Common and distinct brain activation to viewing dynamic sequences of face and hand movements

James C. Thompson, a,⁎ Jillian E. Hardee, b,f Anita Panayiotou, d,e David Crewther, d and Aina Puce b,c,f

a Department of Psychology and Krasnow Institute for Advanced Study, George Mason University, MSN 3F5, Fairfax, VA 22030, USA
b Center for Advanced Imaging, West Virginia University, Morgantown, WV, USA
c Department of Radiology, West Virginia University, USA
d Brain Sciences Institute, Swinburne University, Melbourne, Australia
e School of Psychological Science, La Trobe University, Melbourne, Australia
f Department of Neurobiology and Anatomy, West Virginia University, USA

Received 28 February 2007; revised 17 April 2007; accepted 23 May 2007
Available online 14 June 2007

The superior temporal sulcus (STS) and surrounding lateral temporal and inferior parietal cortices are an important part of a network involved in the processing of biological movement. It is unclear whether the STS responds to the movement of different body parts uniformly, or if the response depends on the body part that is moving. Here we examined brain activity to recognizing sequences of face and hand movements as well as radial grating motion, controlling for differences in movement dynamics between stimuli. A region of the right posterior STS (pSTS) showed common activation to both face and hand motion, relative to radial grating motion, with no significant difference between responses to face and hand motion in this region. Distinct responses to face motion relative to hand motion were observed in the right mid-STS, while the right posterior inferior temporal sulcus (pITS) and inferior parietal lobule (IPL) showed greater responses to hand motion relative to face motion. These findings indicate that while there may be distinct processing of different body part motion in lateral temporal and inferior parietal cortices, the response of the pSTS is not body part specific. This region may provide input to other parts of a network involved with processing human actions with a high-level visual description of biological motion.

Keywords: Biological motion; STS; IPL; IFG; fMRI

The ability to accurately interpret another’s behavior is critical for social communication and understanding. Recent neuroimaging studies have revealed that regions in the inferior frontal, ventral premotor, lateral temporal and inferior parietal cortex form a network associated with the perception of human movement (Beauchamp et al., 2002; Bonda et al., 1996; Buccino et al., 2001, 2004; Grezes et al., 2001; Pelphrey et al., 2005; Puce et al., 1998; Rizzolatti et al., 2001; Wheaton et al., 2004). A key factor in understanding biological movement recognition is to determine if the movement of different body parts is processed in a uniform manner within the network, or whether regions in the network respond differentially to various body parts. Functional magnetic resonance imaging (fMRI) studies suggest that within the ventral premotor cortex viewing the movement of different body parts activation is consistent with the known somatotopy of motor responses in this region: viewing leg motion activates a region dorsal and posterior to areas active to viewing face motion, while a region between the two activates to viewing hand motion (Buccino et al., 2001; Wheaton et al., 2004). However, it is less clear if there is any organization in activation to viewing the motion of different body parts in other brain regions making up the biological movement recognition network.

The superior temporal sulcus (STS) makes an important contribution to the recognition of other people’s movements (Allison et al., 2000; Puce and Perrett, 2003), in particular for movements of articulated joints (Beauchamp et al., 2002) and for the body as a whole (Thompson et al., 2005). One role that has been attributed to the STS is to provide other brain regions in the network with a visual description of the details of human actions (Iacoboni et al., 2001). It is not known if this visual description maintains details of the body part that generated the motion, implying an intermediate-level representation, or if instead a higher-level representation is generated which is independent of moving body part. In a recent study, Pelphrey and colleagues (2005) described a posterior-to-anterior organization of fMRI responses to viewing hand, eye, and mouth movement within the superior temporal sulcus (STS). Other studies have indicated a different response profile to face and hand motion across the STS versus parietal cortex (Thompson et al., 2004; Wheaton et al., 2004). These data suggest that an intermediate-level description is

⁎ Corresponding author.
E-mail address: jthompsz@gmu.edu (J.C. Thompson).
Available online on ScienceDirect (www.sciencedirect.com).
formed in the STS in which details of the body part generating the motion are maintained, or that there may be a preference for face motion in the STS. In contrast, a number of other studies have demonstrated that the STS also responds to non-biological geometric objects that move in a manner that implies animacy or intentionality (Castelli et al., 2000; Blakemore et al., 2003; Schultz et al., 2005; Martin and Weisberg, 2003). It is possible therefore that the STS may form a high-level representation of moving stimuli, rather than maintaining and conveying information about the body part that generated the motion.

