



# fMR-adaptation reveals invariant coding of biological motion on the human STS

Emily D. Grossman<sup>1\*</sup>, Nicole L. Jardine<sup>2</sup> and John A. Pyles<sup>3</sup>

<sup>1</sup> Department of Cognitive Sciences, Center for Cognitive Neuroscience, University of California-Irvine, Irvine, CA, USA

<sup>2</sup> Department of Psychology, Vanderbilt Vision Research Center, Vanderbilt University, Nashville, TN, USA

<sup>3</sup> Center for the Neural Basis of Cognition, Carnegie Mellon University, Pittsburgh, PA, USA

## Edited by:

Patrik Vuilleumier, University Medical Center and University Hospital Geneva, Switzerland

## Reviewed by:

Michael S. Beauchamp, University of Texas, USA

Marius Peelen, University of Trento, Italy

## \*Correspondence:

Emily D. Grossman, Department of Cognitive Sciences, Center for Cognitive Neuroscience, University of California Irvine, Irvine, CA 92697-5100, USA.  
e-mail: grossman@uci.edu

Neuroimaging studies of biological motion perception have found a network of coordinated brain areas, the hub of which appears to be the human posterior superior temporal sulcus (STSp). Understanding the functional role of the STSp requires characterizing the response tuning of neuronal populations underlying the BOLD response. Thus far our understanding of these response properties comes from single-unit studies of the monkey anterior STS, which has individual neurons tuned to body actions, with a small population invariant to changes in viewpoint, position and size of the action being viewed. To measure for homologous functional properties on the human STS, we used fMR-adaptation to investigate action, position and size invariance. Observers viewed pairs of point-light animations depicting human actions that were either identical, differed in the action depicted, locally scrambled, or differed in the viewing perspective, the position or the size. While extrastriate hMT+ had neural signals indicative of viewpoint specificity, the human STS adapted for all of these changes, as compared to viewing two different actions. Similar findings were observed in more posterior brain areas also implicated in action recognition. Our findings are evidence for viewpoint invariance in the human STS and related brain areas, with the implication that actions are abstracted into object-centered representations during visual analysis.

**Keywords:** biological motion, superior temporal sulcus, visual recognition, fMRI, vision

## INTRODUCTION

We live in a dynamic social environment in which the swift and accurate assessment of people's actions facilitates our social interactions and enriches our daily experiences. To make sense of these complex social situations, one must engage a variety of cognitive tasks that include recognizing the actions and intentions of others. Although humans make these assessments quickly and seemingly effortlessly, there is an abundance of evidence that this is a cognitively demanding task requiring the coordination of multiple brain systems.

Neuroimaging and computational studies have implicated both dorsal and ventral visual brain areas in action recognition, and the premotor cortex (PMC) as a part of a larger putative "mirror" system (Saygin et al., 2004; Wheaton et al., 2004). In visual analysis, dorsal brain areas are proposed to support the encoding of action kinematics, and ventral brain areas are proposed to analyze the underlying body postures (Grossman and Blake, 2002; Beauchamp et al., 2003; Giese and Poggio, 2003; Michels et al., 2005; Thompson et al., 2005). Most of these brain areas are not believed to be specialized for action recognition. For example, the motion-sensitive human MT complex (hMT+) has cortical activity strongly linked to motion perception (e.g., Huk et al., 2001), and is likely involved in analysis of body kinematics (Peuskens et al., 2005), but does not appear to be specialized in any way for biological motion perception (Grossman et al., 2000; Peelen et al., 2006). Ventral stream brain areas implicated in biological motion include a region on the fusiform sometimes referred to as the fusiform body area (FBA), and a

region on the inferior temporal sulcus (ITS) that may correspond to the extrastriate body area (EBA), which is often characterized as "ventral" despite being situated anatomically dorsal to the hMT+ (Vaina et al., 2001b; Grossman and Blake, 2002; Taylor et al., 2009). Each of these brain areas has neural signals that dissociate action kinematics from non-biological motion, however, the EBA and FBA both also support recognition of stationary body postures (Taylor et al., 2007; Peelen et al., 2009), while the ITS also has neural signals that support recognition of dynamic and articulated novel objects "Creatures" (Pyles et al., 2007).

Anatomically situated at the center of this large cortical system is the superior temporal sulcus (STS), which has been proposed as the integration site of the two visual processing streams and the hub of this larger network. The STS is the brain area that has most critically been implicated in action recognition (STS, e.g., Blake and Shiffrar, 2007; Adolphs, 2009), and appears to have subregions that support action observation (Bonda et al., 1996; Grossman et al., 2000), action imagery (Kourtzi and Kanwisher, 2000; Grossman and Blake, 2001), and recognition of action verbs (Bedny et al., 2008). Patients with lesions on the STS have difficulty recognizing actions (Battelli et al., 2003; Pavlova et al., 2003), an effect that is reproduced by creating reversible "lesions" over the STS through repetitive TMS (Grossman et al., 2005).

To date, our best understanding of the neuronal specialization on the STS comes from single-unit investigations of the macaque superior temporal polysensory area (STPa, sometimes referred to as STSa), the likely homologue to human STS (Puce and Perrett, 2003).

Investigations of the macaque STSa have identified a population of neurons that fire during action recognition (Oram and Perrett, 1994) and appear to be tuned to the movements of bodies or specific body parts such as the face, arms and legs (Jellema et al., 2004; Barraclough et al., 2006; Vangeneugden et al., 2009). The majority of these STSa neurons are tuned to specific combinations of body postures and movements that render them effectively viewpoint specific (Oram and Perrett, 1996). A smaller population of STSa neurons, however, appear to be viewpoint invariant, as demonstrated by more generalized tuning across changes in viewing perspectives (i.e., a person walking backwards as seen from the front, or back). Both populations of biological motion selective neurons have large receptive fields (Bruce et al., 1981), tend to be position and size invariant (Wachsmuth et al., 1994), and are anatomically intermixed (Perrett et al., 1985b).

