere basis to demonstrate that there was a strong anterior-to-posterior gradient in activity.

Counts of activated voxels, so defined, within each ROI, were converted to percentages relative to the number of voxels in that ROI. The percentages were analyzed with repeated-measures ANOVAs to determine if the percentage of activated voxels differed by ROI, hemisphere, and/or stimulus condition.

2.5.3. Time–activation waveforms for activated voxels

Time–activation waveforms were computed for the activated voxels from each ROI. Each participant contributed 14 waveforms (seven anatomical regions by two hemispheres) for each of the five stimulus conditions (70 total). Repeated measures ANOVAs were performed to evaluate differences

in HDR amplitude as a function of gaze shift context at each time point in the time–activation waveforms for selected ROI.

We measured the peak amplitude and latency to peak amplitude of the HDR from 3 to 12 s. For these measurements, the time–activation waveforms for each participant, hemisphere, ROI, and stimulus condition were fit with a cubic spline. This spline function was resampled at a 0.25-s resolution. Shifts in latency to peak amplitude as a function of gaze shift timing were evaluated with repeated-measures ANOVAs.

2.5.4. Voxel-based analyses

The analysis procedures described in the preceding sections involved an anatomical ROI approach. This class of strategy offers several advantages. For example, ROI analyses facilitate detailed inspection of the HDR time courses from ROI, respect individual differences in brain anatomy, and afford a high degree of certainty about localization of activity. However, unpredicted regions of activation not chosen as ROI and unexpected patterns of activity might be overlooked in an ROI analysis.

To this end, we performed voxel-based analyses on group-averaged data. Across-subjects functional time course volumes and *t*-statistic activation maps were computed for each stimulus condition, combining data from all 15 participants. Before averaging, the images were spatially normalized (i.e. translated, stretched, and rotated) to a template image set from a representative subject. Alignment factors for the functional images were calculated on a slice-by-slice basis using custom software. The normalization algorithm used the high-resolution anatomical images without regard to the functional data. Before normalization, the brain was extracted from each subject's anatomical images to eliminate the influence of high contrast but extraneous regions such as the skull and neck.

The averaged and spatially normalized data were used to identify and interrogate unexpected regions of group-consistent positive or negative activation. The group-averaged data were also used to compare the pattern of activation observed in response to observed shifts in gaze to the pattern observed when only the checkerboard pattern was present. We also searched for regions that were differentially activated by the context of gaze shifts. That is, we computed a *t*-statistic for each voxel, comparing correct and incorrect separately for 1 and 3 s and the all shifts to the no shift conditions. This analysis was performed for each time point.

3. Results

3.1. Eye movement results

The results of a 2 (experience) × 5 (stimulus condition) mixed ANOVA confirmed that the average amounts of eye movements did not differ by stimulus condition or experience with the stimuli. The interaction between these two factors was not significant. These results indicate that differences in functional activation by stimulus conditions cannot be explained by disparities in the total amount of eye movements made by the subjects during stimulus viewing. The absence of a difference in the average total scanpath length as a function of task experience suggests that the subjects did not systematically change viewing habits over time.

3.2. fMRI results

3.2.1. Average activation time courses for all voxels in ROI

The average HDR time courses for the voxels comprising each ROI are illustrated in Fig. 2A. These waveforms represent the contribution of all voxels within the ROI across the four conditions involving an observed eye–gaze shift. Thus, they were uninfluenced by statistical criteria or correlation analyses. The mean MR signal value of the prestimulus epoch across subjects was 1279. Across subjects and voxels, the standard error of the prestimulus baseline was 0.25. Significant positive HDRs were observed 6–9 s following stimulus onset in all but the STSant and SYF/INS. A negative HDR waveform was observed in SYF/INS. Note, in each ROI, HDRs were likely evoked by S1 and S2. The amplitude of the waveforms varied by ROI. Although each ROI produced a HDR with the same general time course, the HDRs from the FFG and STSpc were larger than the HDRs from the STSpost, STSal, and IPS.

3.2.2. Anterior-to-posterior gradients in activity

Activity within each ROI (across the four conditions involving a gaze shift) was examined on a slice-by-slice and hemisphere basis to evaluate anterior-to-posterior gradients in activity and to determine if these gradients differed by hemisphere. A striking anterior-to-posterior gradient in activity was observed along the STSant, STSpost, and STSal. The HDR waveforms for all voxels within each slice of these three components of the STS complex are shown in Fig. 2B by hemisphere. There was a strong anterior-to-posterior gradient in activity such that the highest levels of activity (in terms of HDR waveform amplitudes) occurred in and around the crux of the STS complex where the STSpost becomes the STSal (30–50 mm posterior from the AC). Smaller amplitudes were observed anterior (primarily in the STSant, 0–25 mm posterior from the AC) and posterior (mainly in superior portions of the STSal, 55–60 mm posterior from the AC) to these slices. There was also a strong right hemisphere bias in the slices with the highest amplitudes. Within the STSpc, there was no activity gradient. However, across the STSpc, peak HDR amplitudes were higher in the right hemisphere. In the FFG, no gradient was observed, but amplitudes were marginally higher in the right hemisphere for three posterior slices of the FFG (60–70 mm from the AC). In the IPS, a gradient of activity was observed—amplitudes were greater in six posterior slices (65–90 mm posterior from

the AC). Furthermore, amplitudes were greater in the right hemisphere for four posterior slices of the IPS (75–90 mm from the AC). In the SYF/INS, an activation gradient was not observed. However, HDR peak amplitudes were greater in the right SYF/INS.

These findings led us to subdivide the IPS into anterior (IPSant (35–60 mm posterior from the AC)) and posterior (IPSpost (65–90 mm posterior from the AC)) segments. In subsequent analyses, these subdivisions supplemented the original ROI. Moreover, because it was clear that activations in the STSant did not exceed levels above noise, this region will not be discussed further.

3.2.3. Correlation with a reference waveform

Before evaluating the influences of our experimental manipulations on the HDR time courses, we calculated the correlation between the empirically defined reference waveform (see inset of Fig. 2B) and the time activation waveforms of each voxel. Activated voxels were those whose time–activity waveforms significantly correlated with the reference waveform ($t \geq 1.96$, $P \leq 0.05$, uncorrected). By using a relatively low threshold for statistical significance, we purposely defined as activated some voxels whose actual waveform shapes could vary considerably from the reference waveform shape.

Fig. 3 presents the average percentage of activated voxels (relative to the total number of voxels) in ROI, by hemisphere. These averages were generally moderate ($M = 29\%$) but varied widely by ROI from a low of 4% for the left SYF/INS to a high of 48% for the right STSpc. We examined the spatial distribution of the activated voxels within each ROI by superimposing them on anatomical images. Activated voxels were generally clustered together within each ROI.

