

# Psychological Science

<http://pss.sagepub.com/>

---

## **Action Understanding in the Superior Temporal Sulcus Region**

Brent C. Vander Wyk, Caitlin M. Hudac, Elizabeth J. Carter, David M. Sobel and Kevin A. Pelphrey  
*Psychological Science* 2009 20: 771  
DOI: 10.1111/j.1467-9280.2009.02359.x

The online version of this article can be found at:  
<http://pss.sagepub.com/content/20/6/771>

---

Published by:



<http://www.sagepublications.com>

On behalf of:



[Association for Psychological Science](http://www.apa.org)

**Additional services and information for *Psychological Science* can be found at:**

**Email Alerts:** <http://pss.sagepub.com/cgi/alerts>

**Subscriptions:** <http://pss.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.com/journalsPermissions.nav>

## Research Article

# Action Understanding in the Superior Temporal Sulcus Region

Brent C. Vander Wyk,<sup>1</sup> Caitlin M. Hudac,<sup>1</sup> Elizabeth J. Carter,<sup>2</sup> David M. Sobel,<sup>3</sup> and Kevin A. Pelphrey<sup>1</sup>

<sup>1</sup>Yale Child Study Center, Yale University; <sup>2</sup>Department of Psychology, Carnegie Mellon University; and <sup>3</sup>Department of Cognitive and Linguistic Sciences, Brown University

**ABSTRACT**—*The posterior superior temporal sulcus (STS) region plays an important role in the perception of social acts, although its full role has not been completely clarified. This functional magnetic resonance imaging experiment examined activity in the STS region as participants viewed actions that were congruent or incongruent with intentions established by a previous emotional context. Participants viewed an actress express either a positive or a negative emotion toward one of two objects and then subsequently pick up one of them. If the object that was picked up had received positive regard, or if the object that was not picked up had received negative regard, the action was congruent; otherwise, the action was incongruent. Activity in the right posterior STS region was sensitive to the congruency between the action and the actress's emotional expression (i.e., STS activity was greater on incongruent than on congruent trials). These findings suggest that the posterior STS represents not only biological motion, but also how another person's motion is related to his or her intentions.*

Being social animals in a physical environment, humans must understand the people around them, just as they must understand the physical environment. To achieve this understanding, whether of the environment or other people, the neural system must carry out processes of perception, recognition, integration, and prediction. In the social domain, this means that the neural system must have mechanisms dedicated to perceiving and recognizing other people, integrating these perceptions across

diverse informational sources, and predicting what people are likely to do next.

A large body of research shows that the superior temporal sulcus (STS), which forms the inferior boundary of the superior temporal gyrus in the temporal lobe, plays an important role in these social-cognitive processes. In particular, the posterior end of the STS appears to process perceptual information about bodily action. It responds to the observation of other moving bodies (Allison, Puce, & McCarthy, 2000; Bonda, Petrides, Ostry, & Evans, 1996) and moving faces (Puce, Allison, Bentin, Gore, & McCarthy, 1998). Individuals with deficits in connectivity between other regions and the STS show impairments in perceiving biological motion (Pavlova, Staudt, Sokolov, Birbaumer, & Krageloh-Mann, 2003).

However, the STS region also has a broader function in social cognition. For example, in the macaque, some STS cells change their response rates as a function of the history of observed biological motions (Jellema & Perrett, 2003), and the pattern of these changes suggest that these cells play a role in action prediction. And in addition to responding to aspects of biological motion, the STS is active during evaluation of the intentions behind other people's actions. For instance, Pelphrey, Singerman, Allison, and McCarthy (2003) showed that this region exhibited a stronger hemodynamic response when participants observed an actor direct his or her gaze to a visual stimulus in an incongruent manner (i.e., the actor looked in the direction opposite to an interesting visual event) than when they observed an actor direct his or her gaze in a congruent manner (i.e., the actor looked at the stimulus). Furthermore, individuals with deficits in connectivity between the STS and other regions show impairments in understanding the intentions behind actions (Pavlova, Sokolov, Birbaumer, & Krageloh-Mann, 2008). The posterior STS region even exhibits activity in response to the motion of geometric shapes that are suggestive of underlying

Address correspondence to Brent C. Vander Wyk, Yale University, Child Study Center, 230 South Frontage Rd., New Haven, CT 06520-7900, e-mail: brent.vanderywyk@yale.edu.

intentional activity but do not actually contain biological motion (Castelli, Happé, Frith, & Frith, 2000; Heider & Simmel, 1944). Finally, the STS region, and particularly the posterior STS region in the right hemisphere, exhibits differentially elevated activation when the details of biological motion do not match expectations set up by previous cues (Brass, Schmitt, Spengler, & Gergely, 2007; Pelphrey, Morris, & McCarthy, 2004).