The purpose of this study is to use fMRI to investigate viewpoint dependence and neural tuning of the human STS, and functionally related cortical areas. Based on the range of stimuli used to identify functional specialization on the human STS, we hypothesized that this region, much like the monkey STSa, consists of heterogeneous populations of neurons. Therefore, if some viewpoint dependencies do exist, they may exist only within some subpopulation of the STS neurons. Standard neuroimaging techniques would be unlikely to detect such subtle differences among subpopulations in the millions of neurons driving the BOLD in a brain area, given the average voxel size of 1.5–4 mm and larger effective spatial resolution due to spatial correlations of the BOLD response.

To overcome the spatial limitations inherent to the BOLD response, we have used fMR-adaptation (also referred to as repetition suppression) to “tag” subpopulation of neurons underlying the fMRI response (e.g., Henson et al., 2003; Grill-Spector, 2006; Krekelberg et al., 2006). fMR-adaptation was derived from the observation that repeated presentations of identical stimuli tend to result in reduced hemodynamic responses as compared to pairs of different stimuli (Buckner et al., 1998; Grill-Spector and Malach, 2001). This is not entirely unlike the double-pulse or stimulus specific adaptation effects observed in omnibus electrode recording and unit physiology in which prior exposure to an image reduces the neural response for that image in subsequent presentations (Musselwhite and Jeffreys, 1983; Baylis and Rolls, 1987; Fahy et al., 1993; Li et al., 1993; Sobotka and Ringo, 1994). The physiological mechanism of adaptation is currently unknown, but the effect has been linked to habituation, neuronal fatigue, and repetition priming (Muller et al., 1999; Ganel et al., 2006; Grill-Spector et al., 2006).

In the rapid fMR-adaptation paradigm, the suppression of the neural response to the second exposure is measured as a slightly reduced peak (on the order of 0.05–0.1% signal reduction) in the hemodynamic impulse response function. This is because the sluggish BOLD response reflects summation of the two events (i.e., two impulse response functions convolved with a 16-s hemodynamic response function). Of most importance to these studies, the fMRI-adaptation paradigm has been exploited to reveal dimensions of visual stimuli coded by a particular neural population (sometimes referred to as “invariance”). For example, this technique has successfully revealed invariant properties of shape, object and face

perception in ventral temporal cortex (Kourtzi and Kanwisher, 2001; Andrews and Ewbank, 2004; Self and Zeki, 2005; Andresen et al., 2009; Pyles and Grossman, 2009).

In four experiments we tested five specific properties of STS neurons: action specificity, invariance across mirror reversals, cross-adaptation between intact actions and motion-matched controls, and invariance across changes in position and size. We compared the peak hemodynamic response in targeted brain areas (identified by independent localizers) and explored possible adaptation effects in a whole-brain general linear model analysis. We found that the STS processes biological motion in a viewpoint-invariant manner. The STS BOLD response shows specificity for individual actions, has a subpopulation of neurons that does not adapt from the motion features alone, and is invariant across mirror-reversals, changes in position and size. Together, these findings are evidence for high-level action tuned populations of neurons on the STS, much like what has been shown in monkey STSa.

## MATERIALS AND METHODS

### PARTICIPANTS

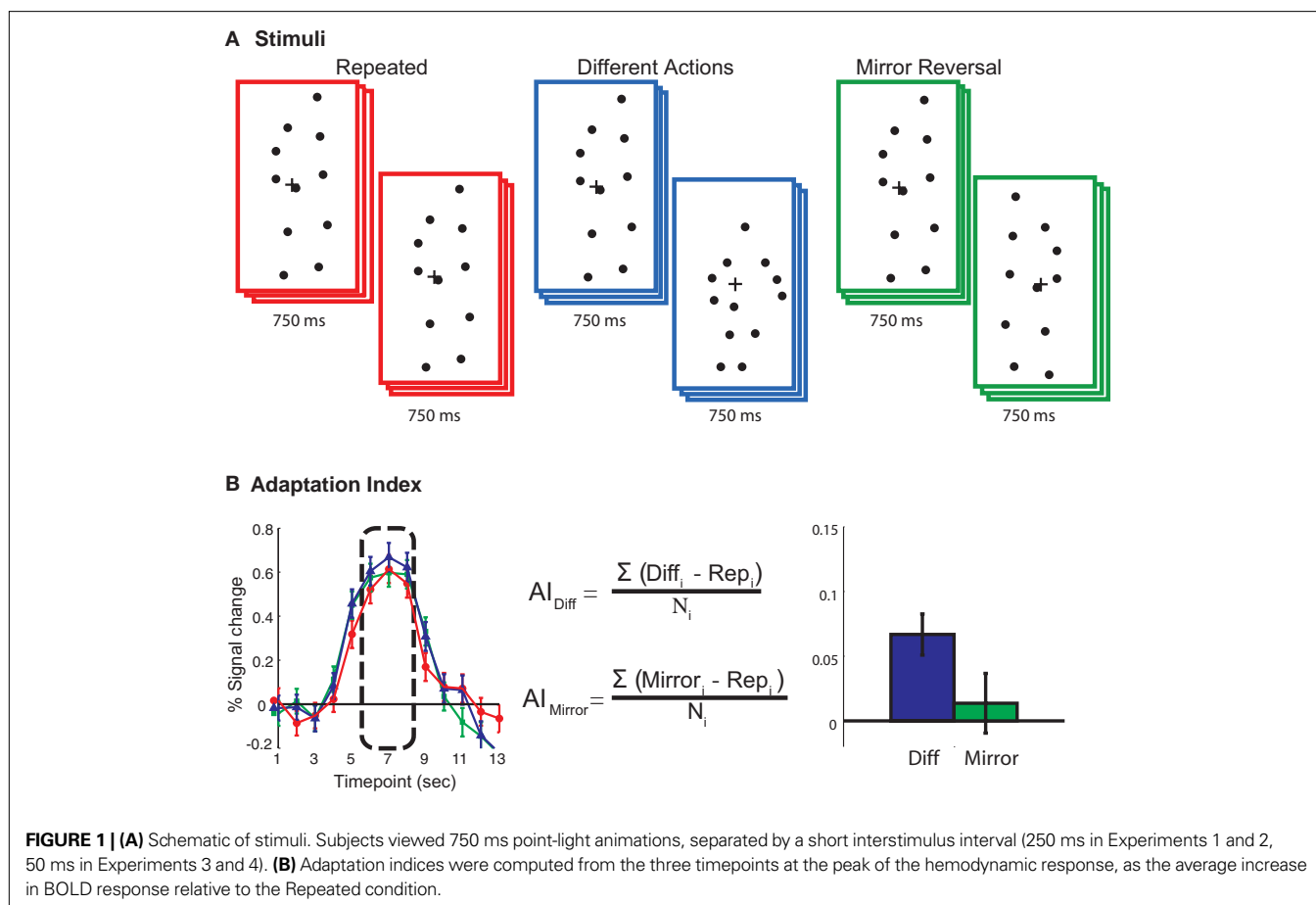
A total of 19 participants (6 male, 13 female) from the UC Irvine campus and community gave informed and written consent as approved by the University of California Irvine Institutional Review Board. Fourteen observers with normal or corrected-to-normal vision participated in Experiments 1 and 2. One subject was excluded from the analysis due to excessive head movements, and another was excluded due to unusual features in the anatomical images. Of the remaining 12, nine were naïve to biological point-light stimuli. Eight subjects participated in Experiments 3 and 4, two of whom had participated in the previous experiments, with the remaining all naïve to point-light biological motion. All subjects had normal or corrected-to-normal vision.