Percentages of activated voxels were submitted to repeated measures ANOVAs to evaluate potential hemisphere effects in each ROI. For this analysis and subsequent analyses, the STSpost and STSal were grouped together (STSpost + al). More voxels were activated in the right hemisphere than in the left for the STSpost + al ($F(1, 14) = 43.32$, $P < 0.001$), IPS ($F(1, 14) = 15.00$, $P < 0.05$), IPSant ($F(1, 14) = 11.90$, $P < 0.01$), IPSpost ($F(1, 14) = 6.50$, $P < 0.05$), and SYF/INS ($F(1, 14) = 18.71$, $P < 0.01$). Significant hemisphere effects were not observed for the voxel percentage activated counts from the FFG and the STSpc. More voxels were activated in the IPSpost as compared to the IPSant, $F(1, 14) = 15.00$, $P < 0.01$. The activated voxel counts did not differ by stimulus condition in any ROI.

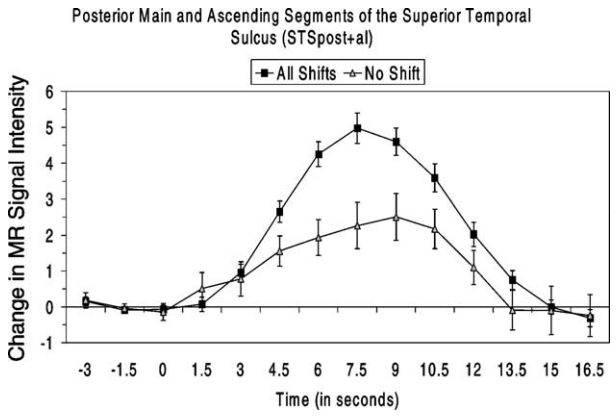

Fig. 3. The percentage of voxels whose correlation with the reference waveform was significant ($t \geq 1.96$, $P < 0.05$) by ROI and hemisphere.

We focused subsequent analyses on four subsets of activation—the STSpost + al, STSpc, IPSpost, and FFG. Using these areas, we conducted a second HDR waveform analysis on those voxels that were activated (defined earlier) within each ROI. The resulting waveforms were similar to those calculated for unselected voxels. However, the amplitudes were approximately twice as large. These waveforms were used to address three questions. (1) Do the HDR waveforms differ when we compare the time course waveforms in response to eye movements (all shifts) to those in response to the movement of the checkerboard stimulus alone (no shift)? (2) Are the HDR waveforms modulated by the timing of observed eye–gaze shifts? (3) Are the HDR waveforms influenced by the context of an observed eye–gaze shift?

3.2.4. Responses to eye–gaze shifts

We calculated the average amplitude of the HDRs across the five time points between 4.5 and 10.5 s. Repeated measures ANOVAs on these scores indicated that the average amplitude values were significantly greater for the all shifts than for no shift in the STSpost + al ($F(1, 14) = 27.65$, $P < 0.001$), STSpc ($F(1, 14) = 11.52$, $P < 0.01$), IPSpost ($F(1, 14) = 12.18$, $P < 0.01$), and FFG ($F(1, 14) = 6.54$, $P < 0.05$). The largest effect size for this comparison occurred for the STSpost + al ($F = 12.18$). The two HDR time courses from this region are displayed in Fig. 4. In all shifts, the signal was greater in amplitude at each averaged time point, as compared to no shift.

3.2.5. Influence of gaze shift timing on HDRs

To evaluate the influence of gaze shift timing on the HDR time courses, 2 (latency) × 2 (context) × 2 (hemisphere) repeated measures ANOVAs were conducted on the latency to peak amplitude measures. Peak amplitudes occurred significantly earlier for the 1 s gaze shift conditions relative to the 3 s gaze shift conditions in the STSpost + al ($M = 7.5$ s (1 s) versus 8.3 s (3 s), $F(1, 14) = 9.58$, $P < 0.01$) and STSpc ($M = 7.4$ s (1 s) versus 8.1 s (3 s), $F(1, 14) = 5.27$, $P < 0.05$) (see Fig. 5A and B). There were consistent trends in the direction of longer latencies to peak amplitude in the 3 s conditions as compared to the 1 s conditions in the IPSpost (M(S.D.) = 7.5 s (1 s) versus 8.1 s (3 s)) and FFG (M(S.D.) = 7.1 s (1 s) versus 7.5 s (3 s)) (see Fig. 5C and D). However, the effect of latency was not significant for either ROI. No other main effects or interactions were significant.

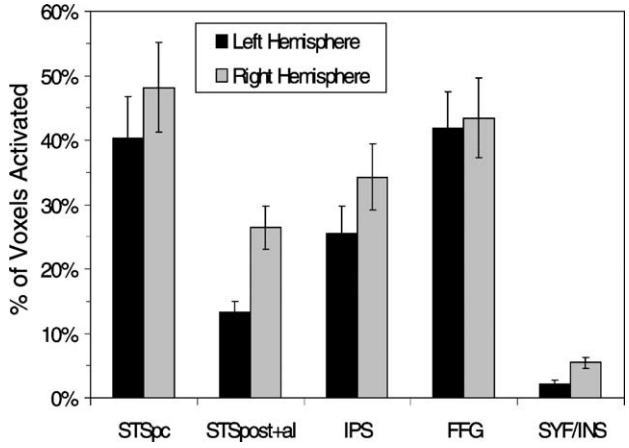

Fig. 4. Comparison of the HDR waveforms in response to observed shifts in eye–gaze and no shifts the STSpost + al. In this and every ROI, the HDR to all shifts condition was stronger than that to no shift.

3.2.6. *Influence of gaze shift context on HDRs*

The mean activation time courses for the correct 1 s and incorrect 1 s conditions are depicted by ROI in Fig. 6A–D. The effect of context was apparent in late-occurring amplitude differences in each ROI. The peak amplitude of the HDR was consistently greater for incorrect 1 s than for correct 1 s for the last 2–5 post S2 time points. The onset (∼9 s post S1 or 10 s post S2) and general form (largest ∼15.0–16.5 s post S1 onset) of this difference was generally consistent across ROI, but the magnitude varied by ROI. Peak differences were larger for the IPSpost and STSpc, than for the STSpost + al, which showed a greater difference than did the FFG.

To evaluate the effect of context on the activation waveforms, 2 (context) × 2 (hemisphere) repeated measures ANOVAs were conducted on the mean amplitude values at each of the eight time points between 6.0 and 16.5 s for each ROI and latency condition (see Table 1). In Fig. 6A–D, each post S2 time point at which mean signal values were significantly greater for incorrect 1 s compared to correct 1 s is marked with an asterisk. Significant differences were not observed before 10.5 s. At 10.5 s, the HDR amplitudes differed for the IPSpost. At 12, 13.5, 15 and 16.5 s the effect of context was significant for the STSpost + al, STSpc, and IPSpost. At 15 s and 16.5 s, the correct versus incorrect comparison was significant for the FFG. In contrast to the robust effects of context upon the 1 s delay data, no context effects were observed when comparing correct and incorrect conditions at 3 s.