These findings suggest that the STS does not merely represent the surface features of biological motion, but also participates in integrating biological motion, or actions, with the social context. If so, the STS must be sensitive to environmental or cognitive sources of information that inform this context. Such sources might include language, the personality of the perceiver, the ascribed qualities of the perceived (e.g., “How much are they like me?”), and cultural norms. However, an especially salient component of the social context is perceived emotion, particularly when emotion indicates a preference for or against an object or event. Even very young children recognize that another person’s emotional response indicates his or her preference for or against objects in the environment (Phillips, Wellman, & Spelke, 2002; Repacholi & Gopnik, 1997). The goal of the experiment we report in this article was to examine whether the STS exhibits differences in activity that are dependent on a previous emotional context that has provided cues regarding an observed person’s preferences and underlying intentions.

For this experiment, we used a paradigm adapted from Phillips et al. (2002). Participants observed a videotape of an actress who expressed positive or negative regard toward one of two objects. She then reached and picked up that same object or the other one. Viewing the actress’s emotional expression allowed participants to attribute an intention to her: Positive expressions toward an object would warrant the attribution of the intent to pick up the object; negative expressions would warrant the attribution of the intent not to pick up the object. The actress’s subsequent reaching gesture could then be interpreted as being either congruent or incongruent with the intention.

Much of the previous work outlining the brain mechanisms for the analysis of intentions has remained open to the alternative explanation that observed activation differences were due to uncontrolled attentional factors. Indeed, Hopfinger, Buonocore, and Mangun (2000) found that the STS region was involved in the allocation of spatial attention in response to visual cues. Therefore, previously observed differences in STS activity that were associated with differences in intention might have been due to participants having to switch their attention. For this reason, it is important that by using both positive and negative emotions directed toward an object, this study balanced the attentional demands between the congruent and incongruent conditions. When participants viewed an actress who had positive regard toward an object, they had to shift attention twice to understand the incongruent display, in which she reached toward the other object, but they had to shift attention only once

to understand the congruent display, in which she reached for the object she held with positive regard. In the negative case, the incongruent display involved only one shift of attention (as the actress reached for the object toward which she expressed negative regard), whereas the congruent display required two shifts of attention. If results showed that the STS region responded differentially as a function of congruency, this would be consistent with the view that the role of the STS in action understanding is greater than simply representing the visual properties of biological motion, and the differential activation could not be attributed to attentional differences.