An important consideration when examining brain responses to motion of different body parts is that patterns of responses might, in part, be due to differences in the movement kinematics and dynamics of different body parts (Gentilucci et al., 2001). Observers are highly sensitive to the dynamic features of biological movement, and can use abstract temporal cues alone to discriminate between self- and other-produced actions (Flach et al., 2004). The importance of dynamic information when viewing the motion of different body parts is reinforced by findings that attending to the timing of dynamic sequences of non-biological visual stimuli can activate the ventral premotor, inferior parietal, and superior temporal cortex (Coull, 2004; Schubotz and von Cramon, 2004).

Previous studies examining brain responses in the STS to viewing the movement of different body parts have used passive viewing (Pelphrey et al., 2005; Puco et al., 1998; Wheaton et al., 2004), or detection tasks that may have allowed different strategies for processing mouth and hand movements (Thompson et al., 2004). Here, we used computer animated stimuli to represent face, hand and non-biological (control) movement. We asked subjects to focus on the dynamics of the stimuli and detect target motion sequences that were identical across stimulus types. If the STS forms an intermediate representation of biological movement where details of body part are maintained, we would expect to see non-overlapping responses to face and hand motion, once movement dynamics and meaning or task-relevance of the stimuli are controlled for.

In contrast, if there are overlapping responses to face and hand motion in STS this may suggest a high-level representation of biological movement that is not body-part dependent.

Materials and methods

Subjects

Eleven healthy, neurologically normal right-handed subjects (5 males) aged 22–40 years participated in a study approved by the Ethics Committee of Swinburne University of Technology, Melbourne, Australia. All subjects had normal or corrected-to-normal vision.

Experimental design

Subjects lay supine and viewed movies of face, hand, or radial grating motion presented through the MRI scanner’s control room window on a screen placed at the edge of the patient’s gurney of the MRI scanner via a mirror mounted on the quadrature head coil. Stimuli were presented on a PC running Presentation (Neurobehavioral Systems, Albany CA).

Each stimulus category was presented in blocks of mouth, finger, or radial grating motion each lasting 40 s alternating with blocks of fixation lasting 20 s. During fixation, subjects fixated on a red cross presented in the center of a grey screen. Pseudorandom sequences of short and long movements within each block were presented in each block (Fig. 1). Both the left and right hands were used as stimuli. A total of 4 runs were presented, with 2 blocks of each stimulus per run. Motion consisted of short (212 ms) or long (696 ms) movements of a face with opening mouth, an extending index finger on a hand, or inwardly moving rings of a radial motion stimulus (Fig. 1, top panel). Between each movement a static version of a face with closed mouth, a closed hand, or a radial pattern were presented with a random ISI of 1–3 s. Subjects responded with a button press when a target motion sequence of short–short–long movements occurred, and with another button press when they observed a second target sequence of long–long–short movements (Fig. 1, bottom panel). Successive target sequences were separated by at least one non-target motion stimulus triplet. A total of 28 target sequences were presented for each stimulus type (mouth, hand, radial pattern).

Stimulus generation

Stimuli were created using Poser 4.0 (Curious Labs, Santa Cruz, CA). The face stimulus covered a total area approximately 3° ×4.5° of horizontal and by vertical visual angle, respectively. The hand and the radial gratings covered approximately 3° ×3° of horizontal and vertical visual angle. Movies of face and hand motion were created by first initializing the face with mouth closed or the index finger curled in a closed hand (Fig. 1). A keyframe of the face with mouth fully open or finger fully extended was then created as the middle frame of 7 (short) or 23 (long) keyframes using a rate of 33 frames/s. Smooth motion from the initial keyframe to the middle keyframe and back was then performed with spline interpolation. For the radial grating stimuli, each ring was approximately 0.16° wide, and only the inner six of the nine rings moved. The color of the darkest and lightest rings matched the minimum and maximum RGB value of the face and hand stimuli. Gratings moved outwards for either 7 or 23 frames at 33 frames/s.