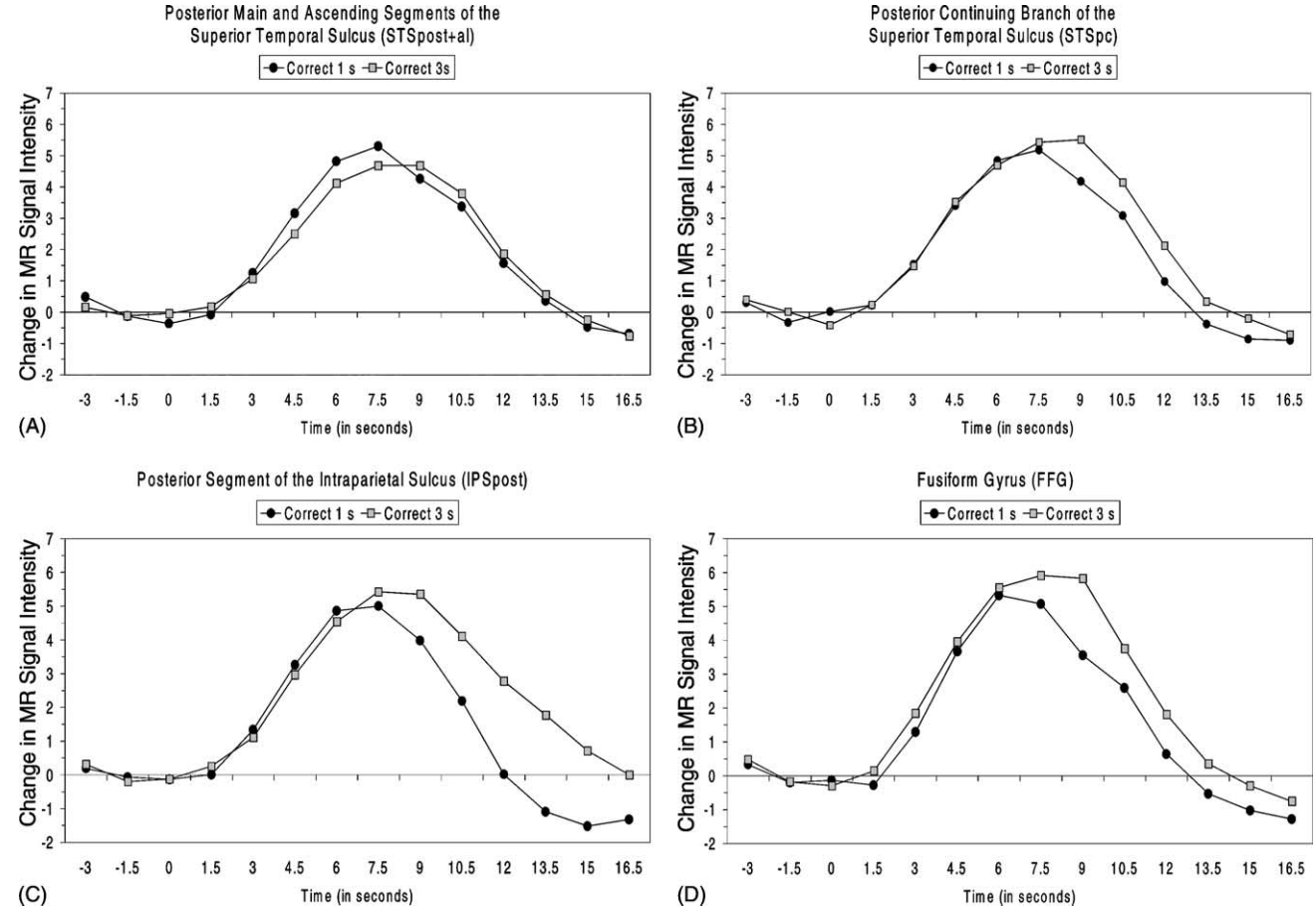

Fig. 5. Mean HDR time courses from the comparison between correct 1 s and correct 3 s by ROI (A–D). Significant (3 s > 1 s) shifts in latency to peak amplitude were observed for the STSpost + al and the STSpc, but not the IPSpost and FFG.