## METHOD

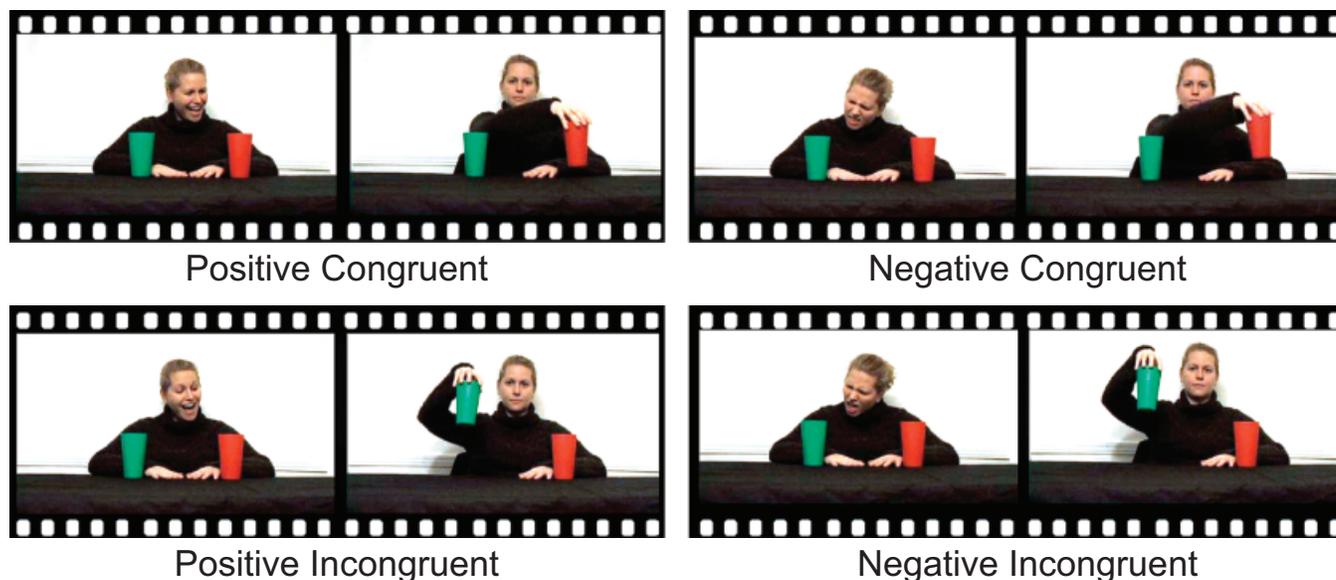
### Participants

Participants were 16 healthy adults (10 male, 6 female; 15 right-handed) ranging in age from 19 to 31 years (mean age = 25 years). All participants had normal or corrected-to-normal vision and were screened for psychiatric and neurological disorders. The project was approved by the Institutional Review Board of Carnegie Mellon University. All participants gave written informed consent to participate and received financial compensation for their time.

### Experimental Design

The experimental design for this event-related functional magnetic resonance imaging study consisted of four conditions: positive-congruent, negative-congruent, positive-incongruent, and negative-incongruent. The conditions are illustrated schematically in Figure 1.

Each trial began with the actress facing forward. She had a neutral expression, and her gaze was directed at the participant. A red and a green cup stood on the surface in front of her. Then, the actress moved her head and eyes to look at one of the cups (whether it was the red or green cup was randomly determined for each trial). While looking at the cup, she smiled with raised eyebrows while saying “Ahhh!” (in the two positive conditions) or frowned with furrowed eyebrows while saying “Yuck!” (in the two negative conditions). This expression of emotion lasted 2 s. The actress then resumed her neutral expression and looked straight ahead, resting for 2 s. Over the following 4 s, she kept her eyes forward and her expression neutral while reaching toward, lifting, and replacing either the cup toward which she had previously directed her attention or the other cup. Thus, she acted in line with (congruent) or in contrast to (incongruent) the expectation created by her positive or negative expression. In this way, four conditions were created: (a) a positive-congruent condition, in which the actress expressed a positive emotion toward an object and then reached for that object; (a) a positive-incongruent condition, in which the actress expressed a positive emotion toward an object and then reached for the other, ignored object; (c) a negative-congruent condition, in which the actress



**Fig. 1.** Schematic illustration of the four conditions in the experiment. In the positive conditions, the actress expressed positive regard toward one of the cups, and in the negative conditions, she expressed negative regard toward one of the cups. In the congruent conditions, the actress's reach was expected, given her previous expression. In the incongruent conditions, her reach was unexpected, given her previous expression.

expressed a negative emotion toward an object and then reached for the other, ignored object; and (d) a negative-incongruent condition, in which the actress expressed a negative emotion toward an object, but then reached to that object.

A single trial lasted 8 s. During the 12-s intertrial intervals, the actress maintained eye contact with the camera and posed a neutral expression. Trials appeared in a pseudorandom order subject to the constraint that no one trial type could occur more than twice in a row. Participants completed four runs, each consisting of 12 s of the actress sitting motionless with a neutral expression, 24 trials (6 in each condition), 16 s of the actress sitting motionless with a neutral expression, and finally 12 s of blank screen (total of 508 s). Stimuli were presented using E-Prime 2.0 software (Psychological Software Tools, Inc., Pittsburgh, PA). Both the cups' positions and the hand used to lift the cup were randomized within each run. Participants were instructed only to watch the video attentively. The analyses presented here included data from 62 runs. Two runs, one from each of 2 participants, were excluded because of motion artifacts resulting from movement of the participants while in the scanner.