### MRI ACQUISITION

Neuroimaging data were collected on a 3T Philips Achieva whole-body MRI scanner equipped with an eight-channel head-coil on campus at the University of California, Irvine. We collected high-resolution whole-brain anatomical images (T1-weighted MPRAGE,  $1 \times 1 \times 1 \text{ mm}^3$ , TE = 3.7 ms, TR = 8.4 ms, flip angle =  $8^\circ$ , SENSE factor = 2.4) from each individual to be used for co-registration of the functional scans. Subjects participated in two types of functional scans (both single-shot T2\*-weighted gradient EPI, TE = 30 ms, flip angle =  $90^\circ$ , right-left phase encoding, SENSE factor = 1.5, slices acquired ascending and interleaved). Localizer scans were designed to identify regions of interest (ROIs) on the STS and in more posterior areas, and were acquired using standard imaging parameters (28 oblique axial slices,  $2.05 \text{ mm} \times 2.05 \text{ mm} \times 4 \text{ mm}$  voxels, 0 mm gap, TR = 2 s). The rapid event-related scans were designed to more accurately estimate the amplitude of the hemodynamic response, and thus were acquired rapidly, with fewer slices targeting the intersection of the temporal, occipital, and parietal lobes (17 oblique axial slices,  $4 \times 4 \times 4 \text{ mm}^3$  voxels, 1 mm gap, TR = 1000 ms).

### PROCEDURE

During scanning participants viewed point-light biological motion sequences extracted from videotaped segments of an individual performing actions with lights attached to their joints (**Figure 1A**).



The point-light animations in these experiments depicted 25 unique actions, including walking, running, jumping, and throwing, and have been used in previous psychophysical and neuroimaging studies (Grossman et al., 2004; Garcia and Grossman, 2008). These sequences were digitized and encoded as (x, y, t) joint positions in Matlab (Mathworks, Inc.), then displayed as 12 small black dots ( $0.17^\circ$  of visual angle, with overall figure subtending approximately  $8 \times 3.5^\circ$  of visual angle) on a gray background. Stimuli were displayed on a Christie DLV 1400-DX DLP projector using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) controlled by a Macintosh G4 computer. Subjects viewed the animations through a periscope mirror mounted on the birdcage head-coil and directed at a screen positioned at the head end of the scanner.

To eliminate issues of circularity in the data analysis (e.g., Kriegeskorte et al., 2009), all subjects participated in two types of scans: localizer scans and fMR-adaptation experimental scans. Localizer scans were designed to identify the areas of the brain that respond selectively to biological motion (e.g., Pyles et al., 2007). This was achieved by comparing neural responses when subjects viewed point-light biological motion to when subjects viewed motion-matched “scrambled” animations. Scrambled animations were created by randomizing the starting position of the point-light dots within a region approximating the target figure, then leaving their motion vectors intact. Seven blocks each of biological motion and scrambled motion were interleaved, separated by 6 s passive

fixation intervals. Subjects viewed eight exemplars of motion (1 s each, with 1 s interstimulus interval) within each 16-s block. Subjects also passively fixated the initial 8 s and final 16 s of each scan. During the stimulus blocks, subjects performed a one-back task on each exemplar to indicate if an animation repeated sequentially. Responses were collected on an MR-compatible response box (Current Design, Inc.).

We also collected an hMT+ localizer in 11 of the 12 subjects that participated in Experiments 1 and 2, and in all of the subjects who participated in Experiments 3 and 4. The purpose of this scan was to identify a motion-selective control area that would reflect the low-level motion properties of the stimulus (e.g., position). To achieve this, we positioned two fields of expanding and contracting optic flow motion, centered  $14.6^\circ$  to the left and right of a central fixation cross. In alternating 14 s intervals separated by 6 s fixation, subjects viewed the moving optic flow in the left hemifield or right hemifield, or viewed both fields of dots held stationary. These epochs were repeated five times within the 5-min scan. Subjects passively fixated the initial 8 s and final 16 s of the scan.

fMR-adaptation was measured in separate rapid event-related scans. A single trial in these scans consisted of a pair of animations, each 750 ms, with a brief fixation (50–250 ms, described below) between the items in the pair. The first animation in each pair depicted one of 25 actions such as walking, squatting, throwing, or pushing. The second animation in a pair was determined by

experiment and condition, as explained below. For all experiments, subjects viewed 10 trials of each experimental condition per scan. To optimize deconvolution of the hemodynamic response and separate the BOLD amplitude estimates for each condition, we jittered the inter-trial interval from 3 to 9 s between pairs (Dale, 1999; Serences, 2004). The initial 18 s and final 24 s of each scan consisted of passive fixation to allow for MR saturation, to acquire an estimate of passive fixation, and to allow measurement of the hemodynamic response associated with the final trial.