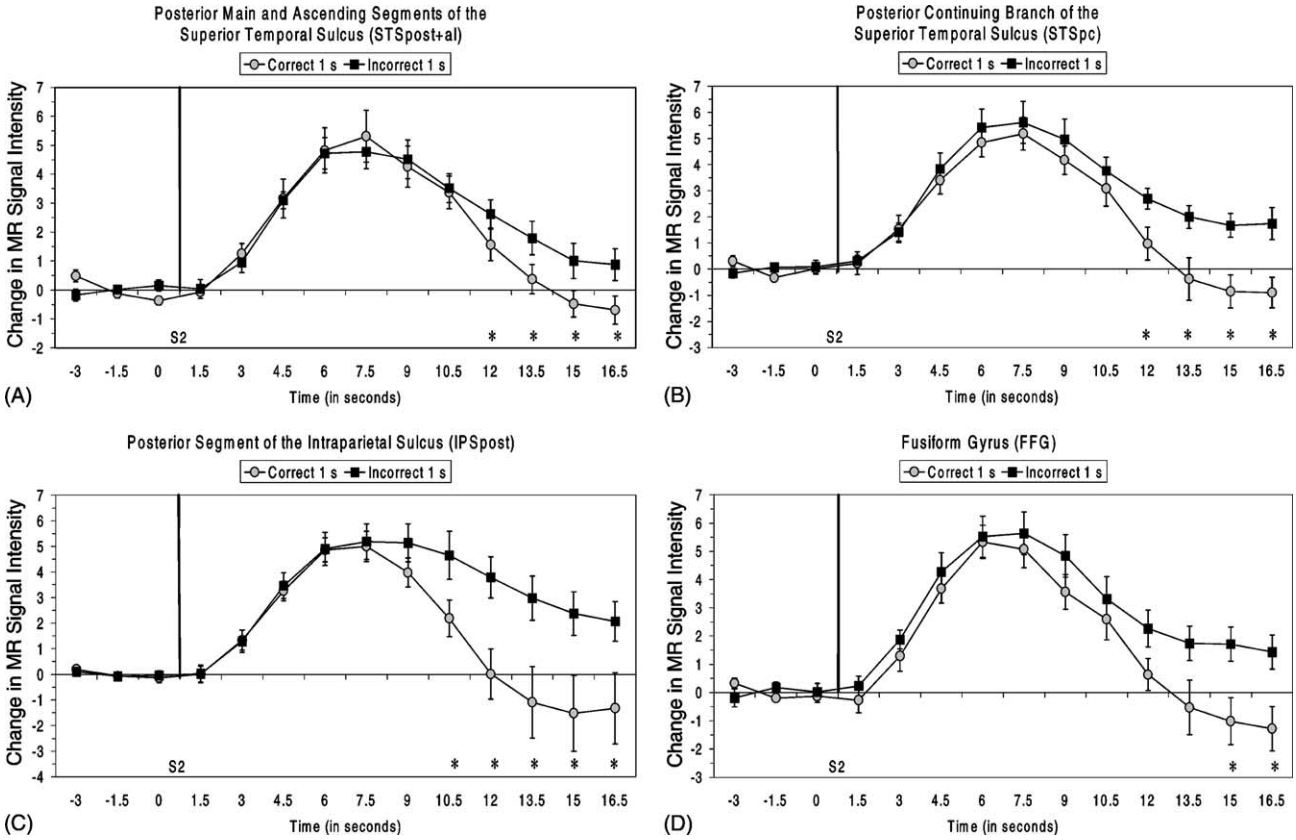

Fig. 6. Mean HDR waveforms activated voxels are plotted by ROI (A–D). Each post S2 time point at which mean signal values were significantly greater for the incorrect 1 s condition compared to the correct 1 s condition is marked with an asterisk.

Table 1
Effects of observed gaze shift context on HDR time courses by ROI

Correct 1 s vs. incorrect 1 s	10.5 (s)		12.0 (s)		13.5 (s)		15.0 (s)		16.5 (s)	
	F	*P*	*F*	*P*	*F*	*P*	*F*	*P*	*F*	*P*
Posterior main and ascending segments of STS										
Context (correct vs. incorrect)	0.169	0.688	4.677	0.048*	4.708	0.048*	6.169	0.026*	7.301	0.017*
Hemisphere (right vs. left)	1.464	0.246	3.147	0.098	0.005	0.942	0.099	0.758	0.663	0.429
Context × hemisphere	0.276	0.607	0.104	0.752	0.434	0.521	0.489	0.496	1.250	0.282
Posterior continuing branch of STS										
Context (correct vs. incorrect)	1.198	0.292	9.930	0.007*	10.87	0.005*	22.60	0.000*	34.75	0.000*
Hemisphere (right vs. left)	0.408	0.533	0.301	0.592	0.359	0.559	0.103	0.753	0.378	0.549
Context × hemisphere	0.105	0.751	0.008	0.930	0.066	0.800	0.085	0.775	0.547	0.473
Posterior IPS										
Context (correct vs. incorrect)	4.823	0.045*	8.078	0.013*	7.290	0.017*	6.359	0.024*	6.176	0.026*
Hemisphere (right vs. left)	5.168	0.039*	5.862	0.030*	2.122	0.167	0.943	0.348	0.493	0.494
Context × hemisphere	3.115	0.099	2.307	0.151	1.925	0.187	1.598	0.227	1.271	0.279
Fusiform gyrus										
Context (correct vs. incorrect)	0.040	0.844	1.876	0.192	1.845	0.196	4.243	0.052*	6.481	0.023*
Hemisphere (right vs. left)	0.093	0.765	1.258	0.281	0.685	0.422	0.004	0.950	1.533	0.236
Context × hemisphere	0.009	0.925	0.246	0.628	0.513	0.486	0.021	0.888	0.053	0.822

Note. $d.f. = (1, 14)$ for all statistical comparisons.
* $P \leq 0.05$.

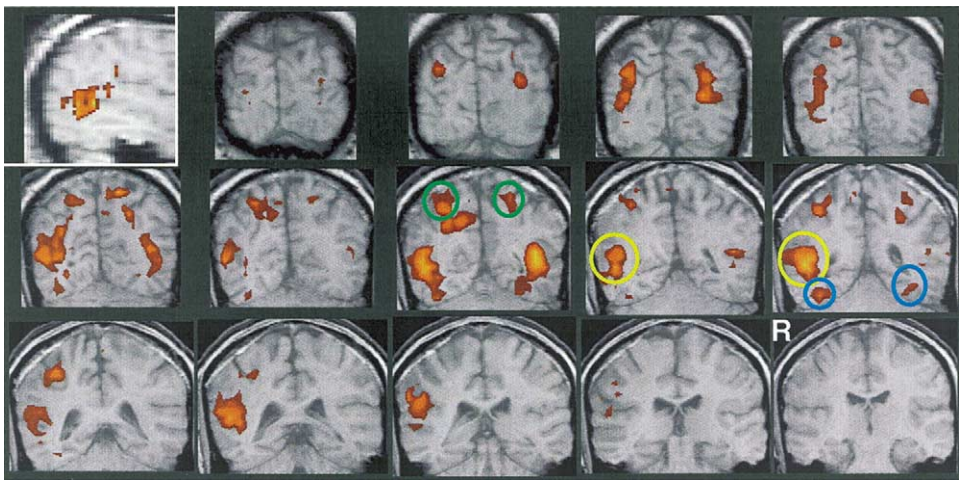

Fig. 7. The across-subjects activation *t*-statistic map for the all shifts condition overlaid on the template subject's anatomical images. The activation map represents the magnitude of the correlation (on a voxel-by-voxel basis) between the reference waveform and the HDR time course of voxels averaged over the four conditions containing a shift in gaze. The correlations were converted to *t* statistics and a threshold set at $t \geq 2.5$. Positive activations were observed in STS (yellow circles), IPS (green circles), and FFG (blue circles). A sagittal section showing the approximate location and extent of right hemisphere STS activation replaces the final posterior coronal image.

3.2.7. *Voxel-based analyses*

Fig. 7 presents the across-subjects *t*-statistic map for the all shifts condition overlaid on the template subject's anatomical images from 30 mm (lower right corner) to 95 mm (upper left, last coronal image) posterior from the AC. The activation map represents the magnitude of the correlation (on a voxel-by-voxel basis) between the reference waveform and the HDR time course of voxels averaged