MRI data acquisition and analysis

A total of 160 volumes/condition were collected in four imaging runs in each subject. Gradient echo EPI BOLD signal was acquired from whole-brain at 3T (GE Medical Systems Horizon LX MRI scanner) from 17 oblique slices (+30° from axial; 5 mm thick, 1 mm gap, in plane resolution 1.875 mm, TR/TE=2000/40 ms). This slice prescription allowed whole brain coverage except for the ventral-most surface of the posterior occipital lobes. Anatomical images consisting of low-resolution T2 and high-resolution SPGR (0.5 mm ×0.5 mm ×1.5 mm) images were also collected.

Group analysis

Data analysis was performed using SPM2 (FIL, UCL London, UK), and Matlab 6.1 (Mathworks, Natick, MA). The low-resolution T2 volume was first coregistered to the high resolution SPGR. GE-EPI images were then co-registered to the T2 volume and corrected for motion and slice timing differences. The SPGR image was normalized to the Montreal Neurological Institute template, and the parameters determined for this normalization were applied to functional images. Voxels from the SPGR and functional images were resampled to 2 mm³ isovoxel resolution.
Finally, the functional images were smoothed with a Gaussian kernel with a full width half maximum (FWHM) of 8 mm.

Regressors for each of the experimental conditions were created by modeling the BOLD response to each stimulus condition as a box car with an ‘ON’ cycle equal to the length of each block, convolved with a hemodynamic response function. Included in the model were six motion covariates (three translation and three rotation parameters) determined from the motion correction, and a constant term to account for drift. A high-pass temporal filter of 1/128 s was also applied to the data to remove low frequency drifts in the MR signal.

Parameter estimates were created for each subject, and contrasts comparing all experimental conditions to fixation, in addition to contrasts comparing face or hand movement to the radial grating were created. The significance of each of these contrasts was examined by using one-sample t-tests at the group level in a random effects analysis. As the focus of this study was on posterior temporal and parietal cortex, we created a volume of interest (VOI) that contained the temporal, parietal and lateral occipital cortex using the WFU_Pickatlas (www.rad.wfubmc.edu/fmri) (Fig. 2). Results of the contrast for all experimental conditions relative to fixation were thresholded at \( p < 0.05 \) (corrected for multiple comparisons at the cluster level). Within the voxels that survived this threshold, contrasts between face or hand movement relative to radial gratings were thresholded at \( p < 0.005 \) (uncorrected) and clusters of >30 voxels. Overlap in activation to face and hand movement relative to radial gratings was determined by finding the intersection of each of these contrasts thresholded as above (Nichols et al., 2005). For the comparison between face and hand movement, we extracted the beta weights from the voxel of peak significance from the comparison of face or hand movement relative to radial gratings and compared them using a \( t \)-test.

We also performed a whole-brain, voxel-wise analysis by identifying regions that responded to all experimental conditions relative to baseline, thresholded with a cluster level threshold of \( p < 0.01 \) (corrected for multiple comparisons). Voxel-wise contrasts of face or hand movement relative to radial gratings were then thresholded at \( p < 0.005 \) (uncorrected) with clusters >30 voxels. Overlap in activation was determined by finding the intersection of activation as outlined above.

**Single-subject analysis**

To examine the possible overlap in response to face and hand movement relative to radial gratings in more detail, we also conducted a single-subject analysis. Procedures used to transform individual brains into stereotaxic coordinates such as the MNI or Talairach coordinate systems align on the basis of anatomy rather than function (Brett et al., 2002; Beauchamp et al., 2004). In particular, volume-based warping procedures such as those used by SPM2 do not match individual brains to the template on the basis of