Experiment 1 tested for action specificity in the STS response, and for invariance across changes in viewing perspective. Subjects viewed the following three types of trials: (1) Repeated trials in which the same action was repeated twice, (2) Different trials in which two different actions were viewed, and (3) Mirror Reversal trials in which the same action was repeated, but was left-right reversed in viewing perspective. Each action animation was 750 ms in duration, with a 250-ms interstimulus interval between the two animations in the pair. Subjects were asked to indicate whether the animation depicted the same or different action by a keypress (2-alternative forced choice discrimination). All actions were viewed foveally, centered on a fixation cross in the center of the image. The first four subjects were instructed that the correct response for the Mirror Reversal condition was “same”. Because this task confounded the task decision with viewpoint invariance (namely, Mirror Reversal trials and Repeated trials were both associated with a “same” response), we instructed the remaining eight subjects to assign the Mirror Reversal trials as “different”. This experimental scan was repeated eight times for each subject, yielding 80 estimates for each predictor of the hemodynamic response, per condition and per subject.

Experiment 2 tested for possible adaptation of low-level features, unspecific to action interpretation. In these scans, subjects viewed four types of trials: (1) Biological motion alone trials in which a single action was shown once, (2) Scrambled trials in which a single motion-matched non-biological “scrambled” control animation was viewed (see Section “Materials and Methods” for construction of the scrambled animations), (3) Biological + Scrambled trials in which first a point-light biological animation was shown, followed by the same action scrambled such that all the local motion features were matched with the prior biological sequence (e.g., walker + scrambled walker, kicker + scrambled kicker), and (4) Scrambled + Biological trials in which a scrambled sequence was followed immediately by the same action intact (e.g., scrambled walker + walker). Each action animation was 750 ms in duration, with a 250-ms interstimulus interval between the two animations in the pair. Observers passively viewed the foveally positioned sequences, which were centered on a fixation cross in the central part of the screen.

Experiment 3 tested the position specificity of the neural response on the STS. In these scans, subjects viewed the following three types of trials: (1) Fovea + Fovea in which the same action was presented twice in the central vision (same as the Repeated condition from Experiment 1), (2) Fovea + Left in which the same point-light action was positioned first in the fovea, then centered 5.2° into the left hemifield, and (3) Fovea + Right in which the same point-light action was viewed first in the fovea, then centered 5.2° in the right hemifield. For all conditions, the animations

could depict one of the 25 unique actions, but the same action was always repeated within the pair and subtended  $8 \times 3.5^\circ$  in visual angle. Each action animation was 750 ms in duration, with a 50-ms interstimulus interval between the two animations in the pair. This scan was repeated four times for a total of forty estimates for each predictor of the hemodynamic response, per condition, per subject.

Subjects were instructed to maintain fixation when animations were positioned in the visual peripheral field. At the time of data collection, we were not able to measure eye movements to verify that subjects successfully maintained fixation; however, we will report on the visual response from retinotopic hMT+, which demonstrates a high degree of retinotopic specificity (e.g., Wandell et al., 2007). To maintain attention throughout the scans, observers were required to make a leftward/rightward facing judgment on each trial. Because the action was identical, the observers could make this assessment on the basis of the first animation, which was always positioned in the foveal field.

Experiment 4 measured for size invariance in the BOLD response. In these scans, subjects viewed three types of trials: (1) Medium + Medium in which the same action was presented in the fovea twice, both of the same size (5.2° of visual angle), (2) Medium + Small in which the same action was viewed first as 5.2°, then as 3°, (3) and Medium + Large in which the same animation was viewed first as 5.2°, then as 9°. For all trial types, the same animation was repeated within the pair and could depict any one of the 25 unique actions. Each action animation was 750 ms in duration, with a 50-ms interstimulus interval between the two animations in the pair. Subjects viewed ten trials of each type per each 3:36 min scan, and the scan was repeated for a total of forty estimates for each predictor of the hemodynamic response, per condition, per subject. To maintain attention throughout the scans, observers were required to make a leftward/rightward facing judgment on each trial. Because the action was identical, the observers could make this assessment on the basis of the first animation.