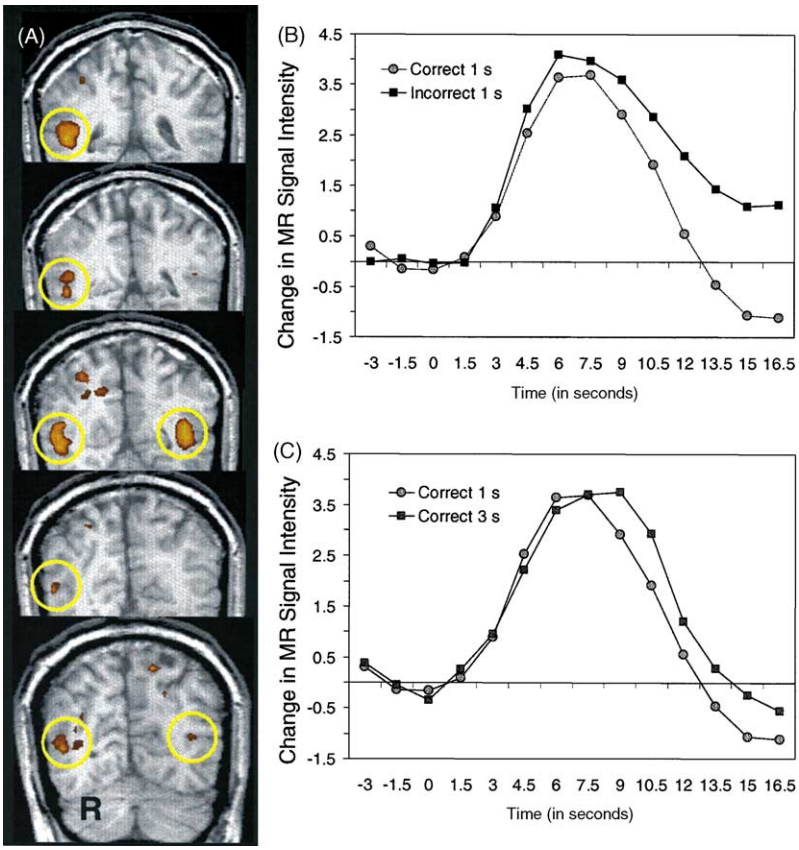

Fig. 8. (A) We interrogated activated ($t \geq 3.6$) voxels in the STS (yellow circles) for slices 45–65 mm (left to right) posterior from the AC. The top panel (B) shows the effect of context, and the bottom panel (C) illustrates the effect of timing.

over the four conditions containing a shift in gaze. Robust positive activations were observed in the STSpost, STSal, STSpc, IPS, and FFG. The activations appeared to be more spatially extensive and of greater magnitude in the right hemisphere for the three STS areas and the FFG. Bilateral activations were observed for the IPS. Activation was not observed in the final posterior coronal image; therefore, we substituted an image showing the approximate location and extent of right hemisphere STS complex activation in a sagittal section (top left corner). This pattern of activations was consistent with the results of our ROI based analyses.

We interrogated activated ($t \geq 3.6$, $P < 0.001$, uncorrected) voxels in the STS complex (yellow circles in Fig. 8A) for slices 45–65 mm posterior form the AC. HDR waveforms and effects obtained form these voxels were entirely consistent with the results form our ROI analyses. The top panel (Fig. 8B) shows the effect of context. The bottom panel (Fig. 8C) illustrates the effect of timing.

An across-subjects *t*-statistic map comparing the all shifts and no shift conditions at 7.5 s (overlaid upon the corresponding template anatomical images) was generated by calculating (on a voxel-by-voxel basis) the difference between the two conditions at each time point (Fig. 9A). Voxels showing a significant all shifts > no shift differences are indicated ($t \geq 1.96$, $P < 0.05$, uncorrected) for five slices 40–60 mm posterior from the AC (left to right in 5-mm increments). Differences were observed in the STSpost, STSal, STSpc, IPS, and FFG. Areas of difference were more extensive and the mean differences were larger in the right three STS regions and IPS. This analysis employed an activation mask such that only those that were significantly activated in the all shifts condition were included ($t \geq 1.96$, $P < 0.05$, uncorrected).

Fig. 9B presents results from the correct 1 s versus incorrect 1 s comparison at 15 s for a single slice 55 mm posterior from the AC. Incorrect > correct differences were observed bilaterally in the STS and FFG and in the right IPS ($t \geq 1.96$, $P < 0.05$, uncorrected). Consistent with the ROI analyses, differences were not observed for the incorrect 3 s > correct 3 s comparison.

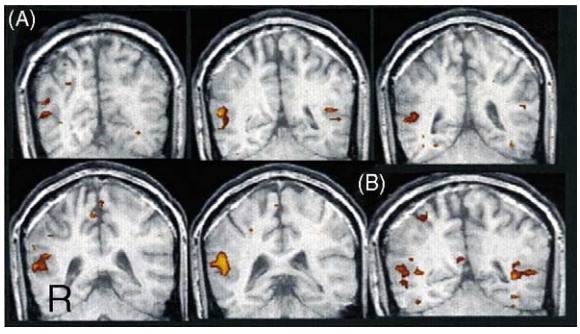

Fig. 9. (A) Across-subjects *t*-statistic map comparing the all shifts and no shift conditions at 7.5 s. Voxels showing a significant all shifts > no shift difference ($t \geq 1.96$, $P < 0.05$) are indicated for slices 40–60 mm posterior from the AC (bottom to top). (B) Comparison between correct 1 s and incorrect 1 s at the 15 s time point. One slice (50 mm posterior from the AC) is shown. Voxels showing a significant (incorrect > correct) difference ($t \geq 1.96$, $P < 0.05$) are indicated. Red-to-yellow color map indicates lower to higher for *t*-values from 1.96 to +6.0.

4. Discussion

The current findings confirm the results of previous studies that reported activation in discrete brain regions elicited by the perception of a gaze shift. The present research extends prior work by demonstrating that the context in which an eye movement occurs modulates activity in brain regions associated with gaze shift perception. The pattern of results observed in portions of the STS, IPS, and FFG was similar, and thus a single explanation may account for activation in all of these regions. We believe this is unlikely, however, given the different functions ascribed to these regions in prior studies.

4.1. Superior temporal sulcus

Our study focused upon the STS, as prior studies from our group [55,56] and others [31,66] have implicated the STS in eye–gaze perception. Within the STS, the highest levels of activity were observed near the crux of the STS complex where the posterior portion of the straight segment joins the ascending branch and in the posterior continuing branch of the STS. Activity was much greater in magnitude and spatial extent in the right than left STS. This right hemisphere bias was also observed by Puce et al. ([55], see their Fig. 6). In contrast, Wicker et al. [66] reported bilateral STS activations elicited by averted gaze, while Hoffman and Haxby [31] reported greater left hemisphere STS activity when subjects attended to gaze. These inconsistencies in hemispheric bias may be related to stimulus properties. The current study and Puce et al. [55] used a stimulus face in which the eyes moved, while Hoffman and Haxby [31] and Wicker et al. [66] had subjects view static faces in which the eyes did not move.

Although our stimuli were complex, consisting of a visual target followed by a delayed gaze shift, we assert, based on two key results, that the gaze shift per se evoked the major activation. First, the HDR was significantly larger when the character made a gaze shift than when it did not. Indeed, for the same voxels, the HDR to the all shifts condition was generally twice the amplitude of the HDR elicited during the no shift condition. Second, the latency to the peak amplitude of the HDR was longer in the 3 s gaze-delay conditions compared to the 1 s conditions.

