Behavioral/Systems/Cognitive

Object-Based Attentional Modulation of Biological Motion Processing: Spatiotemporal Dynamics Using Functional Magnetic Resonance Imaging and Electroencephalography

Ashley S. Safford,[1,2] **Elizabeth A. Hussey,**[2] **Raja Parasuraman,**[2] **and James C. Thompson**[2]

[1]Graduate Program in Neuroscience and [2]Department of Psychology, George Mason University, Fairfax, Virginia 22030

Although it is well documented that the ability to perceive biological motion is mediated by the lateral temporal cortex, whether and when neural activity in this brain region is modulated by attention is unknown. In particular, it is unclear whether the processing of biological motion requires attention or whether such stimuli are processed preattentively. Here, we used functional magnetic resonance imaging, high-density electroencephalography, and cortically constrained source estimation methods to investigate the spatiotemporal effects of attention on the processing of biological motion. Directing attention to tool motion in overlapping movies of biological motion and tool motion suppressed the blood oxygenation level-dependent (BOLD) response of the right superior temporal sulcus (STS)/middle temporal gyrus (MTG), while directing attention to biological motion suppressed the BOLD response of the left inferior temporal sulcus (ITS)/MTG. Similarly, category-based modulation of the cortical current source density estimates from the right STS/MTG and left ITS was observed beginning at ∼450 ms following stimulus onset. Our results indicate that the cortical processing of biological motion is strongly modulated by attention. These findings argue against preattentive processing of biological motion in the presence of stimuli that compete for attention. Our findings also suggest that the attention-based segregation of motion category-specific responses only emerges relatively late (several hundred milliseconds) in processing.

Introduction

The human body in action produces a complex pattern of motion, with multiple points of articulation and many degrees of freedom. Despite this complexity, human observers show an impressive ability to recognize the movements and actions of others (Johansson, 1973; Blake and Shiffrar, 2007). Considering the ecological significance of biological motion, it would seem reasonable to expect that processing should be efficient and act in an automatic, preattentive manner. While there is substantial evidence that the recognition of biological motion is rapid (Mather et al., 1992), highly robust to noise (Neri et al., 1998), and may involve mechanisms present at birth (Simion et al., 2008), the role of attention is less clear. Computational modeling of the perception of human movements has largely assumed it to be a bottom-up process (Giese and Poggio, 2003). In contrast, a number of behavioral studies have suggested that perception of biological motion might require top-down processing (Cavanagh et al., 2001; Thornton et al., 2002; Parasuraman et al., 2009). In addition, a study by Pavlova et al. (2006) reported that magnetoencephalographic (MEG) responses to biological motion were modulated by attention (Pavlova et al., 2006). These findings indicate that attention might be important for the processing of

biological motion, but provide only limited information about the neural mechanisms involved.

In the present study, we examined the spatiotemporal effects of attention on the neural processing of biological motion by using a modified, dynamic version of the "double-exposure" paradigm (O'Craven et al., 1999; Furey et al., 2006). Using overlapping static images of faces and houses, previous studies have revealed that the response in category-specific cortex (e.g., fusiform face area) was suppressed when attention was directed toward the nonpreferred category (e.g., houses). To illustrate the importance of examining the timing of neural activity (as opposed to only its cortical localization), a recent study used the double-exposure paradigm to show that an early MEG response to faces was not modulated when attention was directed toward houses (Furey et al., 2006). In contrast, attention modulated a later MEG response to both faces and houses. These findings suggest that certain classes of stimuli, particularly those that we have had extensive experience viewing and are ecologically important, such as faces and possibly also biological motion, might undergo initial preattentive processing followed by later, attention-modulated processing.

To examine the contribution of attention to the processing of biological motion, we had participants view movies consisting of overlapping point-light biological motion and point-light tool motion. Each motion category could only be segregated through the integration of the form and motion cues that defined the category. Consequently, we were able to characterize the effects of attention at the level of object recognition (Furey et al., 2006; Peelen et al., 2009). We examined the spatiotemporal effects of
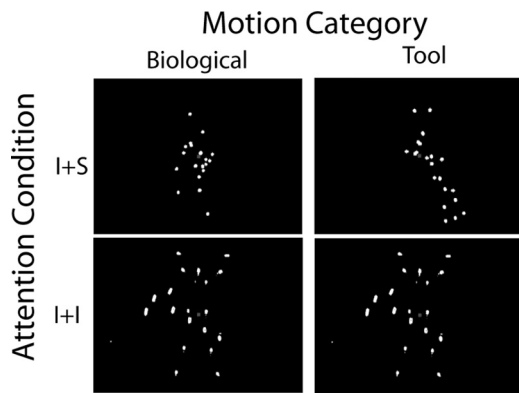
Figure 1. Static illustration of the four stimulus conditions. Stimuli were point-light animations of human or tool motion presented during the fMRI and EEG testing sessions. Videos of these stimuli can be found in the online supplemental material, movies S1–S3.

attention on the processing of biological motion using functional magnetic resonance imaging (fMRI) and surface-based analyses, combined with high-density electroencephalography (EEG) and novel cortical source localization methods.

Materials and Methods

Participants. Thirteen healthy individuals (4 males; age range = 18–35 years; mean = 22.6, SD = 4.1) participated in both the fMRI and EEG components of this study during separate testing sessions. The order of the fMRI and EEG studies was counterbalanced across participants to reduce practice effects. All participants were right handed with normal or corrected-to-normal vision. Participants were compensated $15 per hour and provided written informed consent in accordance with the Human Subjects Review Board at George Mason University (Fairfax, VA). Three of the participants are authors.

Stimuli and task. Visual stimuli, point-light animations of human motion, were created by videotaping an actor who was dressed in black clothing with points of light affixed to the head and major joints (shoulders, elbows, wrists, hips, knees, ankles) in a dark room. Actions performed included performing jumping jacks, walking up stairs, sitting up, kicking right, kicking left, bending over to touch toes, and walking in place. Similarly, point-light-animations of tool motion were created by placing lights on several tools and moving them in an appropriate manner. The tools used included scissors, pitcher, broom, hammer, tongs, saw, and pliers. Adobe Premiere Pro 2.0 (Adobe Systems) was used to edit the videos so that they were of uniform length, and the tools and humans were of relative size. Scrambled (Scram) versions of the biological motion and tool motion videos were created in Matlab (MathWorks). The motion of each light point was tracked on a frame-by-frame basis using a luminance-based clustering algorithm. The starting point, orientation, and temporal phase of each point were then scrambled, reapplied to the points, and converted into movie files. The tool and human videos were then overlaid with each other or the scrambled versions of the other type of motion, and a red central fixation cue was added (Fig. 1) (see supplemental movies S1–S3, available at www.jneurosci.org as supplemental material).

Stimuli were presented during neuroimaging data acquisition using Presentation software (Neurobehavioral Systems). Participants' attention was directed toward either human or tool motion in the single and overlapping stimuli by a 1-back task, creating four experimental conditions: attend to biological motion (Biological$_{Intact}$ plus Tool$_{Scram}$ and Biological$_{Intact}$ plus Tool$_{Intact}$) and attend to tool (Tool$_{Intact}$ plus Biological$_{Scram}$ and Tool$_{Intact}$ plus Biological$_{Intact}$). Participants were asked to respond with a button press when they saw a repeat in the motion category to which they had been instructed to attend. During fMRI, 2 s videos with 1 s interstimulus interval (ISI) were presented in a block design with separate blocks for each condition. There was a 50% probability that a block would contain a repeating stimulus from the attended category. Eight 24 s blocks were presented in each of 6 runs with 12 s blank periods

between each block for a total of 48 blocks, 24 in each of the 4 conditions. For the EEG component, 2 s videos were presented with a 1 s ISI during each run of 64 trials (12.5% probability of a repeat), with a single condition used in each run and three runs per condition (12 runs total). Button presses were recorded for behavioral analysis. Reaction times (RT) were converted to log(RT) to reduce skewness, and hits (H) and false alarms (FA) were converted to sensitivity (d′) using the formula: $d' = z(H) - z(FA)$ (Green and Swets, 1966).

fMRI data acquisition and analysis. fMRI data were collected using a Siemens Allegra 3T scanner at the Krasnow Institute for Advanced Study at George Mason University. Visual stimuli were displayed on a rear projection screen and viewed by participants on a head coil-mounted, angled mirror. The following parameters were used to acquire functional gradient-echo, echoplanar imaging scans: 33 axial slices (4 mm slice thickness; 1 mm gap), repetition time (TR)/echo time (TE) = 2000/30 ms, flip angle = 90, 64 × 64 matrix with 3.75 × 3.75 mm in-plane resolution, field of view = 24 cm. In each run 184 volumes were collected. At the end of the fMRI scanning session, two T1 whole-head anatomical structural scans were collected using a three-dimensional, magnetization-prepared, rapid-acquisition gradient echo (MPRAGE) pulse sequence (160 1-mm-thick slices, 256 × 256 matrix, field of view = 260 mm, 0.94 mm voxels, TR/TE = 2300/3 ms).