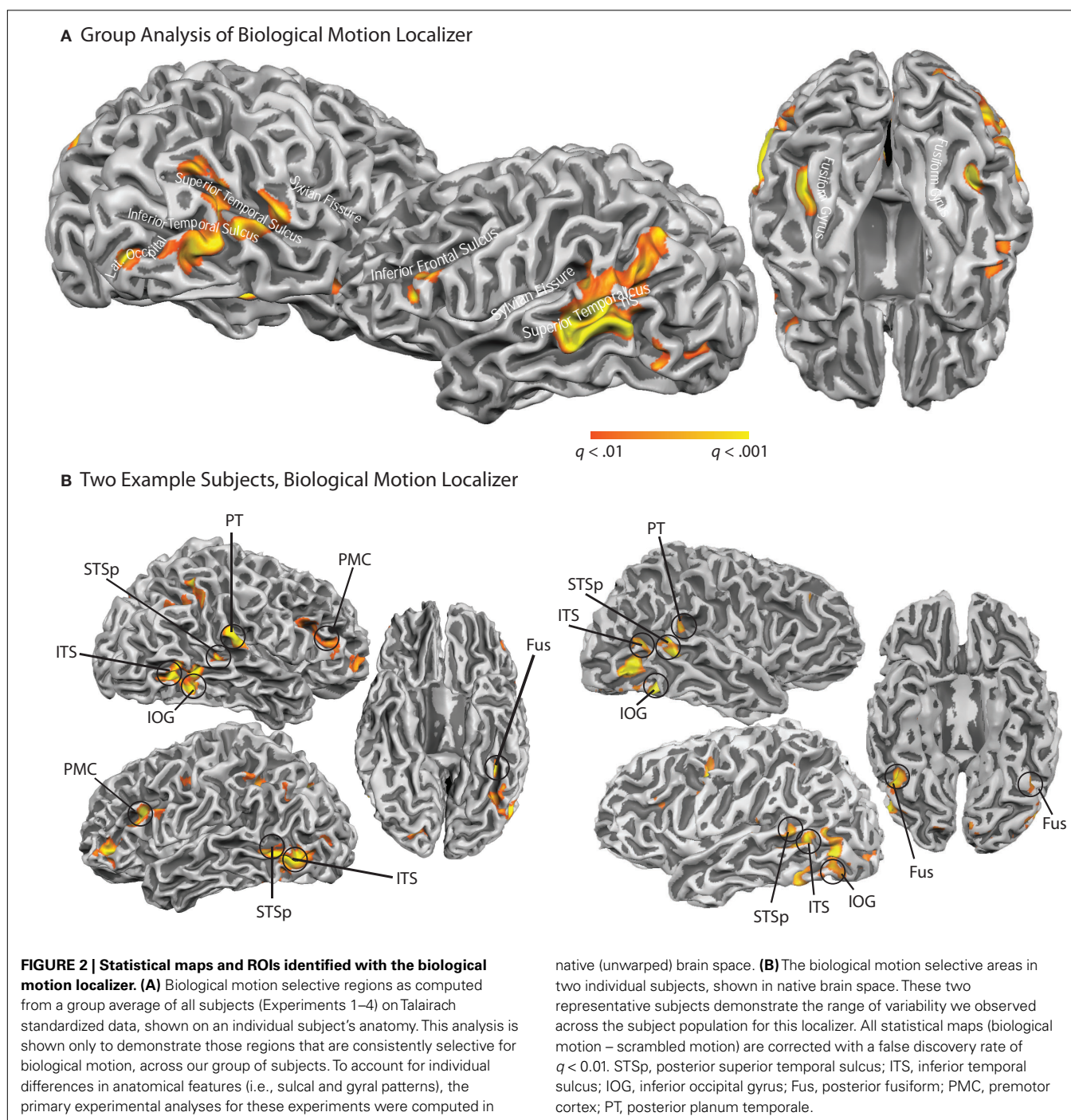
## ANALYSIS

Preprocessing and fMRI data analysis were conducted with BrainVoyager QX (Brain Innovations, Inc.). All fMRI images were corrected for slice acquisition order, corrected for any subject movement both within and across scans, corrected for any linear signal drift, temporally high-pass filtered at three cycles per scan (approximately 0.01 Hz), and co-registered to the individual subjects' anatomical images. All timecourses were normalized to percent signal change and all subsequent statistical maps were thresholded using a false discovery rate (FDR, with a corresponding  $q$ -value). The FDR corrects for the family-wise error rate by controlling the proportion of expected false positives, and has the advantage of being adaptive to the signal levels in data while still correcting for multiple comparisons (Genovese et al., 2002).

Localizer scans were analyzed in native brain space (not warped to a standard head model) using a general linear model analysis with boxcar predictors convolved with a model of the hemodynamic response (Boynton et al., 1999), and statistical maps were thresholded at  $q < 0.01$ . Biological motion selective brain areas were localized as the regions with significantly higher beta estimates for the biological epochs as compared to the scrambled estimates.

This statistical comparison typically yields a number of brain areas within occipital, temporal and parietal cortex (**Figure 2**, see **Table 1** for Talairach coordinates across our population of subjects). On the basis of previous literature and the known anatomical location of these regions, we identified the following ROIs in our subjects: the STSp was identified along the dorsal extent of the STS, near the intersection with the ITS; the posterior extent of the ITS, a region on the lateral surface of the posterior temporal lobe, on the inferior occipital gyrus (IOG); a region on the ventral surface of the temporal lobe, adjacent to the inferior occipital sulcus and

the fusiform gyrus (Fus), a region on the most posterior extent of the Sylvian Fissure likely corresponding to the multisensory subregions of the planum temporale (PT, Meyer et al., 2007; Hickok and Saberi, in press), and a region on the inferior frontal gyrus, likely corresponding to PMC. In some subjects the ITS and IOG ROIs formed a continuous region with different peaks. In these instances, the ROIs were restricted so as to consist of non-overlapping voxels centered at the peaks. We should also note that some of these ROIs may correspond to brain areas discussed in the literature and identified with other functional localizers. For example, the Fus



**Table 1 | Mean Talairach coordinates for the ROIs isolated using the independent localizers on each subject.**

ROI	Left hemisphere				Right hemisphere			
	N	X	Y	Z	N	X	Y	Z
<b>EXPERIMENTS 1 AND 2. DIRECTION INVARIANCE AND CROSS-ADAPTATION</b>								
STSp	10/12	-51.8	-57.9	9.5	8/12	48.1	-51.1	11.7
ITS	6/12	-41.2	-63.7	6.3	8/12	41.9	-62.1	7.0
IOG	6/12	-48.8	-69.8	4.5	8/12	47.8	-70.0	0.6
Fus	9/12	-37.3	-52.7	-12.4	8/12	37.6	-53.9	-11.3
PMC	9/12	-44.4	23.3	12.1	7/12	45.0	26.4	10.3
PT	8/12	-53.3	-45.5	21.1	7/12	49.4	-44.0	20.0
hMT+	11/11	-44.8	-70.5	2.7	10/11	36.4	-69.1	3.9
<b>EXPERIMENTS 3 AND 4. SIZE AND POSITION INVARIANCE</b>								
STSp	6/8	-48.2	-58.0	14.2	8/8	49.3	-50.3	14.9
ITS	6/8	-42.8	-73.0	8.2	7/8	44.3	-66.6	9.0
IOG	3/8	-51.7	-73.0	11.7	4/8	46.8	-73.3	4.0
Fus	6/8	-39.8	-56.3	-10.8	6/8	39.0	-52.7	-11.2
PMC	7/8	-42.9	30.9	14.1	6/8	42.8	25.5	12.5
PT	4/8	-53.8	-50.3	27.3	6/8	50.5	-44.2	24.5
hMT+	8/8	-44.1	-78.4	5.9	8/8	43.3	-71.4	1.7

*N* indicates the number of subjects in which this ROI was identified, out of the total possible. All ROIs except the hMT+ were isolated using the biological motion localizer. One subject from Experiments 1 and 2 did not participate in the hMT+ localizer. STSp, posterior superior temporal sulcus; ITS, inferior temporal sulcus; IOG, inferior occipital gyrus; Fus, posterior fusiform; PMC, premotor cortex; PT, posterior planum temporale, ventral angular gyrus; hMT+, human middle temporal area complex.

region we have identified likely corresponds to the human FBA, while the ITS region may overlap with the functionally identified EBA (Grossman et al., 2004; Michels et al., 2005; Peelen et al., 2006). However, because we did not collect those traditional localizers on these subjects we have used anatomical (as opposed to functional) labels for these regions. All of the Talairach coordinates for all of these brain areas shown in **Table 1**, including the number of subjects and hemispheres in which they were observed.