We did, however, observe a small but significant HDR in the STS during the no shift condition. A flickering checkerboard is a potent visual stimulus, yet it is unlikely that the evoked activity was elicited solely because of its visual properties. Others have found that the STS region is activated by static views of eyes, mouths, hands, and faces

[2,10,29,31,38,47,54,56]. Allison et al. [1] suggested that these activations could be interpreted in the framework of implied motion [61]. That is, while static face and eye stimuli may not involve motion per se, the appearance of the visual target near the character's eyes may imply motion of the eyes toward the target, or the subject may strongly anticipate such motion.

It could be argued that the activation differences in the STS region and other ROIs did not result from the perception of gaze, but rather were evoked by differential eye movements made by the subjects. Our eye movement monitoring study refutes this interpretation, as *all* stimulus conditions evoked the same total amount of eye movements. Although it is possible that subjects engaged in qualitatively different eye movements for each of the different experimental conditions, inspection of the average visual scanpaths from the five conditions revealed no obvious qualitative differences.

If the STS activation reflects the activation of a simple eye movement detector, there should be no difference in activation for gaze shifts made to a visual target compared to identical gaze shifts made to empty space. Our results show that gaze shifts to empty space evoked STS activation with significantly longer duration compared to gaze shifts that acquired the visual target. Thus, context had a strong effect upon gaze-related activation. We note, however, that these context effects were only present when the gaze shift occurred in close temporal proximity to the appearance of the visual target. With a longer delay, the context effect did not occur.

What process might be responsible for the context effects observed the STS? When the checkerboard appears, an observer might expect or anticipate that the character will shift its gaze to the target, and when the eyes capture the target, this expectation is met and STS activation ceases. When the target appears and the character shifts its gaze toward an empty location in space, the observer's expectation is violated and activity in the STS region is prolonged—perhaps related to a reformulation of an expectation or anticipation of a 'correcting' gaze shift. The disappearance of the target and the return of gaze to the center finally caused a cessation of STS activation.

Why were there no context effects for the 3 s gaze sifts in the STS? Temporal contiguity between onset of the checkerboard and the eye movement might have played an unexpectedly important role. Regardless of whether the gaze shift was correct or incorrect, if it happened shortly after the appearance of the target, the target appearance and gaze shift were likely perceived as linked events. In contrast, when the appearance of the checkerboard and the eye movement were separated by a delay of 3 s, the link between the two events could have been ambiguous to the observer. This explanation could be explored by parametrically varying the delay between the target onset and gaze shift. If temporal contiguity is an important factor, we should observe a relationship between the length of the delay and the magnitude of the difference in HDRs evoked by incorrect versus correct gazes.

4.2. Intraparietal sulcus

The highest levels of IPS activity were observed in the posterior half of this ROI and were biased to the right hemisphere. The latency to peak of the HDR in IPS was numerically longer for the 3 s than the 1 s delayed gaze shifts, but this latency difference did not reach significance. Like the STS region, the IPS was sensitive to the context of the gaze shift.

In the current study, the pattern of results within the IPS closely mirrors that of the STS. However, because the IPS is activated during tasks involving spatial perception, particularly covert shifts in spatial attention [12–14,33,34,48], gaze-related activity in the IPS has been interpreted as reflecting the engagement of this attentional system for encoding the spatial direction of another's gaze and for mediating covert shifts of spatial attention in the direction of another's gaze [31]. This interpretation is consistent with the current findings.

Previous studies of gaze have reported IPS activation [31,55], but not to the extent observed here. In all previous functional neuroimaging studies concerned with gaze perception, the stimulus face gazed towards empty space, while here we provided a target for the gaze shift. Thus, our design might have confounded two kinds of processing associated with shifts in spatial attention—involuntary shifts of attention to the visual target and shifts of attention cued by eye–gaze. On correct trials, the observer's attention is attracted to the flickering visual target, and the character's gaze shift directs attention to the same location. In contrast, on incorrect trials the observer shifts attention away from the target to the new location in space indicated by the character's gaze. We presume that this reallocation engages the IPS for a second time to create an updated visuospatial representation. Thus, it is possible that the observed late increases in activity for the incorrect 1 s condition relative to the correct 1 s condition reflect the reformulation of the visuospatial representation.

Why there were no context effects for the 3 s gaze sifts in the IPS? It is likely that the onset of the checkerboard-patterned target was a powerful attractor of the subject's attention. However, it is also likely that its strength as an attention-holding agent declined as a function of time since stimulus onset. Thus, when the eyes moved in the correct 1 s and incorrect 1 s conditions, attention was likely focused near the checkerboard. In contrast, after 3 s of viewing the target stimulus, we suggest that the subject was no longer attending to the target, but rather had begun to shift attention elsewhere, perhaps to the character's face. The eye movement study allowed us to explore this explanation. We reasoned that if the checkerboard was a powerful attractor of attention then, in a free viewing setting, a large proportion of viewing time would have been devoted to it. Furthermore, if it lost its power to hold attention over time, then the proportion of viewing time at its location should decline. We calculated the mean POR for each 1 s time interval between

1 and 5 s for the no gaze shift condition. With each second that the target was present, the mean POR position shifted away from the locale of the checkerboard and towards the face, with the largest shift occurring between 2 and 3 s after the onset of the target. By 3 s, the POR had shifted an average of 2.2°. Thus, when a gaze shift finally occurred, the subject was no longer attending to the visual target, and the processing effort involved in reallocating attention to a new location was the same for both correct and incorrect conditions.

4.3. Fusiform gyrus

The HDR within the FFG was larger in all conditions involving an eye movement compared to the no gaze shift condition. The latency to peak of the HDR in the FFG was not significantly influenced by the timing of a gaze shift; although a trend towards a delayed latency to peak amplitude was evident. Activity in the FFG was also influenced by the context of the gaze shift, albeit to a lesser degree than in the STS and IPS. The sensitivity of FFG activation to gaze shifts and, in particular, to their context was an unexpected finding, as prior studies have emphasized the role of the FFG in processing invariant aspects of faces [27,31,46]. With the exception of the gaze shifts, the stimulus face was present throughout the duration of each run. However, it is possible that subjects averted their eyes to the checkerboard and then returned to the face, thus reprocessing the face with each shift of attention, and evoking a response in FFG. Nonetheless, it is important to emphasize that no effort was made in the present study to identify face specific regions of the FFG in lateral posterior fusiform gyrus. Therefore, these findings do not speak to possible characteristics of the fusiform face area generally. Activity was observed bilaterally along the entire extent of the FFG that was examined, and no spatial gradient in activity was noted.