---

**Fig. 1. Activation task stimuli and sequence.** A grid showing still images of Face (top), Hand (middle) and Radial motion pattern (bottom) stimuli shown before and after the movement (left and right), and during motion endpoint (center). There were two possible motion durations (designated as ‘short’ or ‘long’). A typical ongoing experimental sequence is shown at the bottom of the figure.
If the STS varies considerably in anatomy across individuals, the location of coordinates corresponding to the STS in the template brain may not necessarily correspond to the STS in individuals. It is also possible that normalization procedures may lead to a distortion in the underlying functional distribution of responses to different stimuli. To examine single-subject responses, we conducted a whole-brain analysis on each subject's data in native space. Coregistered EPI data were resampled to 2 mm$^3$ and smoothed with an 8 mm FWHM Gaussian kernel. We then identified regions that responded to all experimental stimuli relative to baseline with a cluster threshold of $p < 0.05$ (corrected for multiple comparisons). Voxel-wise contrasts of face or hand.
movement relative to radial gratings were then thresholded at \( p<0.005 \). Overlap in activation was determined by finding the intersection in activation as outlined for the group analysis. In order to reference the location of the group activation to the single-subject anatomy and activation patterns, we obtained the deformations used for each subject to warp their brain to the MNI template and inverted these deformations, using the Deformations Toolbox in SPM2. Then, for each subject, we applied the inverted deformations to the MNI coordinates of the peak activations in temporal and parietal cortex of the group analysis and compared the location of group activation to single-subject anatomy.

**Results**

**Behavioral data**

No significant differences in mean accuracy for detecting the target sequences were observed between the three movement conditions Mean \( \pm \text{S.D.} \) Accuracy: Face=70.8 \( \pm \) 19.9; Hand=70.6 \( \pm \) 18.0; Radial Gratings=71.7 \( \pm \) 39.0. Similarly, there were no significant differences in mean reaction time (RT) to target sequence detection between the three conditions (Mean \( \pm \) S.D.) RT in ms: Face=755.7 \( \pm \) 134.4; Hand=798.7 \( \pm \) 232.7; Radial Gratings=772.7 \( \pm \) 149.2). 

**Group analysis**

Significant activation to the experimental conditions relative to fixation was observed throughout the temporal and parietal VOI (Figs. 2A and B, green). Within the areas that responded to the experimental conditions relative to fixation, significantly greater activation to the mouth movement than to radial gratings was observed in the right midSTS (\( p<0.001 \)) and right pSTS (\( p<0.001 \)) (Fig. 2A, red; Table 1). Significantly greater activation to hand movement than to radial gratings was observed in the right posterior inferior temporal sulcus (pITS) (\( p<0.0001 \)), right pSTS (\( p<0.001 \)) and right IPL (\( p<0.001 \)) (Fig. 2A, blue; Table 1). Importantly, as the main focus of our experiment we observed overlapping activation across face and hand movements, when compared to radial gratings, in the right pSTS (\( x=56, y=-48, z=12; \) Fig. 2A, purple). There was no significant difference between face movements relative to radial gratings when compared to hand movement relative to radial gratings (\( p>0.2 \)) in this region. Comparing responsivity in the midSTS revealed significantly greater activation to face movements compared to radial gratings relative to activation to hand movements compared to radial gratings (\( p<0.05 \)) (Fig. 2B). In contrast, there was significantly greater activation to hand movements compared to radial gratings than to face movements compared to radial gratings in the right pITS (\( p<0.005 \)) and right IPL (\( p<0.05 \)).

Whole-brain analysis revealed activation all three experimental conditions relative to fixation in the left ventral premotor cortex (vPM), left inferior frontal gyrus (IFG), left putamen, and right medial frontal gyrus (MFG) (Fig. 2C). Significantly greater activation to face motion relative to radial gratings, and hand motion relative to radial gratings, was observed in the left IFG, left putamen, and right MFG (Table 1). All three areas active to both face motion relative to radial gratings, and hand motion relative to radial gratings (Fig. 2D IFG: \( x=-49, y=8, z=2; \) MFG: 2, 29, 52; Put: \( -18, -2, 14 \)); in the left vPM there was significant activation to all three conditions compared to fixation, with no significant difference between conditions (Fig. 2D \(-46, -8, 46 \)).