Cortical surfaces were reconstructed from the two MPRAGE scans using FreeSurfer software (surfer.nmr.mgh.harvard.edu/). This automated processing involves motion correction, averaging of the two images, removal of nonbrain tissue, intensity normalization, and segmentation to create a representation of the pial surface. The pial surface model was also inflated to support visualization of activation occurring within cortical sulci.

Preprocessing of fMRI data included removal of the first four volumes from each run to compensate for the time it took to reach equilibrium magnetization. The FEAT (fMRI Expert Analysis Tool) software tool of the FSL (fMRI of the Brain Software Library) toolbox (www.fmrib.ox.ac.uk/fsl/) was used for fMRI analysis. The fMRI time series were high-pass filtered at 128 s, motion corrected, and intensity normalized. No smoothing was applied at this stage of analysis. For each run, the onset and duration of each block were modeled, creating four regressors (one for each condition) that were convolved with a gamma function (SD = 3; lag = 6) to estimate the response to the stimuli separately for each of the four conditions. In addition, the temporal derivative and parameters from motion correction were added to the model. Prewhitening was also used to remove temporal autocorrelation of the fMRI time series. Contrast-of-the-parameter estimate (COPE) images were calculated, and the estimates were averaged over the six functional runs. The COPE images were then projected onto the FreeSurfer-generated surface of each individual, transformed into Talairach space, and smoothed with an 8 mm full width at half-maximum (FWHM) Gaussian kernel. A surface-based mixed effects ANOVA with fixed factors of category (Biological vs Tool) and attention (Intact plus Intact vs Intact plus Scram) and participants as a random effect was conducted. Results were viewed on the average inflated surface with a height threshold of $p < 0.001$ and a cluster size threshold of $p < 0.05$ (corrected for multiple comparisons using Monte Carlo simulation). Plots of significant effects were created by averaging parameter estimates for each condition from each participant within circular regions of interest (ROI) with radii of 4 mm (∼1 SD of the smoothing kernel) centered on the peak vertex within each significant cluster. The parameter estimates from these ROIs were then averaged across participants, and within-subjects SEs associated with the significant ANOVA effect were calculated using the method described by Loftus and Masson (1994). It should be noted that the error bars from the fMRI data were used only to illustrate the significant ANOVA effects and were not used for inference because of concerns with inferential circularity.

Event-related potential acquisition. High-density EEG was recorded from 116 scalp channels and horizontal and vertical electro-olfactogram channels relative to a centroid reference using an electrocap (Neuroscan). A Synamps2 system (Compumedics) was used to record data at a sampling rate of 500 Hz with a bandpass filter of 0.1–100 Hz and an A-D digitization rate of 500 Hz. Following data acquisition, the locations of each electrode and the three major fiducial points (left and right periau-

ricular points and nasion) were digitized using a Polaris optical camera (Northern Digital) with Brainsight software (Rogue Research). Importantly, during digitization a reference was attached to the participant's head so that head movements would not result in inconsistent measurements. The digitized locations were used to coregister the electrodes to the anatomical MRI scans during source localization.

Analysis with EMSE. EEG data processing and analysis were performed using EMSE Suite software (Source Signal Imaging). Data were visually inspected and bad channels were removed; at most, two channels were removed for an individual participant. Next, the remaining channels were rereferenced to the common average and an infinite impulse response temporal bandpass filter (0.1–40 Hz) was applied. Epochs based on video onset were created to include 200 ms of prestimulus baseline and 598 ms poststimulus. Trials with artifact greater than ±100 μV based on the horizontal and vertical electro-oculogram channels were rejected, as were trials that included the behavioral responses, before the remaining trials were averaged together by condition.

Sensor space analysis. Visual inspection revealed that the P1/N1 response was maximal over posterior locations, consistent with previous studies of the response to biological motion (Baccus et al., 2009). To analyze the scalp sensor data, we divided the 116 channels into 7 approximately equally sized groups: left anterior lateral, left posterior lateral, middle anterior, central, middle posterior, right anterior lateral, and right posterior lateral. The response in each condition was averaged across electrodes within each of these groups, and a mixed-effects ANOVA with fixed factors of category (Biological vs Tool) and attention (Intact plus Intact vs Intact plus Scram) and participants as a random effect was conducted at each time point across the entire epoch. Results were considered significant if $p < 0.001$ for at least 26 ms (13 consecutive time points).

Source space analysis. We estimated the cortically constrained current source density (CCSD) of electrical responses to each of the four conditions using methods similar to those used previously (Appelbaum et al., 2006; Hamalainen and Ilmoniemi, 1994). For each individual participant, a boundary element model (BEM) was created to estimate the electrical conductivity of the head as part of the source estimation procedure. Compartmentalized tissue segmentations defined regions for the cortex, inner skull, outer skull, and scalp and were the basis for the head BEM. The high resolution-averaged and motion-corrected FreeSurfer image underwent gradient correction in EMSE and was coregistered to the digitized electrode positions by selecting corresponding points on the MRI image for the fiducial points and the scalp surface. Voxel intensities were used to approximate CSF and white matter cortical tissue volumes from manually selected start points. The contiguous cortical gray matter surface was specified by the resulting white matter tissue boundaries and subsequently used in an expansion algorithm to determine the inner and outer skull boundaries. A smooth scalp surface was created by manually removing extraneous noise occurring outside the scalp. The Mesh Generation Wizard in EMSE was used to create a mesh for each of these segments to use in the BEM. Conductivities (Ω^{-1}/m^{-1}) were scalp = 0.33, skull = 0.0042, brain/CSF = 0.33. Finally, the cortical surface produced by EMSE was replaced with a more accurate FreeSurfer-generated representation of the pial surface to allow greater correspondence with fMRI results. The number of nodes in the cortical surface was reduced by a factor of 6 in Matlab before loading it into EMSE, resulting in a dipole at approximately every 6 mm.

CCSD estimates of neural activity, based on the scalp potential recordings, were modeled using the L2 minimum norm estimate (MNE) procedure without regularization (Hamalainen and Ilmoniemi, 1994). In this procedure, neural activity on the cortical surface is estimated using linear optimization to determine the magnitudes of a set of distributed dipoles that together have the lowest root-mean-squared power and at the same time are consistent with the scalp recordings (Michel et al., 2004). The location of the dipoles was restricted to the cortical surface using the reduced FreeSurfer cortical surface described above. The orientation of the dipoles was fixed to be orthogonal to the pial surface. In addition, lead field normalization was applied to compensate for the bias of the MNE method toward superficial sources (Crowley et al., 2000). This procedure yields an inverse matrix that described a linear relation-

ship between the measured scalp activity and the estimated cortical current densities.