4.4. Other brain regions involved in processing gaze shifts

We speculate here about the functions of the STS, IPS, and FFG in gaze perception and how they might relate to existing proposals concerning the neurofunctional organization of social cognition. The STS region appears to be performing functions similar to those of Baron-Cohen's proposed ID and ToMM [4]. That is, the STS region is sensitive to eye movements, and like the proposed ID; it is sensitive to the intentionality or goal-directedness of the movements. This conclusion is consistent with previous research showing that the STS region may assess the intention of actions. PET studies in humans show that observation of hand movements activates the STS [21,22,57,58], but meaningful hand movements (i.e. pantomimes) evoke greater regional cerebral blood flow (rCBF) within the STS region than do insignificant gestures [23,24]. Similarly, in monkeys, populations of STS neurons fire when an observed reach is target-directed but not when an observed reach is directed to empty space [35]. Moreover, like the hypothesized ToMM, the STS seems to be participating in the formulation (or reformulation) and monitoring of predictions about the character's behavior and interpreting her behavior in that context. The IPS appears to be serving a function similar to Baron-Cohen's proposed SAM [4]. That is, the IPS appears to support the identification of another person's direction of gaze and facilitate joint attention by formulating visuospatial representations that direct an observer's attention in the direction indicated by another's gaze. However, in this study we did not explicitly attempt to activate the SAM by requiring subjects to move their eyes to the shared target (although this likely occurred). This would require a design in which subjects did or did not share their attention with a target, while making equivalent eye movements in both conditions.

Acknowledgements

We are grateful to Jeremy Goldstein, Lilly Kinross-Wright, Karen Emberger, Sarah Hart, and Ronald Viola for assistance in data acquisition and analysis and manuscript preparation. We thank Dr. Martin J. McKeown for developing the software used in aligning each subject's anatomical images to a common space. We thank Dr. Gary Glover of Stanford University for providing source code for the spiral pulse sequence. We thank Drs. Allen Song and James Voyvodic for assistance with several aspects of this research. This research was supported by the Department of Veterans Affairs and NIH grant MH-05286. Kevin Pelphrey was supported by NICHD 1-T32-HD40127.

References

[1] Allison T, Puce A, McCarthy G. Social perception from visual cues: role of the STS region. Trends in Cognitive Science 2000;4:267–78.
[2] Allison T, Puce A, Spencer DD, McCarthy G. Electrophysiological studies of human face perception. I. Potentials generated in occipitotemporal cortex by face and non-face stimuli. Cerebral Cortex 1999;9:415–30.
[3] Argyle M, Cook M. Gaze and mutual gaze. Cambridge University Press: New York; 1976.
[4] Baron-Cohen S. Mindblindness: an essay on autism and theory-of-mind. MIT Press: Cambridge (MA); 1995.
[5] Baron-Cohen S, Wheelwright S, Hill J, Raste Y, Plumb I. The reading the mind in the eyes test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. Journal of Child Psychology and Psychiatry 2001;42:241–51.
[6] Bassili JN. On-line cognition in person perception. Lawrence Erlbaum: Hillsdale (NJ); 1989.
[7] Bentin S, McCarthy G, Wood CC. Event-related potentials, lexical decision and semantic priming. Electroencephalography and Clinical Neurophysiology 1985;60:343–55.
[8] Brothers L. The social brain: a project for integrating primate behavior and neurophysiology in a new domain. Concepts in Neuroscience 1990;1:27–51.
[9] Bruce V, Young A. Understanding face recognition. British Journal of Psychology 1986;77:305–27.