The hMT+ localizer was analyzed using a GLM with predictors for optic flow and stationary dots, convolved with the estimated hemodynamic response. The hMT+ was identified as the region on the ascending limb of the ITS that is more activated by the optic flow as compared to the stationary dots ( $q < 0.001$ ).

For the fMR-adaptation scans, two types of analyses were performed. First we conducted a ROI-based deconvolution analysis that estimated the hemodynamic response for each stimulus condition, within each independently identified ROI. This was done in “native” brain space (i.e., without transforming the functional data to standardized space), allowing for some flexibility among the variations in individual anatomical structure (e.g., Ono et al., 1991). The focus on pre-defined ROIs (from the biological motion localizer) was chosen because of the known importance of these regions in supporting biological motion recognition. For this ROI-based analysis, the hemodynamic response for each voxel was estimated using a deconvolution GLM with 20 predictors for each of the 20 s following stimulus onset, for each condition. The input to this GLM model was the BOLD timecourse from each ROI, normalized to percent signal change from the mean baseline intensity. This analysis makes no *a priori* assumptions as to the shape or latency of the underlying response profile, and was largely successful in yielding classic hemodynamic response functions from the ROI

timecourses. The latency of the peak amplitude in the deconvolved BOLD responses from Experiments 1 and 2 was approximately 6–8 s following the onset of the first animation in each trial pair, and 5–8 s in Experiments 3 and 4. The difference in the peak latencies across these experiments likely reflects the shorter stimulus duration in the latter experiments (Boynton et al., 1996; Dale and Buckner, 1997).

To test for significant differences among the conditions, planned contrasts computed the statistical significance of the peak of the BOLD responses for each condition, with the peak amplitudes from each condition (e.g., the 5–7 s of the response post-stimulus onset) weighted as  $-1$  and contrasted against a second condition weighted  $+1$ .

In a second analysis, we conducted a whole-brain GLM to probe across the entire brain for regions with evidence of action specificity, or invariance across viewing perspective, position or size. We should note that this analysis is not entirely independent of the ROI-based analysis as it is being conducted on the same data (in part, see below) and using the same statistical hypotheses (Kriegeskorte et al., 2009). Thus the whole-brain analysis should be interpreted as complementary to the ROI-based analysis in that it reveals larger patterns of brain activity engaged in repetition suppression across cortex and across our group of subjects.

This whole-brain GLM analysis was computed across subjects, with functional data normalized to standardized Talairach space (Talairach and Tournoux, 1988). This was achieved by aligning the high-resolution anatomical brain images along the native ACPC axis, then scaling the images to the boundaries of the gray matter. The resulting transformation matrices were then applied to the functional images. Within this standardized space, we then estimated the hemodynamic response function for each voxel and

condition using the same deconvolution analysis procedure as in the ROI-based analysis. We computed statistical contrasts testing for stimulus specificity and invariance (detailed in Section “Results”) and applied a false discovery rate threshold of  $q < 0.01$ .

### QUANTIFYING fMR-ADAPTATION

To determine the difference in BOLD responses for each condition in each ROI, an adaptation index (AI) was computed for each experimental condition (Figure 1B). The AI is a method for calculating the differences in the peak amplitudes between the test conditions (i.e., Different Actions) and the Repeated conditions (including Fovea + Fovea and Medium + Medium), for which one would anticipate the weakest neural response given repetition suppression. However, we found the differences in the deconvolved hemodynamic response for our experimental conditions often extended beyond the peak response, which we hypothesize may be due to the dynamic nature of our stimuli. Therefore, to capture this more robust estimate of differences between our conditions we computed an AI that estimated the mean of this response over a range of timepoints surrounding the peak amplitude response. The AI was calculated as:

$$AI = \frac{\sum (Test_i - Rep_i)}{N}$$

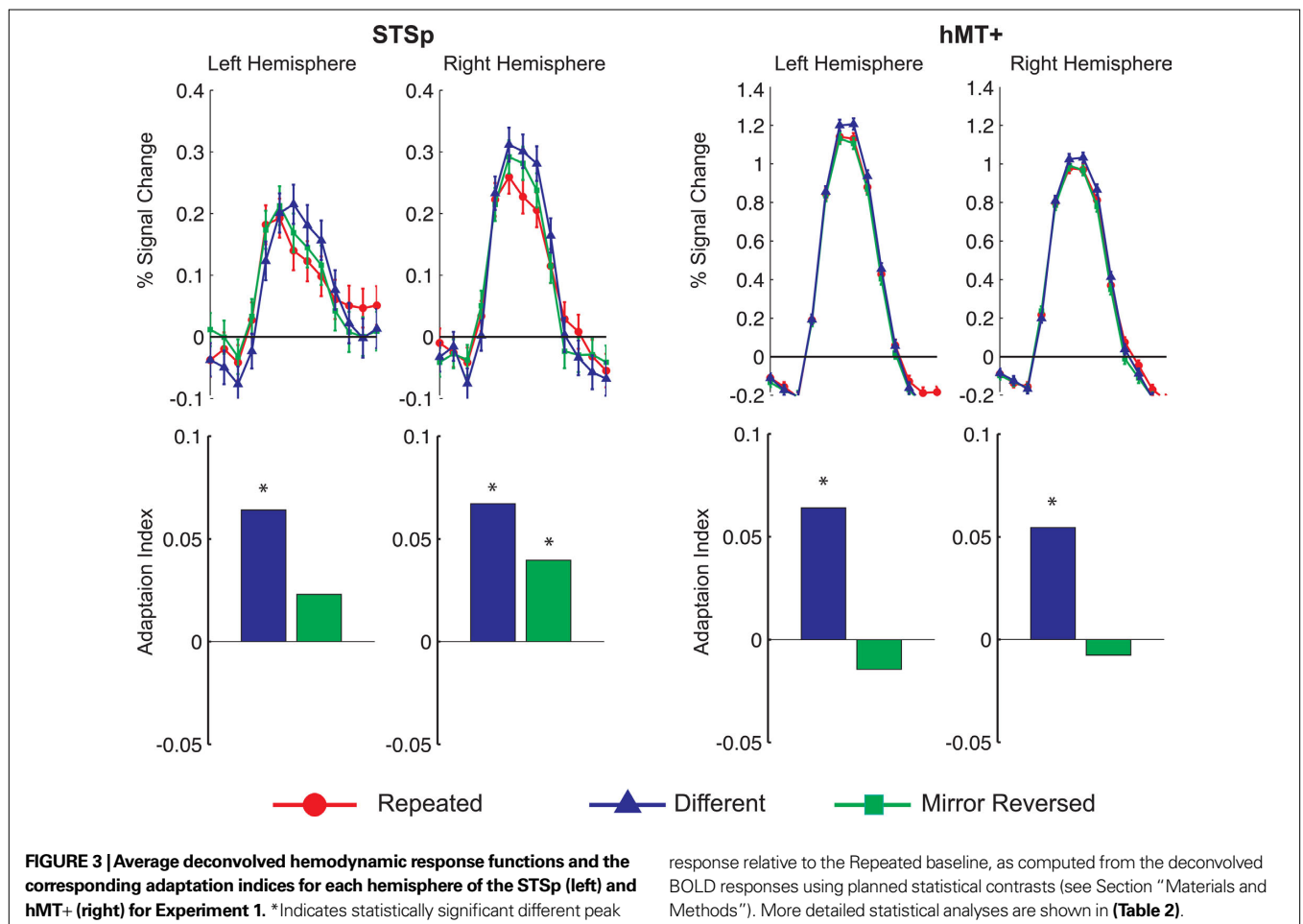
with  $Test_i$  = BOLD percent signal change for the condition of interest (i.e., Different Actions) at timepoint  $i$ ;  $Rep_i$  = percent signal change for the Repeated condition at timepoint  $i$ , and  $N$  = the number of timepoints included in the analysis (always 3).