To increase the signal-to-noise ratio (SNR) of the data and reduce the amount of data for statistical analyses, each poststimulus event-related potential (ERP) waveform at each electrode was averaged within 23 consecutive, nonoverlapping 26 ms windows. We estimated the signal-to-noise ratio of the windowed scalp potential data for each participant by dividing the maximum deflection in the poststimulus period at each electrode by the SD of the prestimulus period at that electrode and then calculated the mean SNR across electrodes. All SNRs were >5, with 10 of 13 participants having SNRs >10 (supplemental Table S2, available at www.jneurosci.org as supplemental material), which is high for cognitive event-related potential recordings (Regan, 1989). Each windowed time point was then multiplied by the inverse matrix to generate a CCSD estimate on the decimated cortical surface for that time point. This CCSD estimate was subsequently interpolated to the dense (1 mm^2) pial surface for each individual participant. The transform used to map the fMRI data to Talairach space was used to transform the CCSD data into Talairach space, and the data were smoothed with an 8 mm FWHM Gaussian kernel. Then, for each time window, a mixed effects ANOVA with fixed factors of category (biological vs tool) and attention (Intact plus Intact vs Intact plus Scram) and participants as a random effect was conducted, and the results were viewed with a height threshold of $p < 0.001$ and a cluster size threshold of $p < 0.05$ (corrected for multiple comparisons using Monte Carlo simulation). An additional constraint was added such that clusters were only considered significant if they survived threshold for at least 2 consecutive 26 ms time windows. Plots of the time course of significant effects were created by averaging parameter estimates for each time point for each condition from each participant within 4 mm radius circular regions of ROIs centered on the peak vertex within each significant cluster. Mean responses and within-subjects SEs from the ANOVA were calculated using the same method as that used for the fMRI data. As was the case with the fMRI data, the time series from the ROIs centered on the peak vertices from the CCSD clusters were used for illustrative, not inferential, purposes.

To compare fMRI results to CCSDs, we first calculated the Euclidean distance between peak vertices of the closest fMRI and CCSD clusters and the percentage of vertices within CCSD clusters that overlapped with vertices in fMRI clusters. We also extracted CCSD time courses from all of the above threshold vertices within all of the fMRI clusters and from 4 mm circular ROIs at the peak vertices from the right middle temporal gyrus (MTG) and superior temporal sulcus (STS) fMRI clusters. We then calculated mean responses and within-subjects SEs as described above. As the fMRI data used to identify these clusters/ROIs were independent of the CCSD data, one can use the means and SE bars to make inferences about statistical significance.

Results

Behavioral data

Participants were presented with visual stimuli in four conditions during separate fMRI and EEG testing sessions counterbalanced to prevent order effects. Task performance is summarized in supplemental Table S1, available at www.jneurosci.org as supplemental material. Analysis of the RTs of correct responses showed that log(RT)s were significantly faster to the two tool motion categories than for biological motion during the EEG session, $F_{(1,12)} = 14.47$, $p < 0.005$, and the fMRI session, $F_{(1,11)} = 6.70$, $p < 0.05$. Analysis of sensitivity, d', showed that in both the fMRI and EEG sessions there was a significant interaction between motion category and attention condition: fMRI session, $F_{(1,11)} = 27.31$, $p < 0.001$; EEG session, $F_{(1,12)} = 16.10$, $p < 0.01$. In the fMRI session there was also a significant main effect of attention (Intact plus Scram vs Intact plus Intact), $F_{(1,11)} = 42.01$, $p < 0.001$. These effects reflect that performance in the Tool$_{Intact}$ plus Biological$_{Scram}$ condition was significantly better than those in the other three conditions.

Table 1. fMRI results: Talairach position of regions showing main effects of motion category and attention condition and motion category by attention condition

Region	Coordinates			Cluster size (mm²)	Max	Hemisphere
	x	y	z			
Bio > Tool						
STS	43.4	−48.5	14.7	142.51	5.338	RH
MTG	45.1	−59.2	25.2	1133.48	6.111	RH
STG	−58.0	−48.0	16.4	76.54	3.948	LH
Tool > Bio						
aIPS	40.8	−32.2	37.6	170.59	−4.871	RH
medFG	27.1	−77.8	−2.7	667.44	−4.392	RH
ITS/MTG	−43.9	−59.9	3.2	680.35	−8.762	LH
midIPS	−24.3	−58.9	31.7	374.20	−7.829	LH
posIPS	−19.3	−82.0	24.9	341.23	−6.819	LH
medFG	−31.0	−55.6	−5.2	200.69	−4.441	LH
I + I > I + S						
MTG	48.3	−54.2	12.2	207.55	4.172	RH
STS	−40.1	−61.5	22.7	159.50	4.460	LH
Interaction						
Post Central	−35.5	−32.9	51.1	49.65	3.704	LH

Values in the Max column correspond to the maximal value of $-\log_{10}(p)$ significance. Bio, Biological; I, Intact; S, Scrambled; LH, left hemisphere; RH, right hemisphere; med FG, medial fusiform gyrus; midIPS, middle IPS; posIPS, posterior IPS.
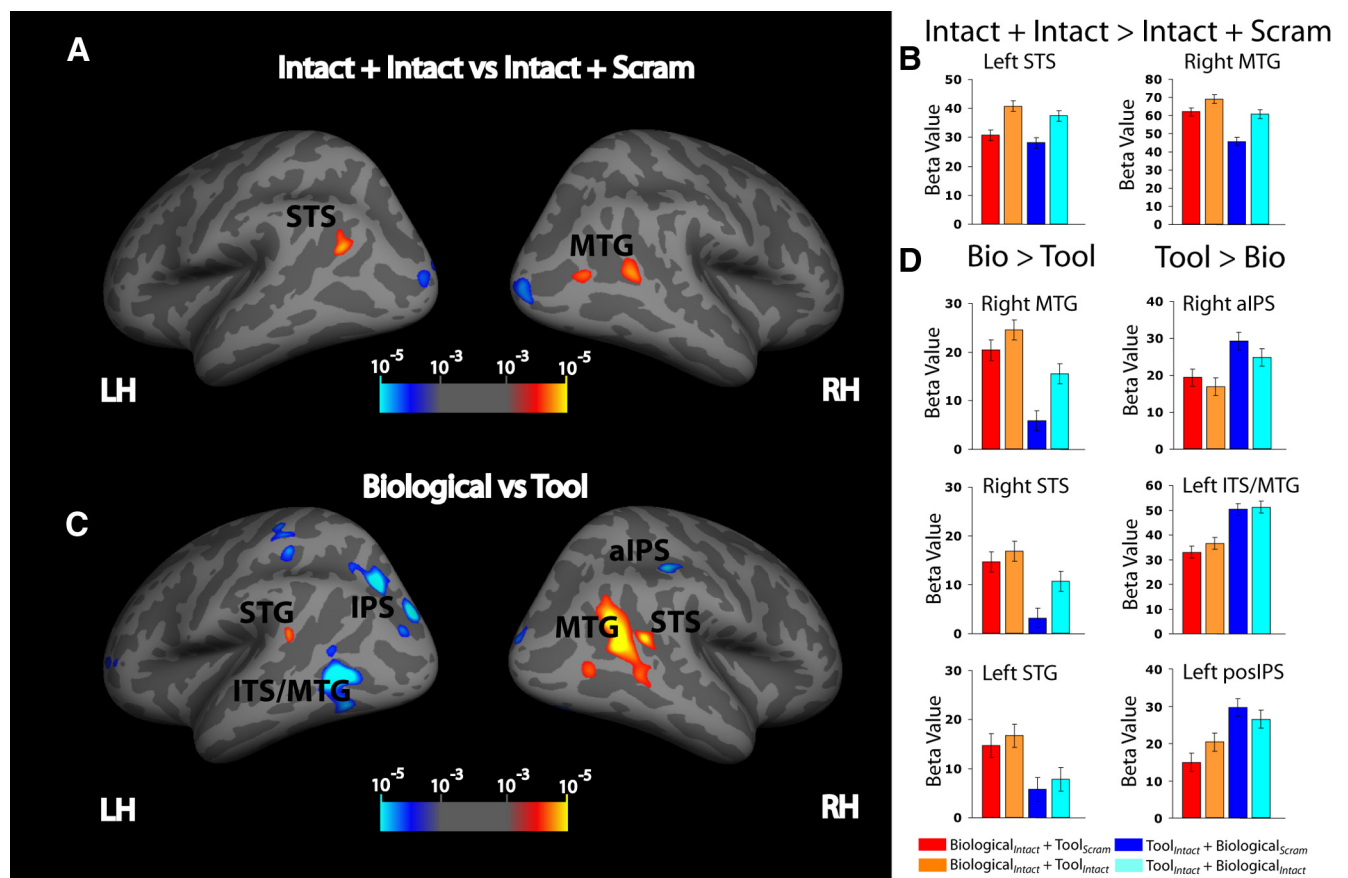

Figure 2. fMRI results. **A**, Displayed are statistical maps of the attention (Intact plus Intact vs Intact plus Scram) main effect projected onto the inflated surface of the FreeSurfer fsaverage brain. Color coding represents $-\log_{10}(p)$ thresholded at $p < 0.001$ uncorrected; cluster $p < 0.05$ corrected. **B**, Plots of beta values for each of the four experimental conditions within a 4 mm radius circular ROI surrounding the local maxima for significant regions. Error bars represent the within-subjects confidence interval. **C, D**, Main effect of motion category. Bio, Biological; LH left hemisphere, RH, right hemisphere posIPS, posterior IPS.