**Single-subject analysis**

The single-subject analysis revealed a distribution of responses to face and hand movement relative to radial gratings that was consistent with the group activation. In particular, the location of coordinates corresponding to the peak activations in the pITS, midSTS, pSTS, and IPL in the group corresponded to these anatomical locations in single-subjects. Overall, the finding of overlapping activation to face movement relative to gratings and hand movement relative to gratings in the pSTS was observed in 7 out of 11 subjects using a threshold of \( p<0.005 \). Lowering the threshold to \( p<0.05 \) revealed that 9 out of 11 subjects showed overlapping activation in pSTS. The overlap of activation to face movement and hand movement relative to radial gratings from the 11 subjects is presented in Fig. 3. These results confirm that the results of the group analysis were not due to distortion of underlying response patterns by spatial normalization and instead reflect common processing of face and hand movement in this region. There was more variation across subjects in the response patterns of the pITS, midSTS and IPL in which there was distinct activation in the group analyses. For example, we observed overlapping activation to both face and hand movement relative to radial gratings in midSTS in several subjects (e.g. subjects AP and TA in Fig. 3).

**Discussion**

The present study examined fMRI responses as subjects monitored and detected dynamic target sequences of face, hand, and radial grating motion. Cortex in and around the STS has been shown to consistently respond to the movements of other biological forms (Allison et al., 2000; Puce and Perrett, 2003;...
Bonda et al., 1996; Puce et al., 1998; Beauchamp et al., 2002). Consistent with these findings, we found greater responses to face and hand movements than to radial grating motion in the right mid-STS and pSTS, as well as the right pITS and IPL. Importantly, the radial pattern motion also conveyed the same dynamic, task-relevant information as that presented by the face and hand. This suggests that these responses are not simply due to greater task-relevance of the biologically relevant stimuli. Of the regions in temporal and parietal cortex that responded significantly to face or hand motion relative to the radial grating, a significantly greater response to face motion relative to hand motion was observed in the right mid-STS, whereas a significantly greater response to hand motion relative to face motion was observed in the right pITS and right IPL. However, analysis of single-subject data revealed that there was some variability in the extent to which these regions showed distinct activation. Importantly, a region of the right pSTS showed overlapping responses to both face and hand motion, relative to radial grating motion, and there was no significant difference between the response to face or hand motion in this region. The overlapping activation in pSTS was consistent across individual subjects, indicating that it was not simply a result of the group averaging process and reflects common processing of face and hand motion in this region.

It has been suggested that the STS provides a visual description of motion details to other regions involved in the processing of other people’s actions, such as the IPL, PMv, and IFG (Iacoboni et al., 1999). However, the nature of this description has been unclear. A study by Pelphrey and colleagues (2005) indicated a somatotopic
arrangement of responses to viewing hand, eye and mouth movements was present in the STS and surrounding cortex. A number of other findings have suggested that there may even be a different pattern of responses to face motion compared to hand motion in the STS and parietal cortex (Thompson et al., 2004; Wheaton et al., 2004). These data suggest that the STS may form an intermediate-level, body-part specific representation of biological motion or that this region may prefer face motion to hand motion. The findings of the present study suggest instead that a region of the pSTS may form a high-level representation of biological movement that is not dependent on the particular body part generating the motion. An alternative hypothesis is that the distribution of responses in STS and surrounding cortex might in part be shaped by particular task requirements. In circumstances when one is focused on higher-level stimulus attributes, such as the dynamics of the motion in the present study, the pSTS might form a high-level representation that is not bound to a particular body part. In circumstances in which the motion of one particular body part might be the primary source of information, effector-specific representations may be formed and passed on to other parts of the network involved in action recognition. Such a flexible mechanism might explain the distribution of responses to different body part motion reported by Pelphrey and colleagues (2005), as during passive viewing subjects might have focused to a greater extent on the specific effector that was moving. Our findings do suggest, however, that it is not just the dynamics of the action that are processed by the pSTS, as the overlapping responses to face and hand motion were significantly greater than the response to radial grating motion even though the task-relevant information did not differ between face, hand and radial gratings.