### Imaging Protocol

Scanning was performed in a Siemens 3-T Allegra head-only scanner (Siemens, Erlangen, Germany), using an eight-channel head coil, at the Brain Imaging Research Center, Carnegie Mellon University. High-resolution, T1-weighted anatomical images were acquired using an MP-RAGE (magnetization-prepared rapid gradient echo) sequence (repetition time = 1,630 ms, echo time =

2.48 ms, field of view = 20.4 cm, flip angle = 8°, image matrix = 256 × 256, voxel size = 0.8 × 0.8 × 0.8 mm; 224 slices). Whole-brain functional images were acquired using a single-shot, gradient-recalled echoplanar pulse sequence (repetition time = 2,000 ms, echo time = 30 ms, field of view = 20.4 cm, flip angle = 73°, image matrix = 64 × 64, voxel size = 3.2 × 3.2 × 3.2 mm; 35 slices) sensitive to blood-oxygenation-level-dependent (BOLD) contrast. Each run consisted of the acquisition of 254 successive brain volumes, after discarding 2 image volumes to allow for steady-state equilibrium.

### Data Analysis

Data were preprocessed and analyzed using the BrainVoyager QX 1.9 software package (Brain Innovation, Maastricht, The Netherlands). Preprocessing of the functional data included slice time correction (using cubic spline interpolation), alignment of slices (using cubic spline interpolation to the first nondiscarded scan time within a scan run), 3-dimensional motion correction (using trilinear interpolation), spatial smoothing with a 4-mm Gaussian kernel, linear-trend removal, and temporal high-pass filtering (fast-Fourier transform based with a cutoff of 3 cycles/time course). The functional data sets were coregistered to the Talairach-transformed (Talairach & Tournoux, 1988), within-session, T1-weighted anatomical image series to create a 4-dimensional data representation. Estimated motion plots and cine loops were examined for each participant in order to identify movements and eliminate runs in which the participant displayed a deviation in the estimated center of mass (in any dimension) that was greater than 3 mm.

A multiparticipant statistical analysis was performed by multiple linear regression of the time course of the BOLD response in each voxel. The general linear model of the experiment was computed for 62  $z$ -normalized volume time courses (4 runs from 14 participants and 3 runs from 2 participants). Model predictors were defined by convolving an ideal boxcar response with a gamma-function model of the hemodynamic response (Friston et al., 1995). Boxcar values were equal to 1.0 during the actress's 2-s emotional expression and 4-s reach and were 0.0 otherwise. To compare activations among experimental conditions, we performed linear contrasts using  $t$  statistics.<sup>1</sup> Given our a priori hypotheses, we were primarily interested in those regions exhibiting stronger activation in the two incongruent conditions (positive-incongruent, negative-incongruent) than in the two congruent conditions (positive-congruent, negative-congruent). We were also interested in whether the effect of congruency varied as a function of the affect. Regions of interest (ROIs) were identified by calculating the  $t$  contrast for positive-incongruent and negative-incongruent activation being greater than positive-congruent and negative-congruent activation. For multiparticipant statistical maps, we adopted a false-discovery-rate (Genovese, Lazar, & Nichols, 2002) threshold of  $q < .05$ , a procedure that deals with the problem of multiple comparisons by automatically identifying a threshold for statistical significance that ensures that, on average, the proportion of false positives among the activated voxels will be less than  $q$ . Activation maps were visualized on a Talairach-transformed template brain, with only clusters of more than 8 contiguous voxels displayed.

## RESULTS

The contrast test for greater activation on incongruent than on congruent trials revealed only one brain region, localized to the right posterior STS (in Brodmann's Area 22), that exhibited significant differential activity at the threshold ( $q < .05$ ). This region (see the illustrations on the left in Fig. 2) was made up of 301 active voxels, the center of which was located at the following Talairach coordinates:  $x = 57$ ,  $y = -47$ ,  $z = 4$ .

The graphs in Figure 2 illustrate the grand-average waveforms from the voxels comprising this right posterior STS region. As expected, the average response in the incongruent conditions was greater than the average response in the congruent condi-

tions. We computed a series of  $t$  contrasts on the four conditions to more fully explore the pattern of effects within this posterior STS ROI. Further inspection of the interaction term (Emotion  $\times$  Congruency) revealed a significantly greater incongruency effect in the positive-emotion condition than in the negative-emotion condition,  $t(15679) = 2.27$ ,  $p < .05$ . However, the incongruency effect was significant both in the positive condition,  $t(15679) = 5.45$ ,  $p < .001$ , and in the negative condition,  $t(15679) = 2.23$ ,  $p < .05$ . There was also an effect of emotion; responses were greater in the positive condition than in the negative condition, collapsed across congruency,  $t(15679) = 2.87$ ,  $p < .01$ . No brain regions exhibited greater activation on congruent than on incongruent trials.