Cortical surface-based analysis of fMRI data

We mapped each participant's fMRI responses onto their cortical surface and then transformed the surface-based fMRI data into standard space for group analysis. Group-based fMRI responses were then analyzed on an average surface using a mixed-effects ANOVA. Cortical regions that showed a significant ($p < 0.001$

uncorrected; cluster $p < 0.05$ corrected) main effect of motion category (Biological vs Tool) are listed in Table 1 and visualized on an inflated average cortical surface in Figure 2C and a pial surface in supplemental Figure S4B (available at www.jneurosci. org as supplemental material). As expected, the BOLD response of regions in the lateral temporal cortex showed a preference for

motion category, and, as has been previously reported, these responses showed a degree of hemispheric lateralization (Beauchamp et al., 2002, 2003). There was significantly greater activation of the biological motion relative to tool motion conditions in right STS, right MTG, and left superior temporal gyrus (STG) ($p < 0.001$ uncorrected; cluster $p < 0.05$ corrected) for both the Intact plus Scram and Intact plus Intact conditions. In right anterior intraparietal sulcus (aIPS), left inferior temporal sulcus (ITS)/MTG, left middle intraparietal sulcus (IPS), left posterior IPS, and bilateral medial fusiform gyrus a significantly greater response was observed to the Intact plus Scram and Intact plus Intact tool motion conditions compared with the Intact plus Scram and Intact plus Intact biological motion conditions ($p < 0.001$ uncorrected; cluster $p < 0.05$ corrected) (Fig. 2C,D).

Importantly, these responses were dependent on selectively attending to the preferred motion category; when biological motion was present but attention was directed to the tool motion (Tool Intact plus Biological Intact condition), the responses in the biological motion preferring regions (e.g., right STS, right MTG, and left STG) were significantly lower compared with when attention was directed toward biological motion (Biological Intact plus Tool Scram and Biological Intact plus Tool Intact conditions) ($p < 0.001$ uncorrected; cluster $p < 0.05$ corrected) (Fig. 2D). Likewise, when tool motion was present but attention was directed to the biological motion (Biological$_{Intact}$ plus Tool$_{Intact}$ condition), the responses in tool motion-preferring regions were significantly lower compared with when attention was directed toward tool motion (Tool$_{Intact}$ plus Biological$_{Scram}$ and Tool$_{Intact}$ plus Biological$_{Intact}$ conditions) ($p < 0.001$ uncorrected; cluster $p < 0.05$ corrected) (Fig. 2D). In addition to showing a greater response to biological motion than to tool motion, a portion of the right MTG also showed a significant main effect of attention (Intact plus Intact vs Intact plus Scram) ($p < 0.001$ uncorrected; cluster $p < 0.05$ corrected), as is shown in Figure 2A and Table 1 as well as on an average pial surface in supplemental Figure S4A (available at www.jneurosci.org as supplemental material). The other area that showed a preference to the Intact plus Intact condition compared with the Intact plus Scram conditions regardless of motion category was the left STS ($p < 0.001$ uncorrected; cluster $p < 0.05$ corrected) (Fig. 2A,B).

In the Biological$_{Intact}$ plus Tool$_{Intact}$ condition and the Tool$_{Intact}$ plus Biological$_{Intact}$ condition, the relative size of the moving human figure and the tool were approximately equal so that they would occupy the same area of visual field. As a result, this could have produced the appearance of a relative depth difference between the two objects; in the real world, if some of the tools (e.g., hammer, scissors) appeared as big as a human, they would be closer to the observer than the human. This potential depth effect was not present in the Biological$_{Intact}$ plus Tool$_{Scram}$ condition or the Tool$_{Intact}$ plus Biological$_{Scram}$ condition, as each object was presented in the presence of a scrambled version of the other. It is well known that attention can be directed in depth (z) as well as in planar (x-y) space (Downing and Pinker, 1985), although there are differences in attention allocation between x, y, and z dimensions (Anderson and Kramer, 1993). If the observed effects in right STS/MTG or left ITS/MTG were caused by the attentional selection of the distant object or the closer object, respectively, we would expect that the difference between Biological$_{Intact}$ plus Tool$_{Intact}$ and Tool$_{Intact}$ plus Biological$_{Intact}$ would be greater than the difference between Biological$_{Intact}$ plus Tool$_{Scram}$ and Tool$_{Intact}$ plus Biological$_{Scram}$ (i.e., a motion category × attention interaction). However, no such interaction was observed in the right STS/MTG cluster or the left ITS/MTG cluster, even when the

height threshold was lowered to $p < 0.05$ uncorrected. These results indicate that variation in depth-based attentional processing cannot account for the observed pattern of activations in the superior and inferior banks of the temporal cortex.

There was a significant interaction between motion category and attention condition observed in a region of the left postcentral gyrus ($p < 0.001$ uncorrected; cluster $p < 00.05$ corrected) (supplemental Figure S1A, available at www.jneurosci.org as supplemental material)). This interaction reflected a lower response in left motor cortex for the Tool$_{Intact}$ plus Biological$_{Scram}$ condition relative to the other three conditions (supplemental Figure S1B, available at www.jneurosci.org as supplemental material). This result might reflect the behavioral finding that right-handed participants responded more accurately in this condition relative to the other three conditions.

Sensor-space analysis of ERPs
ERPs were recorded to measure the temporal dynamics of attentional effects in the biological motion processing stream. Topographic maps from time points corresponding to the P1 peak latency (134 ms), the N1 peak latency (212 ms), and a later time point (438 ms) are shown in Figure 3A. The grand mean average of the four conditions from left and right posterior lateral electrode groups are shown in Figure 3, B and C. Results from a mixed effects ANOVA performed at each time point revealed a significant ($p < 0.001$ for at least 26 ms) main effect of attention only during the interval from 192 to 244 ms following stimulus onset. This effect was significant for the left anterior lateral, middle anterior, middle posterior, and right posterior lateral electrode groups, reflecting a greater posterior N1 (or it's positive reflection anteriorly) in Intact plus Intact conditions, relative to the Intact plus Scram conditions. Neither the effect of stimulus category nor the attention × category interaction reached significance ($p > 0.001$) at any time point.