[10] Calder AZ, Lawrence AD, Keane K, Scott SK, Owen AM, Christoffels I, et al. Reading the mind from eye gaze.. Neuropsychologia 2002;40:1129–38.
[11] Campbell R, Heywood CA, Cowey A, Regard M, Landis T. Sensitivity to eye gaze in prosopagnosic patients and monkeys with superior temporal sulcus ablation. Neuropsychologia 1990;28:1123–42.
[12] Corbetta M, Akbudak E, Conturo TE, Snyder AZ, Ollinger JM, Drury HA, et al. A common network of functional areas for attention and eye movements. Neuron 1998;21:761–73.
[13] Corbetta M, Shulman GL. Human cortical mechanisms of visual attention during orienting and search. Philosophical Transactions of the Royal Society of London—Series B: Biological Sciences 1998;353:1353–62.
[14] Corbetta M, Shulman GL, Miezin FM, Petersen SE. Superior parietal cortex activation during spatial attention shifts and visual feature conjunction. Science 1995;270:802–5.
[15] Duvernoy HM. The human brain: surface, three-dimensional sectional anatomy with MRI, and blood supply. Springer-Wien: New York; 1999.
[16] Ekman P. Emotion in the human face. Cambridge University Press: New York; 1982.
[17] Emery NJ. The eyes have it: the neuroethology, function and evolution of social gaze. Neuroscience and Biobehavioral Reviews 2000;24:581–604.
[18] Farroni T, Johnson MH, Brockbank M, Simion F. Infants' use of gaze direction to cue attention: the importance of perceived motion. Visual Cognition 2000;7:705–18.
[19] Frith CD, Frith U. Interacting minds—a biological basis. Science 1999;286:1692–5.
[20] Goren CC, Sarty M, Wu PY. Visual following and pattern discrimination of face-like stimuli by newborn infants. Pediatrics 1975;56:544–9.
[21] Grafton ST, Arbib MA, Fadiga L, Rizzolatti G. Localization of grasp representations in humans by positron emission tomography. 2. Observation compared with imagination. Experimental Brain Research 1996;112:103–11.
[22] Grafton ST, Fagg AH, Woods RP, Arbib MA. Functional anatomy of pointing and grasping in humans. Cerebral Cortex 1996;6:226–37.
[23] Grezes J, Costes N, Decety J. Top–down effect of strategy on the perception of human biological motion: a PET investigation. Cognitive Neuropsychology 1998;15:553–82.
[24] Grezes J, Costes N, Decety J. The effects of learning and intention on the neural network involved in the perception of meaningless actions. Brain 1999;122:1875–87.
[25] Haith MM, Bergman T, Moore MJ. Eye contact and face scanning in early infancy. Science 1977;198:853–5.
[26] Hasselmo ME, Rolls ET, Baylis GC, Nalwa V. Object-centered encoding by face-selective neurons in the cortex in the superior temporal sulcus of the monkey. Experimental Brain Research 1989;75:417–29.
[27] Haxby JV, Hoffman EA, Gobbini MI. The distributed human neural system for face perception. Trends in Cognitive Science 2000;4:223–33.
[28] Haxby JV, Petit L, Ungerleider LG, Courtney SM. Distinguishing the functional roles of multiple regions in distributed neural systems for visual working memory. NeuroImage 2000;11:145–56.
[29] Haxby JV, Ungerleider LG, Clark VP, Schouten JL, Hoffman EA, Martin A. The effect of face inversion on activity in human neural systems for face and object perception. Neuron 1999;22:189–99.
[30] Heywood CA, Cowey A. The role of the face-cell area in the discrimination and recognition of faces by monkeys. Philosophical Transactions of the Royal Society of London—Series B: Biological Sciences 1992;335:31–7.
[31] Hoffman EA, Haxby JV. Distinct representations of eye gaze and identity in the distributed human neural system for face perception. Nature Neuroscience 2000;3:80–4.
[32] Hood BM, Willen JD, Driver J. Adult's eyes trigger shifts of visual attention in human infants. Psychological Science 1998;9:131–4.
[33] Hopfinger JB, Buonocore MH, Mangun GR. The neural mechanisms of top–down attentional control. Nature Neuroscience 2000;3:284–91.
[34] Hopfinger JB, Woldorff MG, Fletcher EM, Mangun GR. Dissociating top–down attentional control from selective perception and action. Neuropsychologia 2001;39:1277–91.
[35] Jellema T, Baker CI, Wicker B. Neural representation for the perception of the intentionality of actions. Brain and Cognition 2000;44:280–302.
[36] Jha AP, McCarthy G. The influence of memory load upon delay-interval activity in a working memory task: an event-related functional MRI study. Journal of Cognitive Neuroscience 2000;12:90–105.
[37] Johnson MH, Dziurawiec S, Ellis H, Morton J. Newborns' preferential tracking of face-like stimuli and its subsequent decline. Cognition 1991;40:1–19.
[38] Kanwisher N, McDermott J, Chun MM. The fusiform face area: a module in human extrastriate cortex specialized for face perception. Journal of Neuroscience 1997;17:4302–11.
[39] Kawashima R, Sugiura M, Kato T, Nakamura A, Hatano K, Ito K, et al. The human amygdala plays an important role in gaze monitoring. A PET study. Brain 1999;122:779–83.
[40] Kleinke CL. Gaze and eye contact: a research review. Psychological Bulletin 1986;100:78–100.
[41] Langton SR. The mutual influence of gaze and head orientation in the analysis of social attention direction. Quarterly Journal of Experimental Psychology: Human Experimental Psychology 2000;53:825–45.
[42] Langton SR, Watt RJ, Bruce II. Do the eyes have it? Cues to the direction of social attention. Trends in Cognitive Science 2000;4:50–9.
[43] Luria SM, Strauss MS. Comparison of eye movements over faces in photographic positives and negatives. Perception 1978;7:349–58.
[44] Mai JK, Assheuer J, Paxinos G. Atlas of the human brain. Academic Press: San Diego; 1997.
[45] Maurer D, Salapatek P. Developmental changes in the scanning of faces by young infants. Child Development 1976;47:523–7.
[46] McCarthy G. Physiological studies of face processing in humans. In: Gazzaniga MS, editors. The New Cognitive Neurosciences. MIT Press: Cambridge; 1999. p. 393–410.
[47] McCarthy G, Puce A, Belger A, Allison T. Electrophysiological studies of human face perception. II. Response properties of face-specific potentials generated in occipitotemporal cortex. Cerebral Cortex 1999;9:431–44.
[48] Nobre AC, Sebestyen GN, Gitelman DR, Mesulam MM, Frackowiak RS, Frith CD. Functional localization of the system for visuospatial attention using positron emission tomography. Brain 1997;120:515–33.
[49] Noton D, Stark L. Eye movements and visual perception. Scientific American 1971;224:35–43.
[50] Pelphrey KA, Reznick JS, Sasson NJ, Paul G, Goldman BD, Piven J. Visual Scanning of Faces in Autism. Journal of Autism and Developmental Disorders 2002;32:249–61.
[51] Perrett DI, Hietanen JK, Oram MW, Benson PJ. Organization and functions of cells responsive to faces in the temporal cortex. Philosophical Transactions of the Royal Society of London—Series B: Biological Sciences 1992;335:23–30.
[52] Perrett DI, Smith PA, Potter DD, Mistlin AJ, Head AS, Milner AD, et al. Visual cells in the temporal cortex sensitive to face view and gaze direction. Proceedings of the Royal Society of London—Series B: Biological Sciences 1985;223:293–317.
[53] Phillips ML, David AS. Viewing strategies for simple and chimeric faces: an investigation of perceptual bias in normals and schizophrenic patients using visual scan paths. Brain and Cognition 1997;35:225–38.

[54] Puce A, Allison T, Asgari M, Gore JC, McCarthy G. Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. Journal of Neuroscience 1996;16:5205–15.
[55] Puce A, Allison T, Bentin S, Gore JC, McCarthy G. Temporal cortex activation in humans viewing eye and mouth movements. Journal of Neuroscience 1998;18:2188–99.
[56] Puce A, Smith A, Allison T. ERPs evoked by viewing facial movements. Cognitive Neuropsychology 2000;17:221–40.
[57] Rizzolatti G, Fadiga L, Gallese V, Fogassi L. Premotor cortex and the recognition of motor actions. Brain Research and Cognitive Brain Research 1996;3:131–41.
[58] Rizzolatti G, Fadiga L, Matelli M, Bettinardi V, Paulesu E, Perani D, et al. Localization of grasp representations in humans by PET. 1. Observation versus execution. Experimental Brain Research 1996;111:246–52.
[59] Roberts M, Hanaway J, Morest DK. Atlas of the human brain in section. Lea and Febiger: Philadelphia; 1987.
[60] Scaife M, Bruner JS. The capacity for joint visual attention in the infant. Nature 1975;253:265–6.
[61] Senior C, Barnes J, Giampietro V, Simmons A, Bullmore ET, Brammer M, et al. The functional neuroanatomy of implicit-motion perception or representational momentum. Current Biology 2000; 10:16–22.
[62] Talairach J, Tournoux P. Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system: an approach to cerebral imaging. Thieme Medical Publishers: New York; 1988.
[63] Vecera SP, Johnson MH. Gaze detection and the cortical processing of faces. Visual Cognition 1995;2:59–87.
[64] Voyvodic JT. Real-time fMRI integrating paradigm control, physiology, behavior, and on-line statistical analysis. NeuroImage 1999;10:91–106.
[65] Walker-Smith GJ, Gale AG, Findlay JM. Eye movement strategies involved in face perception. Perception 1977;6:313–26.
[66] Wicker B, Michel F, Henaff MA, Decety J. Brain regions involved in the perception of gaze: A PET study. NeuroImage 1998;8: 221–7.
[67] Yamane S, Kaji S, Kawano K. What facial features activate face neurons in the inferotemporal cortex of the monkey? Experimental Brain Research 1988;73:209–14.