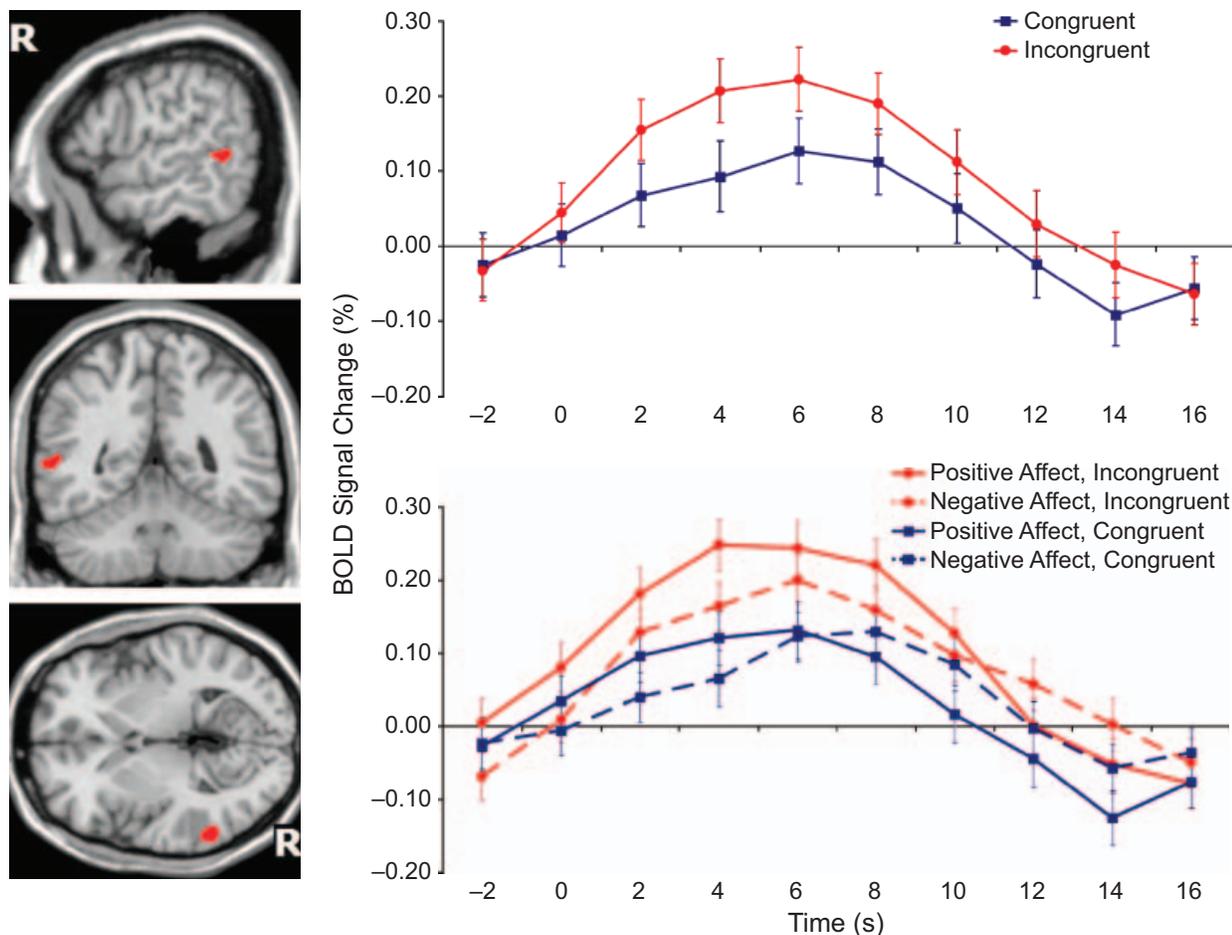
## DISCUSSION

The results of this study indicate that the response of the right posterior STS region to a given biological motion is sensitive to prior emotional context. Specifically, the STS showed a greater response when participants viewed a reach incongruent with a prior emotional expression than when they viewed a reach congruent with the prior expression, regardless of whether expectations were induced by a positive or a negative emotional expression. This finding is in accord with prior research indicating that the posterior STS does more than simply represent the current state of biological motion (e.g., Pelphrey et al., 2003).