Source-based analysis of ERPs
To examine the spatiotemporal distribution of the effects of attention on the neural responses to biological and tool motion, we used cortically constrained MNE. The CCSD estimates from the four conditions were broadly distributed over posterior lateral temporal and parietal cortex, as can be seen in the mean CCSDs from two time points presented in supplemental Figure S2 (available at www.jneurosci.org as supplemental material). The statistical images presented here thus represent locations in which the experimental effects were most reliable across our group of participants. Table 2 lists the coordinates of cortical regions that showed significant ($p < 0.001$ uncorrected; cluster $p < 0.05$ corrected) main effects of motion category (Biological vs Tool) or attention condition (Intact plus Scram vs Intact plus Intact) in the CCSD analysis, as well as the distance to the nearest peak coordinate of fMRI activation and the percentage of the CCSD cluster that is overlapping with the fMRI cluster. These effects are visualized on the inflated surface at the time point of greatest significance in Figure 4, A and B, and on the pial surface in supplemental Figure S4, C and D. At an early time point, ~200 ms following stimulus onset, a significant main effect of attention was found in right MTG, right aIPS, and left IPS ($p < 0.001$ uncorrected; cluster $p < 0.05$ corrected). A similar effect was seen at a later time point (~300 ms) in left superior occipital gyrus ($p < 0.001$ uncorrected; cluster $p < 0.05$ corrected). Figure 4, C–F, shows the time courses derived from the regions right MTG, left IPS, right STS, and left ITS. The patterns of activity between the four conditions were similar in time courses from the other
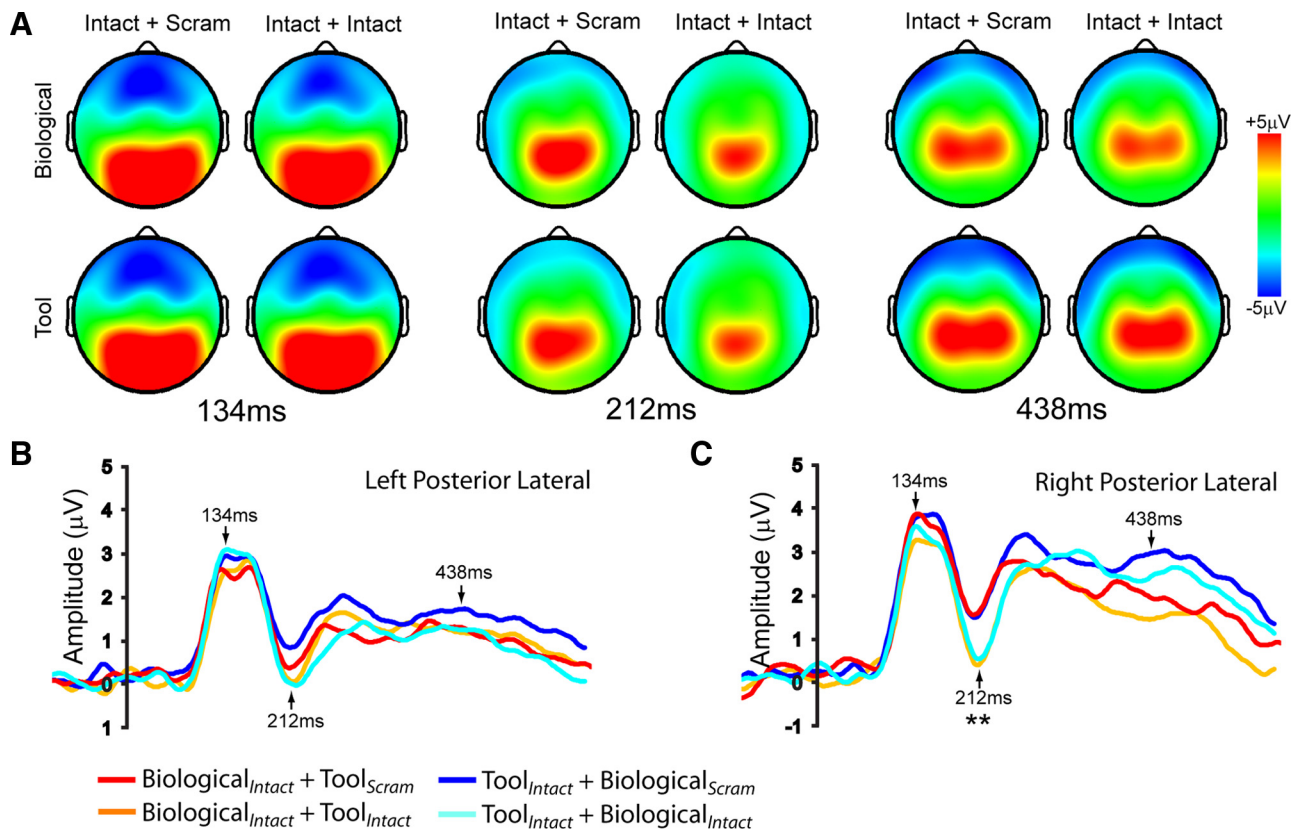
Figure 3. Sensor-space analysis of ERPs. **A**, Topographic maps of the four conditions at time points corresponding to the P1 peak latency (134 ms), the N1 peak latency (212 ms), and a later time point (438 ms). **B**, Plots display the grand mean average ERPs following stimulus onset from the right and left posterior lateral electrode groups.

Table 2. CCSD results: Talairach position of regions showing main effects of motion category and attention condition

Region	Coordinates			Cluster size (mm²)	Max	Hemi	Time	Distance	% Overlap
	x	y	z						
Bio vs Tool									
STS	46.0	−42.2	17.1	114.50	−4.479	RH	478	7.23	8
MTG	46.9	−54.1	11.8	398.70	−7.581	RH	452	14.45	88
aIPS	34.0	−30.9	47.0	491.21	−6.088	RH	530	11.67	0
ITS	−47.4	−49.2	−9.5	54.31	−3.275	LH	478	16.97	0
I + I vs I + S									
MTG	40.7	−61.1	25.1	101.63	−3.911	RH	218	16.49	0
aIPS	33.0	−31.0	45.9	173.14	−4.174	RH	218	NA	NA
IPS	−19.9	−59.6	48.5	395.12	−6.925	LH	218	NA	NA
SOG	−30.9	−80.6	15.7	154.94	−5.181	LH	322	NA	NA

Values in the Max column correspond to the maximal value of $-\log_{10}(p)$ significance. Values in the Time column are the time points at which the effect was maximal (in milliseconds) from stimulus onset. Distance values represent the Euclidean distance from the nearest peak of fMRI activation. % overlap is the percentage of the CCSD cluster that overlaps with a corresponding fMRI cluster. NA indicates no similar activations in fMRI. I, Intact; S, scrambled; Bio, biological; Hemi, hemisphere; LH, left hemisphere; RH, right hemisphere; SOG, superior occipital gyrus.

regions, which are presented in supplemental Figure S3 (available at www.jneurosci.org as supplemental material). Additionally, the CCSD for each of the four conditions at the time points of significant effects are shown in supplemental Figure S2 (available at www.jneurosci.org as supplemental material). These bilateral regions showed a large negative component, consistent with the N1 response from the sensor-space analysis, which was greater for the two Intact plus Intact conditions than for the two Intact plus Scram conditions.

Following this early effect of attention, there was also a later, more prolonged, significant main effect of motion category (Biological vs Tool) found in right STS, right MTG, right aIPS, and left ITS ($p < 0.001$ uncorrected; cluster $p < 0.05$ corrected) starting at ~450 ms and continuing until as late as 600 ms. Consistent with the fMRI results, this effect showed a difference in