## RESULTS

### EXPERIMENT 1: DIRECTION INVARIANCE

In this first experiment we measured the specificity with which the STSp encode actions, and specifically whether these brain areas dissociate between different actions, and between the same action seen as mirror reversals. To do this, subjects viewed pairs of the same actions (Repeated), two different actions (Different), or the same action mirror reversed (Mirror Reversed). Results from these measurements are shown in Figure 3 with the corresponding statistical analyses in Table 2.

We found evidence for action specific encoding in both hemispheres of the STSp, as evidenced by statistically significant fMR-adaptation, meaning that the BOLD response for the Repeated condition was weaker than for the Different condition. Indeed, we observed adaptation for the repeated actions throughout our independently localized ROIs, including in the motion sensitive hMT+ and PMC in the right hemisphere. The fMR-adaptation effect is preliminary evidence that the analysis of different actions is supported by unique neural populations.



**Table 2 | The results of planned statistical contrast on the deconvolved hemodynamic response functions from the STSp and hMT+, for Experiments 1, 3 and 4.**

EXPERIMENT 1. DIRECTION INVARIANCE				
ROI	Hemi	Different–Repeated	Different–Mirror	Mirror–Repeated
STSp	Right	$t = 4.37^{***}$	$t = 1.79, p = 0.07$	$t = 2.56^*$
	Left	$t = 3.66^{**}$	$t = 2.34^*$	$t = 1.32, p = 0.19$
hMT+	Right	$t = 3.63^{**}$	$t = 4.13^{***}$	$t = -0.50, p = 0.62$
	Left	$t = 3.88^{***}$	$t = 4.76^{***}$	$t = -0.87, p = 0.38$
EXPERIMENT 3. POSITION INVARIANCE				
ROI	Hemi	Contralateral	Ipsilateral	
STSp	Right	$t = 1.78, p = 0.08$	$t = -0.60, p = 0.55$	
	Left	$t = 0.22, p = 0.83$	$t = 1.55, p = 0.12$	
hMT+	Right	$t = 3.40^{***}$	$t = -4.39^{***}$	
	Left	$t = 2.16^*$	$t = -5.54^{***}$	
EXPERIMENT 4. SIZE INVARIANCE				
ROI	Hemi	Large	Small	
STSp	Right	$t = -1.00, p = 0.32$	$t = -0.18, p = 0.86$	
	Left	$t = -0.31, p = 0.76$	$t = 0.47, p = 0.64$	
hMT+	Right	$t = -1.72, p = 0.09$	$t = -1.86, p = 0.06$	
	Left	$t = -1.33, p = 0.18$	$t = -1.86, p = 0.06$	

Planned statistical contrasts were computed on the timepoints surrounding the peak response, as deconvolved using a general linear analysis (see Section “Materials and Methods”). Abbreviations are the same as in **Table 1**. Statistical significance at  $*p < 0.05$ ,  $**p < 0.01$ ,  $***p < 0.0001$ .

To determine whether these brain areas were encoding the low-level features of the specific exemplar being observed (such as local motion vectors and specific body postures) or a more generalized representation of the action being depicted, we tested for adaptation across mirror reversals of the same point-light animation. Both hemispheres of the STSp revealed intermediate levels of adaptation for the mirror-reversed actions, with the left hemisphere being completely adapted, and the right hemisphere adaptation being marginally significant.

#### EXPERIMENT 2: CROSS-ADAPTATION

In a second experiment, we considered the extent to which the adaptation findings we observed for the Mirror Reversal trials could be attributed to a population of neurons encoding low-level properties of the stimulus, such as velocity. In particular, the mirror manipulation reversed the horizontal motion trajectories of the sequences, but left intact the vertical trajectories and horizontal mid-level features, which may be critical features in these sequences (Casile and Giese, 2005; Thurman and Grossman, 2008). It is likely that many of our ROIs, and in particular hMT+ and the STSp, contain populations of velocity-tuned neurons (e.g., Bruce et al., 1981), and our adaptation effect may have been due to those neurons alone.

Therefore, to determine whether our adaptation findings reflect the “low-level” (directionally tuned, but not action specific) or “high-level” (action tuned) population of neurons, we made two measurements. First, we measured the BOLD response to single presentations of biological or scrambled sequences, which serves to estimate the relative proportion of low- and high-level neurons in each ROI. Second, we measured for cross-adaptation between biological and motion-matched non-biological (scrambled) animations. Neurons encoding only low-level features should be subject to adaptation from both of these types of stimuli, while neurons encoding the actions (the high level neurons) would only be subject to adaptation by the biological sequences. Because the outputs of action-tuned high-level neurons are believed to be constructed from the outputs of the low-level neurons (Jellema et al., 2000; Giese and Poggio, 2003), a critical factor is the order of the stimuli in the cross-adaptation trials. One should expect an asymmetry of adaptation depending on which stimulus serves as the adaptor, with stronger adaptation in trials where the scrambled animation is shown first (thus resulting in a dampened response in both the low-level and high-level populations). The predictions of this asymmetry are shown in the top panel of **Figure 4**, and generalize to any brain area in which low-level neurons serve as afferents to higher-level computation. Critically, this asymmetry should not be observed in a brain area that encodes only the low-level features of the stimulus.