The data provide evidence against recent arguments that the earlier findings of increased activity in the posterior STS region during viewing of unexpected actions were related to additional shifts in attention required when observed movements differed from expectations (e.g., Corbetta, Patel, & Shulman, 2008). We addressed this alternative hypothesis by varying expectations using emotional cues of different valence that balanced the number of attention shifts in the congruent and incongruent conditions. In the positive-congruent condition, the actress directed participants' attention toward the desired object, so there was no need for viewers to shift attention when she then, seemingly rationally, picked that object up. There was, however, an additional shift in attention in the negative-congruent condition, because the actress would be expected to pick up the object that she did not previously view with disgust. The positive-incongruent condition also required an additional shift in attention, but the negative-incongruent condition did not. Thus, the congruent and incongruent conditions contained the same total number of required shifts in attention.

By our estimation, the attention-reorienting account would predict greater STS activity in the positive-incongruent than in the negative-incongruent condition, because the former required participants to reorient attention twice, and the latter required only one shift in attention. Likewise, this account would predict greater STS activation in the negative-congruent than in the positive-congruent condition, because the latter involved only one shift in attention, whereas the former demanded two shifts in attention (i.e., shifting

<sup>1</sup>For a  $t$  contrast between two predictors, we performed a paired, two-tailed  $t$  test on the hypothesis that the betas for the two predictors were significantly different. Each functional run (62 in all) was concatenated to create a single time course. For a test on a region of interest, the activation of included voxels was averaged at each time point. For whole-brain analyses, each voxel was tested individually. The time course was modeled with the four experimental variables (positive-congruent, positive-incongruent, negative-congruent, and negative-incongruent), as well as a variable for each functional run less 1, to account for variability in BOLD signal that was not experimentally meaningful. The remaining time points in the concatenated time course (15,679) served as degrees of freedom, used for assessing the standard error of the beta for each predictor.



**Fig. 2.** Experimental results. The brain maps on the left show the location of the right posterior superior temporal sulcus (STS) region in which activation was significantly greater on incongruent than on congruent trials ( $q < .05$ ). “R” indicates the right hemisphere. The graphs show the time course of the percentage change in the blood-oxygenation-level-dependent (BOLD) response from the voxels in that region. The graph at the top presents results for the congruent and incongruent conditions, collapsing across affect conditions, and the graph at the bottom presents results separately for each of the four experimental conditions. The graphs show grand averages, obtained by averaging the percentage signal change of individual voxels across all participants and repetitions. Error bars indicate standard errors of the means. The 0-s time point represents the onset of the reaching motion.

attention to the cup at which the actress directed negative regard and then shifting attention to the cup she picked up). Moreover, this attention account would predict equivalent activation in the positive-incongruent and negative-congruent conditions, as well as equivalent activation in the positive-congruent and negative-incongruent conditions. However, we observed that activity in the STS was greater in the incongruent condition than in the congruent condition regardless of affect—a pattern of effects that cannot be accounted for by the attention-reorienting account.

Although our findings cannot directly address the issue, they may have implications for a related debate concerning the functional role of the temporoparietal junction, a region that lies anatomically close to the STS and that is thought to be involved in attributing mental states to other people (i.e., constructing a theory of mind; Saxe & Wexler, 2005). An argument (similar to the one for the STS) has been made that activations in the temporoparietal junction are due to attention-related factors

(Corbetta et al., 2008; Mitchell, 2007). Our findings suggest the possibility that incongruency effects in that region might be present even when attention is tightly controlled.

We also found a significant effect of affect, such that the STS responded more strongly in the positive-affect condition than in the negative-affect condition, and also showed a greater response to incongruency in the positive condition. The latter effect needs to be interpreted with care, because the role of attention was not balanced within each emotion condition. In fact, in the positive condition, the incongruent case required more attentional shifting than the congruent case, whereas the opposite was true in the negative condition. Nevertheless, this finding is consistent with arguments that a central role of the STS is to process actions that indicate positive outcomes (Paulus, Feinstein, Leland, & Simmons, 2005). However, another possibility is that people simply have more experience with positive affect, and that activation in the STS reflects

that experience. A third possibility is that contrasting positive and negative affect established a context of approach and avoidance, and the STS is more responsive to the former (Pelphrey, Viola, & McCarthy, 2004). A fourth, more extreme possibility is that the functional role of the STS in encoding action is to consider other people's positive affect. Our data are agnostic as to these and other possibilities. Critically, the STS exhibited sensitivity to congruency between the actor's emotional preference and action regardless of the valence of the affect.