response between the biological motion and tool motion stimuli for both the Intact plus Intact and Intact plus Scram conditions. Again, this effect was dependent on whether attention was directed toward or away from the preferred stimuli for a given region: in the right STS, the response was significantly less positive when attention was directed to the two biological motion conditions (Biological$_{Intact}$ plus Tool$_{Scram}$ and Biological$_{Intact}$ plus Tool$_{Intact}$) relative to when biological motion was present but attention was directed toward the tool motion (Tool$_{Intact}$ plus Biological$_{Intact}$) ($p < 0.001$ uncorrected; cluster $p < 0.05$ corrected); in left ITS, when attention was directed to biological motion (Biological$_{Intact}$ plus Tool$_{Intact}$), the late response was more positive than when attention was directed toward the tool motion conditions (Tool$_{Intact}$ plus Biological$_{Scram}$ and Tool$_{Intact}$ plus Biological$_{Intact}$) ($p < 0.001$ uncorrected; cluster $p < 0.05$
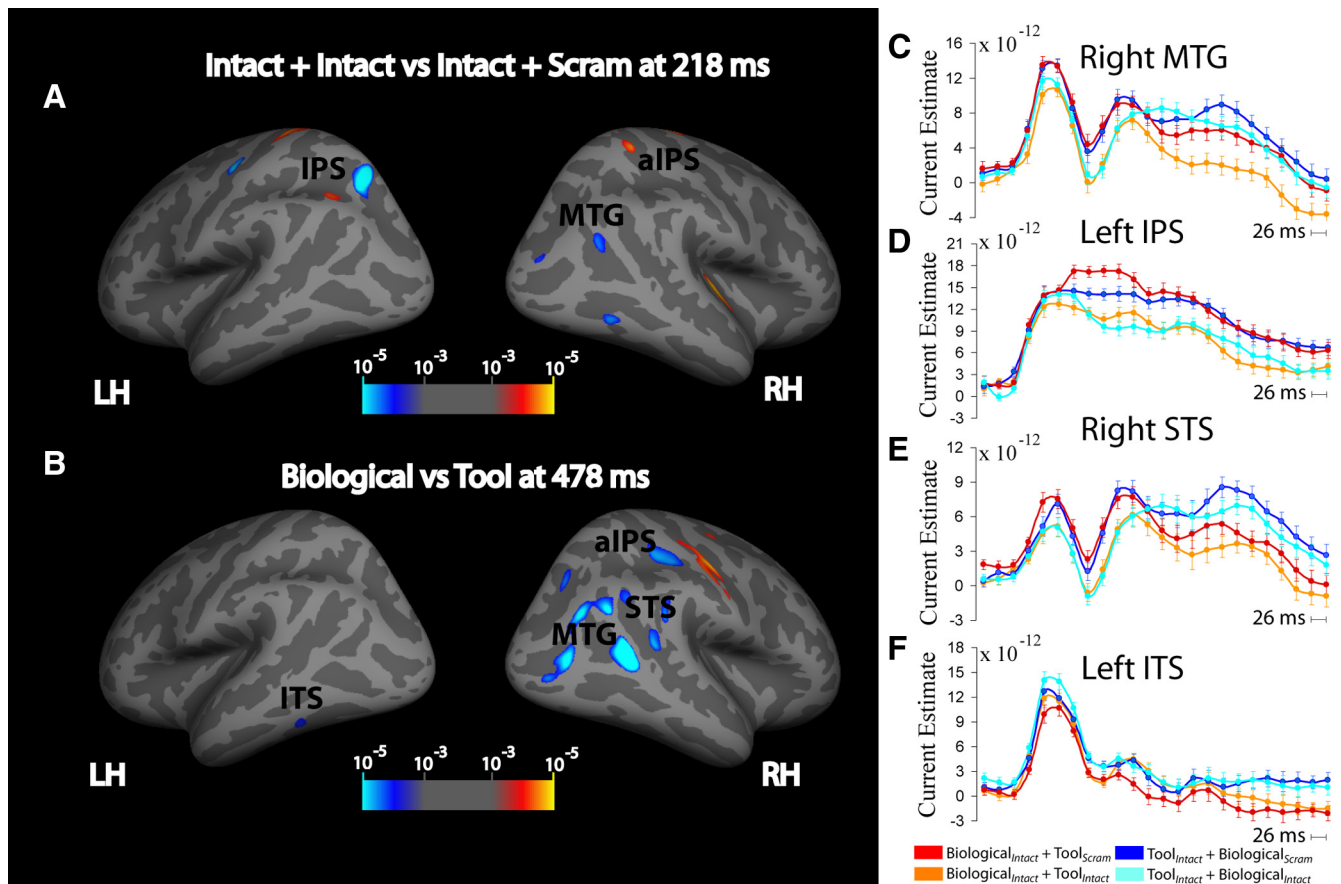
Figure 4. Cortically constrained current source density analysis of attention condition and motion category main effects. **A** and **B**, Attention effects (**A**) and motion category effects (**B**) are displayed. Statistical maps of the between-condition difference in estimated current are displayed on inflated fsaverage surfaces. Color coding represents statistical significance [−log₁₀ (p)] of these differences thresholded at $p < 0.001$ uncorrected; cluster $p < 0.05$ corrected. **C–F**, Time courses of CCSD estimates from each of the four conditions within a 4 mm radius circular ROI surrounding the local maxima for significant regions are shown. Error bars represent the with-in subject confidence interval. The online supplemental material includes time courses for additional regions (supplemental Figure S3, available at www.jneurosci.org as supplemental material) and the CCSD for each of the four conditions at the time points of significant effects (supplemental Figure S2, available at www.jneurosci.org as supplemental material). LH, Left hemisphere; RH, right hemisphere.

corrected). As was the case for the fMRI data, there was no significant category × attention interaction in these regions, indicating that effects were caused by the attentional selection of category-specific information and not by relative depth effects.

Overlap of fMRI and CCSD

While the peak vertices from of a number of fMRI clusters of activation were within 20 mm of peak vertices from clusters identified in the CCSD analysis, such as right aIPS and left ITS/MTG (Table 2), it was only in the right STS/MTG region that overlap was observed (Fig. 5A). In the right MTG, 88% of vertices in the motion category-selective CCSD cluster overlapped with the fMRI cluster that showed the same motion category effect. Similar to the time course from the CCSD peak (Fig. 5B), when the CCSD time course from the whole fMRI cluster was examined we observed the significant main effect of motion category in time points later ∼450 ms (Fig. 5C). In contrast, this effect was not significant at vertices from the peak of the fMRI cluster (Fig. 5D). In the right STS, 8% of vertices in the CCSD cluster that showed a significant main effect of motion category beginning at ∼450 ms overlapped with the fMRI cluster that showed the same motion category effect. While the time course from the CCSD showed the significant main effect of motion category at ∼450 ms (Fig. 5E), the time course from the whole fMRI cluster and the peak of the fMRI cluster did not show this effect (Fig.

5 F, G). Accordingly, the time courses from the fMRI clusters that had no overlap with corresponding CCSD clusters did not show the significant category or attention effects (supplemental Figure S5, available at www.jneurosci.org as supplemental material).

Discussion

Using fMRI, high-density EEG, and cortical source localization methods, we demonstrated that object-based selective attention plays an important role in the neural processing of biological motion. Our fMRI results indicate that the hemodynamic response in cortical regions that prefer biological motion or tool motion is strongly attenuated when attention is directed away from the preferred motion category. Similarly, CCSD analysis of high-density EEG data indicated that the neural response at ∼200 ms reflects the engagement of selective attention, regardless of motion category, in bilateral parietal and right lateral temporal regions. Category-selective CCSD responses that were strongly modulated by attention were observed at time points later than 450 ms in regions consistent with the fMRI responses: both fMRI and late CCSD responses in the right STS/MTG showed a preference for biological motion that was strongly modulated by attention, while the left ITS/MTG showed a preference for tool motion that was also modulated by attention.
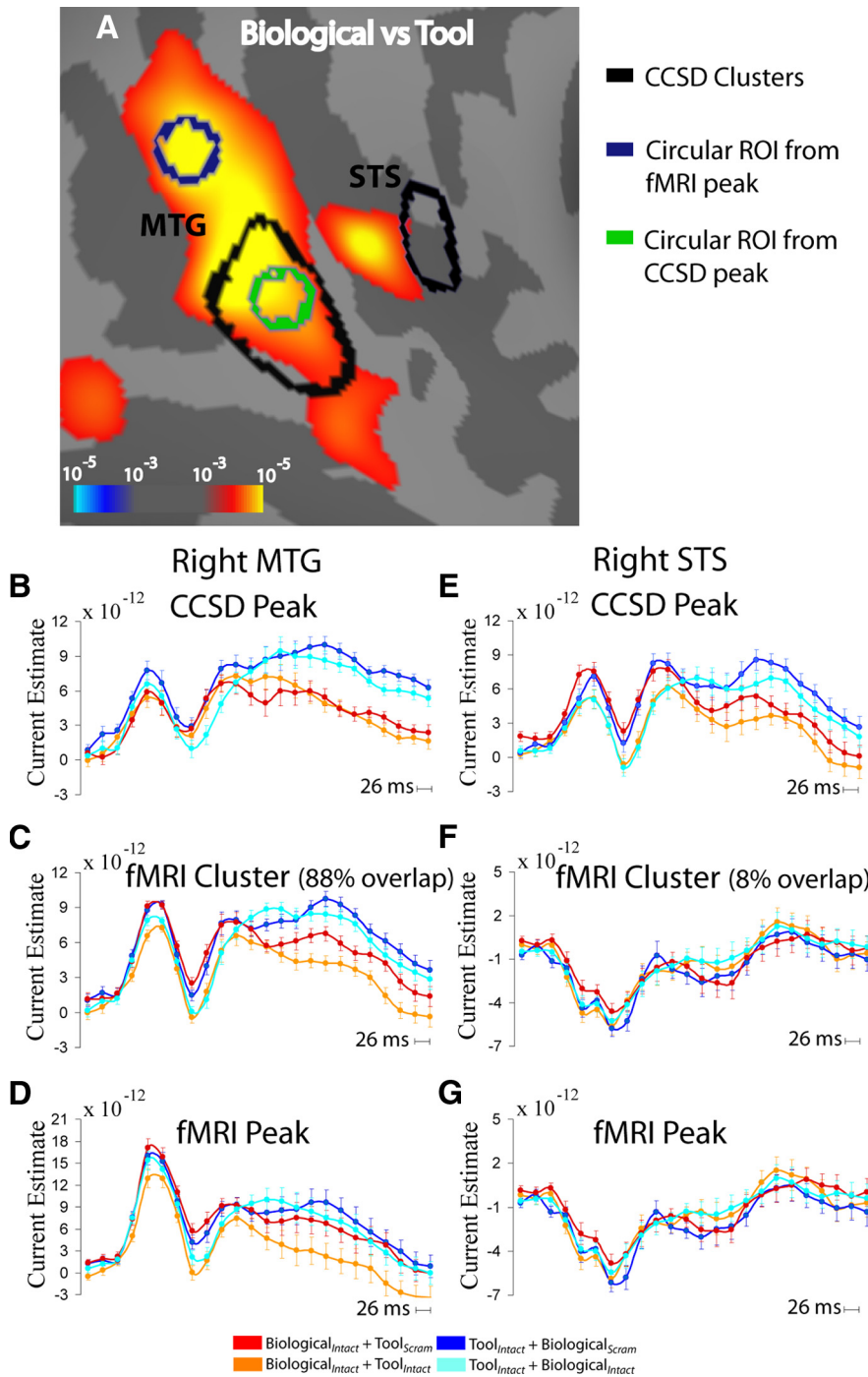
Figure 5. Overlap of fMRI and CCSD regions. *A*, fMRI activation in the lateral temporal cortex for motion category (Biological vs Tool) comparison with CCSD clusters from the same comparison outlined in black. Circular ROIs of 4 mm radius were created surrounding the local maxima and are shown in green (CCSD) and navy blue (fMRI). *B–G*, Time course of CCSD estimates from each of the four conditions within CCSD peak (4 mm radius circular ROI) (*B, E*), fMRI cluster (*C, F*), and fMRI peak (*D, G*) for right MTG (*B–D*) and right STS (*E–G*).