It should be noted that, similar to the findings of Pelphrey and colleagues (2005), we observed greater responses to face motion than hand motion in the right mid-STS, and greater responses to hand motion than mouth motion in a more inferior region of the right lateral temporal cortex (pSTS) and the right IPL. There was variability across subjects in the extent to which each of these regions showed distinct responses to either face movement or hand movement relative to radial gratings, and lowering the threshold for significance indicated that several subjects showed overlapping responses in these regions. In general these results suggest that in addition to the common processing of different body part motion in the pSTS, there is some evidence of distinct responses to different body parts in the neighboring lateral temporal and inferior parietal regions (Pelphrey et al., 2005; Thompson et al., 2004; Wheaton et al., 2004). More anterior regions of the STS, particularly in the right hemisphere, respond to hearing speech and nonspeech human vocalizations (Belin et al., 2000). The mid-STS location of the strongest response to face motion may reflect the recruitment of mechanisms also involved with processing mouth motion associated with vocalizations. The pSTS response in the present study, with greater activation to hand versus face motion, is close to a region responding to static images of headless bodies, hands and hand actions that has been labeled the extrastriate body area, or EBA (Astafiev et al., 2004; Downing et al., 2001). The EBA also responds during the performance of hand actions, even when the hand is not visible (Astafiev et al., 2004). The IPL responds to viewing hand actions (Buccino et al., 2004), as well as the execution of hand actions (Buccino et al., 2004; Ehrsson et al., 2001). Both the EBA and IPL may be involved in visually-guided hand actions, with the right IPL in particular playing an important role in guiding hand actions in space (Mattingley et al., 1998; Husain et al., 2000).

In addition to activation in the temporal and parietal cortex, we also observed greater activation to face and hand movements, relative to radial grating motion, in the right MFG, left IFG and left putamen. Activation in the MFG and IFG to biological movement has been reported previously (Aziz-Zadeh et al., 2006; Grafton et al., 1996; Iacoboni et al., 2005, 1999; Johnson-Frey et al., 2003; Pelphrey et al., 2005; Rizzolatti et al., 2001; Schubotz and von Cramon, 2004). Pelphrey and colleagues (2005) found that the MFG responded to eye and mouth movements, but not hand movements (Pelphrey et al., 2005). Our data show that the MFG responded to hand movements and face movements, with a region of overlap between the two. One could argue that activation in these regions might also be expected as a function of subjects using working memory to keep track of presented stimuli in order to correctly identify the required target sequences. All three stimulus conditions, including the inwardly moving radial pattern, required subjects to keep track of the current sequences of previous three items in working memory. Interestingly, greater activation occurred to face or hand motion sequences than to the radial pattern, yet the working memory load and subjects’ performance did not differ across the three stimulus types, indicating that the ventral prefrontal cortex activation, often associated with manipulation of information in working memory (Wager and Smith, 2003), may not be driven by these processes. In contrast to previous studies, we did not observe distinct responses to face or hand motion in another region of frontal cortex, the so-called ventral premotor (PMv) region (Buccino et al., 2001; Wheaton et al., 2004). In the present study, all three movement conditions activated a large section of the left PMv. A recent study by Schubotz and von Cramon (2004) also reported overlapping activation to detecting sequences of biological and non-biological stimuli in the left PMv. Together with the present results, these findings raise the possibility that this region may encode sequential information that predicts action outcomes regardless of whether the stimuli are biological or not. This would be an interesting avenue of exploration in future studies.

Understanding how different brain regions respond to the motion of different body parts can help provide insight into the role of these regions in action recognition. The common response to face and hand motion in the pSTS suggests that this region may form a high-level representation of biological movement that is independent of body part. This response is, however, greater than the response to a non-biological stimulus matched for motion cues: role of the STS region. Trends Cogn. Sci. 4, 267–278.

Acknowledgments

The authors thank Ms. Mary Pettit for editorial assistance. The collection and initial analysis of data from this study by James Thompson was conducted as part of a postdoctoral fellowship in the Center for Advanced Imaging and Department of Radiology at West Virginia University. This study was funded by a Project Grant awarded to Aina Puce by the National Health and Medical Research Council (Australia).

References