The present findings suggest that the STS region does more than simply identify biological motion. Instead, the STS is likely involved in representing biological motion embedded in a social context—in the present case, one established by a prior emotional expression. Most theories of social cognition agree that humans are able to understand and anticipate others' actions by attributing mental states, such as beliefs, desires, or intentions, to others (Gallese & Goldman, 1998; Gopnik & Meltzoff, 1997; Gordon, 1986; Leslie, 1987). Depending on the theory, these internal representations of others' minds are used in processes of rational inference or simulation. Along these lines, the results of our experiment can be interpreted as showing that the STS region participates in the encoding of such attributed intentions. Many theories require such internal representations of other people's minds on the supposition that human beings seem to be intrinsically unpredictable from the surface cues in the environment.

However, recent work suggests that human cognition is highly sensitive to subtle probabilistic environmental cues. For example, young infants have been shown to be sensitive to transitional probabilities among syllables (e.g., Saffran, Aslin, & Newport, 1996) and sequences of actions (Baldwin, Andersson, Saffran, & Meyer, 2007; Kirkham, Slemmer, & Johnson, 2002). It is possible that human beings make use of the social and environmental context to constrain the representations of likely next actions. In the current study, the emotional expressions may have directly constrained the expected future actions represented in the STS, without explicit attribution of mental states. If so, activity differences in the STS would have been due not to differences in representations of the actress's mind, but rather to the presence versus absence of a match between the predicted and observed action (Jellema & Perrett, 2003). Previous work demonstrated increased STS activation when participants viewed an actor lifting a box if the actor was attempting to deceive them about the weight of the box (Grezes, Frith, & Passingham, 2004a) or if the actor had a false belief about the weight of the box (Grezes, Frith, & Passingham, 2004b). In both cases, the researchers interpreted this activity as resulting from the mismatch between predicted and observed action. Representations of the actor's mind, to the extent they were needed for this task, would have been elsewhere in the brain. From this perspective, understanding other people's actions would require a network of brain regions that instantiate various physical and social aspects of the current situation. The STS would play an important role in this network, integrating probabilistic

information from diverse brain regions and representing the ongoing status of perceived biological motion and expected future actions.

A critical direction for future work is to map the structural and functional connectivity among the STS and regions representing other aspects of social cognition. A number of probabilistic constraint-satisfaction models have been fruitfully applied to diverse cognitive domains, such as language (McClelland & Rumelhart, 1986), causal and conceptual knowledge (Gopnik et al., 2004), theory of mind (Baker, Tenenbaum, & Saxe, 2006), and motor action (Jordan & Rumelhart, 1992). Although these models differ on many algorithmic details, they share the ability to learn from and make use of subtle probabilistic features of the environment in a humanlike way. We expect that applying these computational approaches, in combination with identifying patterns of structural and functional connectivity, would lead to a greater understanding of human action understanding.

**Acknowledgments**—We thank Kwan-Jin Jung, Scott Kurdilla, Deborah Viszlay, Stephen Dewhurst, Marina Donovan, and Susan Perlman for their valuable assistance with this research. This research was supported by the National Institute of Mental Health, the John Merck Scholars Fund, Autism Speaks, and the National Science Foundation.

## REFERENCES

- Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences*, 4, 267–278.
- Baker, C.L., Tenenbaum, J.B., & Saxe, R.R. (2006). Bayesian models of human action understanding. *Advances in Neural Information Processing Systems*, 18, 99–106.
- Baldwin, D., Andersson, A., Saffran, J., & Meyer, M. (2007). Segmenting dynamic human action via statistical structure. *Cognition*, 106, 1382–1407.
- Bonda, E., Petrides, M., Ostry, D., & Evans, A. (1996). Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *Journal of Neuroscience*, 16, 3737–3744.
- Brass, M., Schmitt, R.M., Spengler, S., & Gergely, G. (2007). Investigating action understanding: Inferential processes versus action simulation. *Current Biology*, 17, 2117–2121.
- Castelli, F., Happé, F., Frith, U., & Frith, C. (2000). Movement and mind: A functional imaging study of perception and interpretation of complex intentional movement patterns. *NeuroImage*, 12, 314–325.
- Corbetta, M., Patel, G., & Shulman, G.L. (2008). The reorienting system of the human brain: From environment to theory of mind. *Neuron*, 58, 306–324.
- Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.P., Frith, C.D., & Frackowiak, R.S.J. (1995). Statistic parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, 2, 189–210.
- Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 2, 493–501.