Attentional modulation of the neural response to biological motion

The results of the present study indicate that there is an important contribution of top-down, object-based influences on the neural processing of biological motion. Previous behavioral studies have suggested that top-down control, including attention, plays a role in the processing of biological motion, although the nature of this role has not been clear (Cavanagh et al., 2001; Thornton et al., 2002; Pavlova et al., 2006). The findings of the present study

provide details about the spatiotemporal dynamics of the neural activity underlying this attentional modulation. Our results indicate that selective attention intensifies the response of lateral temporal regions that show a preference for biological motion, in addition to modulating the response of regions that prefer tool motion. When biological motion was present and attention was directed to a different but spatially overlapping object motion category–in this case, tool motion–both the fMRI response and CCSD estimates of electrical activity from the right STS/MTG was considerably suppressed. Likewise, when tool motion was present but attention was directed toward overlapping biological motion, the fMRI and CCSD response of the left ITS/MTG was reduced. The modified double-exposure paradigm used in the present study ensured that biological and tool motion stimuli were overlapping, and thus spatial attention could not be responsible for the observed effects. These findings indicate that attention was acting at the level of object representation, consistent with the biased competition model of attention (Desimone and Duncan, 1995; Beck and Kastner, 2009).

Our findings suggest that while the processing of biological motion might be efficient and robust, it is not attention free. A great deal of research into the perception of biological motion has emphasized the bottom-up nature of processing (Blake and Shiffrar, 2007). An influential and biologically plausible computational model of human biological motion perception implements this perception by using an entirely bottom-up approach (Giese and Poggio, 2003). The results of the present study do not necessarily argue against such an approach if human movements are observed in the absence of any other competing stimuli. However, given our busy, cluttered world, our findings do suggest that such models of biological motion perception should take into account the role of top-down mechanisms such as selective attention.

There has been debate about whether any stimulus category can truly be processed at the level of object representations without the need of attention (Nakayama and Joseph, 1998). Certain classes of stimuli, including those that have considerable ecological significance such as real-world scenes, human bodies, and faces, appear to have privileged access to processing (Rousselet et al., 2002; Downing et al., 2004; Reddy and VanRullen, 2007), although the rapid detection of such stimuli could rely on featural cues rather than object-based representation (Evans and Treisman, 2005). A recent study using fMRI and MEG reported that while hemodynamic responses in category-specific ventral temporal cortex to overlapping faces

and houses showed strong attentional modulation, the early MEG response to faces at ~170 ms (M170) was not modulated by attention (Furey et al., 2006). Faces and biological motion share a number of important characteristics: both depend on elemental detection mechanisms that are present at birth, appear to have been selected at an early stage of mammalian evolution, play important roles in social interaction, and also rely on configural cues for perception (Thompson and Hardee, 2008). However, we found that the cortical electrical response to biological motion and tool motion was modulated by attention and that attentionally dependent responses to the preferred category emerged at ~450 ms following stimulus onset. These findings suggest that despite its similarities with faces, biological motion appears to differ in the need for selective attention when presented with spatially overlapping object motion.

Our finding of a late-emerging category effect that is dependent on attention is consistent with previous ERP studies of the effects of object-based attention using simpler stimuli (Pei et al., 2002). The precise timing of such effects is expected to vary depending on the complexity of the object representation, and, because the point-light motion-defined objects used in the present study require the integration of form and motion, this might explain the latency of this effect. In contrast, spatial and feature-based attentional effects appear to have a much shorter latency (Hillyard and Munte, 1984; Schoenfeld et al., 2007).

Instead of a preattentive response to biological motion, we found that early responses reflected attentional processing, with a greater response at ~200 ms in the two conditions in which participants had to segregate the attended motion category from the unattended category relative to when participants had to segregate the attended motion category from a scrambled version. The sensor-space analysis indicated that this negativity at ~200 ms was the N1 component. The N1 component has been shown to be modulated by attentional demand, with greater amplitude for attended stimuli than unattended stimuli (Parasuraman, 1980; Mangun and Hillyard, 1991; Luck et al., 1993). In addition, the N1 response has been suggested to index discriminative processes (Vogel and Luck, 2000; Hopf et al., 2002). With greater need to discriminate between the two conditions in the overlapping stimulus conditions, it might be expected that the N1 component would show a greater response relative to the single stimulus condition, which is what our sensor-space ERP and source-space CCSD results have shown.

Motion category preferences in lateral temporal cortex
Our results are consistent with previous fMRI evidence that the right STS/MTG shows a preference for biological motion, while the left ITS/MTG prefers tool motion (Beauchamp et al., 2002, 2003). We also showed that responses in these regions to their preferred category emerge around 450 ms and are dependent on the task relevance rather than the simple presence of the stimuli. Such category-specific effects are consistent with previous reports of the category preferences of these two regions (Beauchamp et al., 2002) and with the suggestion that the right STS is a high-level encoder of human actions (Thompson et al., 2007). As is the case with studies that have shown attentional selection of faces in fusiform gyrus when presented simultaneously with houses (O'Craven et al., 1999; Furey et al., 2006), our findings provide further details of the operation of attention at the level of object representations. While different regions in the lateral temporal cortex appear to prefer different categories of object motion, the underlying nature of that specificity is unclear. One proposal is that different regions in lateral temporal cortex

prefer either the articulated motion of human actions (right STS) or the rigid motion of tools (left ITS/MTG) (Beauchamp et al., 2002). While the STS certainly shows a preference for the motion consistent with that generated by an underlying articulated body (Thompson et al., 2005; Pyles et al., 2007), this region is also modulated by the social significance of moving stimuli (Wyk et al., 2009).