- Genovese, C.R., Lazar, N.A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, *15*, 870–878.
- Gopnik, A., Glymour, C., Sobel, D.M., Schulz, L.E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review*, *111*, 3–32.
- Gopnik, A., & Meltzoff, A.N. (1997). *Words, thoughts, and theories*. Cambridge, MA: MIT Press.
- Gordon, R.M. (1986). Folk psychology as simulation. *Mind and Language*, *1*, 158–171.
- Grezes, J., Frith, C.D., & Passingham, R.E. (2004a). Brain mechanisms for inferring deceit in the actions of others. *Journal of Neuroscience*, *24*, 5500–5505.
- Grezes, J., Frith, C.D., & Passingham, R.E. (2004b). Inferring false beliefs from the actions of oneself and others: An fMRI study. *NeuroImage*, *21*, 744–750.
- Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology*, *57*, 243–259.
- Hopfinger, J.B., Buonocore, M.H., & Mangun, G.R. (2000). The neural mechanisms of top-down attentional control. *Nature Neuroscience*, *3*, 284–291.
- Jellema, T., & Perrett, D.I. (2003). Perceptual history influences neural responses to face and body postures. *Journal of Cognitive Neuroscience*, *15*, 961–971.
- Jordan, M.I., & Rumelhart, D.E. (1992). Forward models: Supervised learning with a distal teacher. *Cognitive Science*, *16*, 307–354.
- Kirkham, N.Z., Slemmer, J.A., & Johnson, S.P. (2002). Visual statistical learning in infancy. *Cognition*, *83*, B35–B42.
- Leslie, A.M. (1987). Pretense and representation: The origins of “theory of mind.” *Psychological Review*, *94*, 412–426.
- McClelland, J.L., & Rumelhart, D.E. (Eds.). (1986). *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 2. Psychological and biological models*. Cambridge, MA: MIT Press.
- Mitchell, J.P. (2007). Activity in right temporo-parietal junction is not selective for theory-of-mind. *Cerebral Cortex*, *10*, 262–271.
- Paulus, M.P., Feinstein, J.S., Leland, D., & Simmons, A.N. (2005). Superior temporal gyrus and the insula provide response and outcome-dependent information during assessment and action selection in a decision-making situation. *NeuroImage*, *25*, 607–615.
- Pavlova, M., Sokolov, A.N., Birbaumer, N., & Krageloh-Mann, I. (2008). Perception and understanding of others' actions and brain connectivity. *Journal of Cognitive Neuroscience*, *20*, 494–504.
- Pavlova, M., Staudt, M., Sokolov, A., Birbaumer, N., & Krageloh-Mann, I. (2003). Perception and production of biological movement in patients with early periventricular brain lesions. *Brain*, *126*, 692–701.
- Pelphrey, K.A., Morris, J.P., & McCarthy, G. (2004). Grasping the intentions of others: The perceived intentionality of an action influences activity in the superior temporal sulcus during social perception. *Journal of Cognitive Neuroscience*, *16*, 1706–1716.
- Pelphrey, K.A., Singerman, J.D., Allison, T., & McCarthy, G. (2003). Brain activation evoked by perception of gaze shifts: The influence of context. *Neuropsychologia*, *41*, 156–170.
- Pelphrey, K.A., Viola, R.J., & McCarthy, G. (2004). When strangers pass: Processing of mutual and averted social gaze in the superior temporal sulcus. *Psychological Science*, *15*, 598–603.
- Phillips, A.T., Wellman, H.M., & Spelke, E.S. (2002). Infants' ability to connect gaze and emotional expression to intentional action. *Cognition*, *85*, 53–78.
- Puce, A., Allison, T., Bentin, S., Gore, J.C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *Journal of Neuroscience*, *18*, 2188–2199.
- Repacholi, B., & Gopnik, A. (1997). Early understanding of desires: Evidence from 14- and 18-month-olds. *Developmental Psychology*, *33*, 12–21.
- Saffran, J.R., Aslin, R.N., & Newport, E.L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.
- Saxe, R., & Wexler, A. (2005). Making sense of another mind: The role of the right temporo-parietal junction. *Neuropsychologia*, *43*, 1391–1399.
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system: An approach to cerebral imaging*. New York: Thieme Medical.

(RECEIVED 8/5/08; REVISION ACCEPTED 11/13/08)