Combining fMRI and high-density EEG
In this experiment we used high-density EEG and source localization techniques that allowed us to simultaneously take advantage of spatial resolution approaching that of fMRI along with temporal resolution far higher than that achievable with fMRI. Apart from the analysis of each dataset using surface-based methods, the fMRI data and the CCSD data were independent of each other. In general, the location of effects from these two techniques was remarkably similar. This was particularly the case in right lateral temporal cortex, where we observed considerable overlap between fMRI and CCSD clusters that showed a preference for biological motion compared with tool motion. However, the peak effects for the two different techniques did differ by the order of 10–20 mm. Given that the differences in the physiological and hemodynamic factors that contribute to EEG and fMRI, it is not surprising to observe some differences in the location of activity between these two methods (Nunez and Silberstein, 2000). In addition, the accuracy of even the best source localization techniques is in the order of 1–2 cm when data from real subjects are used (Michel et al., 2004). Our results do however indicate the utility in combining fMRI and EEG methods to examine the spatiotemporal distribution of neural processing.

References

Anderson GJ, Kramer AF (1993) Limits of focused attention in three-dimensional space. Percept Psychophys 53:658–667.

Appelbaum LG, Wade AR, Vildavski VY, Pettet MW, Norcia AM (2006) Cue-invariant networks for figure and background processing in human visual cortex. J Neurosci 26:11695–11708.

Baccus W, Mozgova O, Thompson JC (2009) Early integration of form and motion in the neural response to biological motion. Neuroreport. Advance online publication. Retrieved June 11, 2008. doi:10.1097/WNR.0b013e328330a867.

Beauchamp MS, Lee KE, Haxby JV, Martin A (2002) Parallel visual motion processing streams for manipulable objects and human movements. Neuron 34:149–159.

Beauchamp MS, Lee KE, Haxby JV, Martin A (2003) fMRI responses to video and point-light displays moving humans and manipulable objects. J Cogn Neurosci 15:991–1001.

Beck DM, Kastner S (2009) Top-down and bottom-up mechanisms in biasing competition in the human brain. Vision Res 49:1154–1165.

Blake R, Shiffrar M (2007) Perception of human motion. Annu Rev Psychol 58:47–73.

Cavanagh P, Labianca AT, Thornton IM (2001) Attention-based visual routines: sprites. Cognition 80:47–60.

Crowley TA, Haupt CD, Kynor DB (2000) A weighting matrix to remove depth bias in the linear biomagnetic inverse problem with application to cardiology. In: Biomag 96: Proceedings of the Tenth International Conference on Biomagnetism, Vol I (Aine CJ, Okada Y, Stroink G, Swithenby SJ, Wood CC, eds), pp 97–200. Berlin: Springer.

Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. Annu Rev Neurosci 18:193–222.

Downing CJ, Pinker S (1985) The spatial structure of visual attention. In: Attention and performance XI (Posner MI, Marin OSM, eds), pp 171–187. Hillsdale, NJ: Erlbaum.

Downing PE, Bray D, Rogers J, Childs C (2004) Bodies capture attention when nothing is expected. Cognition 93:27–38.

Evans K, Treisman AM (2005) Perception of objects in natural scenes: is it really attention free? J Exp Psychol Hum Percept Perform 31:1476–1492.

Furey ML, Tanskanen T, Beauchamp MS, Avikainen S, Uutela K, Hari R,

Haxby JV (2006) Dissociation of face-selective cortical responses by attention. Proc Natl Acad Sci U S A 103:1065–1070.

Giese M, Poggio T (2003) Neural mechanisms for the recognition of biological movements. Nat Rev Neurosci 4:179–192.

Green DM, Swets JA (1966) Signal detection theory and psychophysics. New York: Wiley.

Hamalainen MS, Ilmoniemi RJ (1994) Interpreting magnetic fields of the brain: minimum norm estimates. Med Biol Eng Comput 32:35–42.

Hillyard SA, Munte TF (1984) Selective attention to color and location: an analysis with event-related brain potentials. Percept Psychophys 36:185–198.

Hopf JM, Vogel E, Woodman G, Heinze HJ, Luck SJ (2002) Localizing visual discrimination processes in time and space. J Neurophysiol 88:2088–2095.

Johansson G (1973) Visual perception of biological motion and a model for its analysis. Percept Psychophys 14:201–211.

Loftus GR, Masson MEJ. (1994) Using confidence intervals in within-subject designs. Psychonomic Bull Rev 1: 476–490.

Luck SJ, Fan S, Hillyard SA (1993) Attention-related modulation of sensory-evoked brain activity in a visual search task. J Cogn Neurosci 5:188–195.

Mangun GR, Hillyard SA (1991) Modulations of sensory-evoked brain potentials indicate changes in perceptual processing during visual-spatial priming. J Exp Psychol Hum Percept Perform 17:1057–1074.

Mather G, Radford K, West S (1992) Low-level visual processing of biological motion. Proc Biol Sci 249:149–155.

Michel CM, Murray MM, Lantz G, Gonzalez S, Spinelli L, de Peralta RG (2004) EEG source imaging. Clin Neurophysiol 115:2195–2222.

Nakayama K, Joseph JS (1998) Attention, pattern recognition and pop-out in visual search. In: The attentive brain (Parasuraman R, ed), pp. 279–298. Cambridge, MA: MIT.

Neri P, Morrone M, Burr DC (1998) Seeing biological motion. Nature 395:894–896.

Nunez P, Silberstein R (2000) On the relationship of synaptic activity to macroscopic measurements: does co-registration of EEG with fMRI make sense? Brain Topogr 13:79–96.

O'Craven KM, Downing PE, Kanwisher N (1999) fMRI evidence for objects as the units of attentional selection. Nature 401:584–587.

Parasuraman R (1980) Effects of information processing demands on slow negative shift latencies and N100 amplitude. Biol Psychol 11:217–233.

Parasuraman R, de Visser E, Clarke E, McGarry WR, Hussey E, Shaw T, Thompson JC (2009) Detecting threat-related intentional actions of others: effects of image quality, response mode, and target cuing on vigilance. J Exp Psycho Appl 15:275–290.

Pavlova M, Birbaumer N, Sokolov A (2006) Attentional modulation of cortical neuromagnetic gamma response to biological movement. Cereb Cortex 16:321–327.

Peelen MV, Fei-Fei L, Kastner S (2009) Neural mechanisms of rapid natural scene categorization in human visual cortex. Nature 460:94–98.

Pei F, Pettet MW, Norcia AM (2002) Neural correlates of object-based attention. J Vis 2:588–596.

Pyles JA, Garcia JO, Hoffman DD, Grossman ED (2007) Visual perception and neural correlates of novel 'biological motion'. Vis Res 47:2786–2797.

Reddy L, VanRullen R (2007) Spacing affects some but not all visual searches: implications for theories of attention and crowding. J Vis 7:1–17.

Regan D (1989) Human brain electrophysiology: evoked potentials and evoked magnetic fields in science and medicine. New York: Elsevier.

Rousselet G, Fabre-Thorpe M, Thorpe S (2002) Parallel processing in high-level categorization of natural images. Nat Neurosci 5:629–630.

Schoenfeld M, Hopf JM, Martinez A, Mai H, Sattler C, Gasde A, Heinze, HJ, Hillyard SA (2007) Spatio-temporal analysis of feature-based attention. Cereb Cortex 17:2468–2477.

Simion F, Regolin L, Bulf H (2008) A predisposition for biological motion in the newborn baby. Proc Natl Acad Sci U S A 105:809–813.

Thompson JC, Hardee JE (2008) The first time ever I saw your face. Trends Cogn Sci 12:283–284.

Thompson J, Clarke M, Stewart T, Puce A (2005) Configural processing of biological motion in human superior temporal sulcus. J Neurosci 25:9059–9066.

Thompson J, Hardee JE, Panayiotou A, Crewther D, Puce A (2007) Common and distinct brain activation to viewing dynamic sequences of face and hand movements. Neuroimage 37:966–973.

Thornton IM, Rensink RA, Shiffrar M (2002) Active versus passive processing of biological moion. Perception 31:837–853.

Vogel EK, Luck SJ (2000) The visual N1 component as an index of a discrimination process. Psychophysiology 37:190–203.

Wyk BC, Hudac CM, Carter EJ, Sobel DM, Pelphrey KA (2009) Action understanding in the superior temporal sulcus region. Psychol Sci 20:771–777.