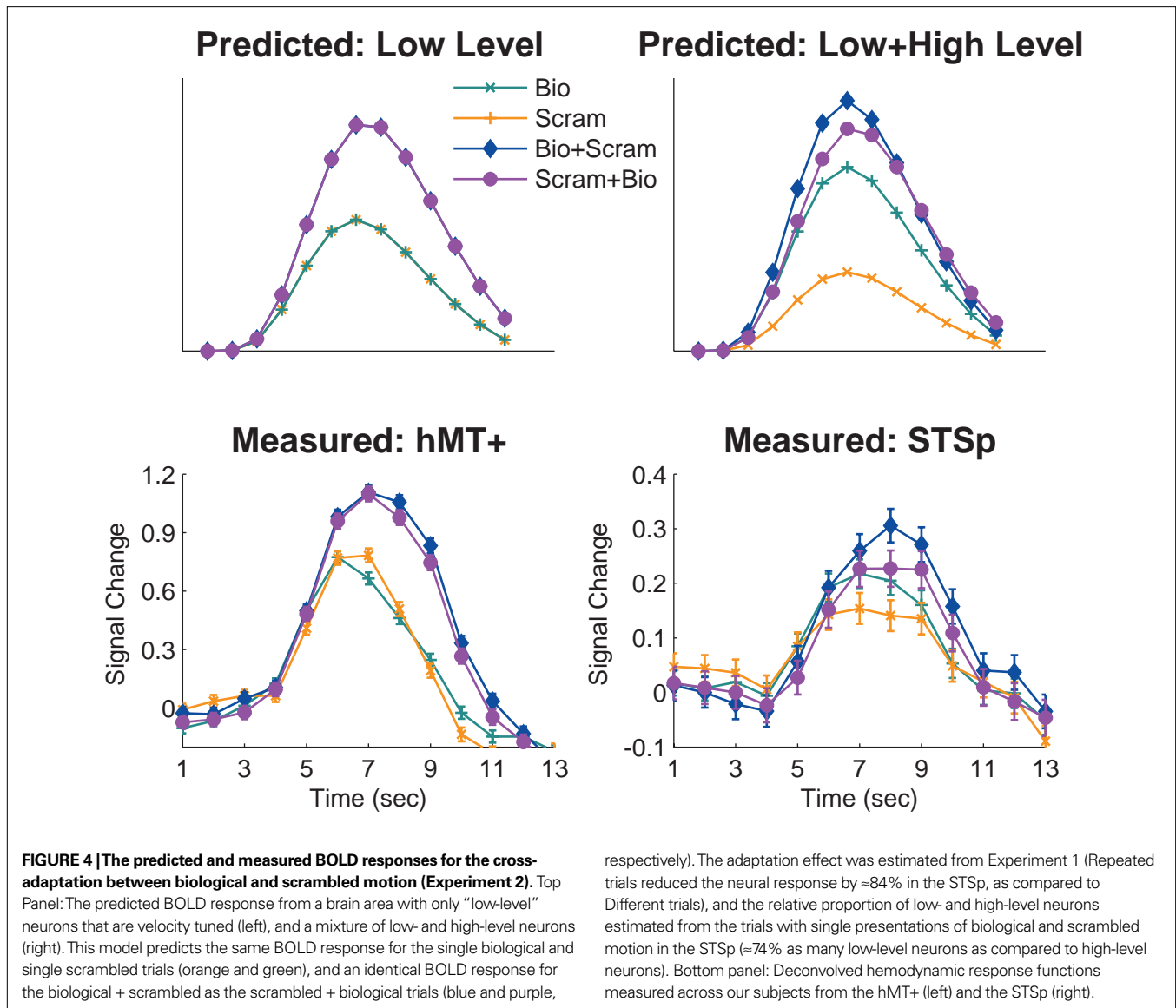
We found that the measured BOLD response in the STSp was indicative of a mixture of low- and high-level neurons, with a significantly stronger response to the biological as compared to scrambled motion ( $t = 3.33, p < 0.0001$ ). Moreover, we found evidence for cross-adaptation, with significantly stronger adaptation when the scrambled animation (the low-level stimulus) preceded the biological animation in the pair ( $t = 2.51, p < 0.01$ ). As a means for comparison, the deconvolved BOLD response in hMT+ favors the scrambled sequences of the biological ( $t = 2.4, p < 0.05$ ), but only weakly differentiates between the order of the pairs in the cross-adaptation trials ( $t = 2.052, p < 0.05$ ). Thus while hMT+ has the neural response expected of a brain area with predominantly low-level velocity tuned responses, the STSp has the signature of a population of neurons encoding the action-specificity of the stimulus.

#### EXPERIMENT 3: POSITION INVARIANCE

In object recognition, it is believed that the ability to recognize an object in different positions of space is supported in part by neurons that discard position information, likely in ventral visual cortex (e.g., Riesenhuber and Poggio, 2002). Moreover, we know from previous research that individuals are able to accurately recognize and discriminate actions, with some limitations, outside foveal vision (Gibson et al., 2005; Ikeda et al., 2005; Thompson et al., 2007). Therefore, in a second set of experiments we explored the extent to which the viewing invariance observed in our initial measurements generalizes to position (in retinal coordinates).

To test for position invariance, we measured the BOLD response for pairs of the same action repeated in the same position (Repeated Foveal) or at two different positions (first foveally, then centered 5.2° into the right or left hemifields). To account for possible





hemifield representations (such as in hMT, Huk et al., 2002), the shifted position trials were analyzed relative to the hemisphere, as Contralateral or Ipsilateral shifts. Results from these measurements are shown in **Figure 5A** with the corresponding statistical analyses in **Table 2**.

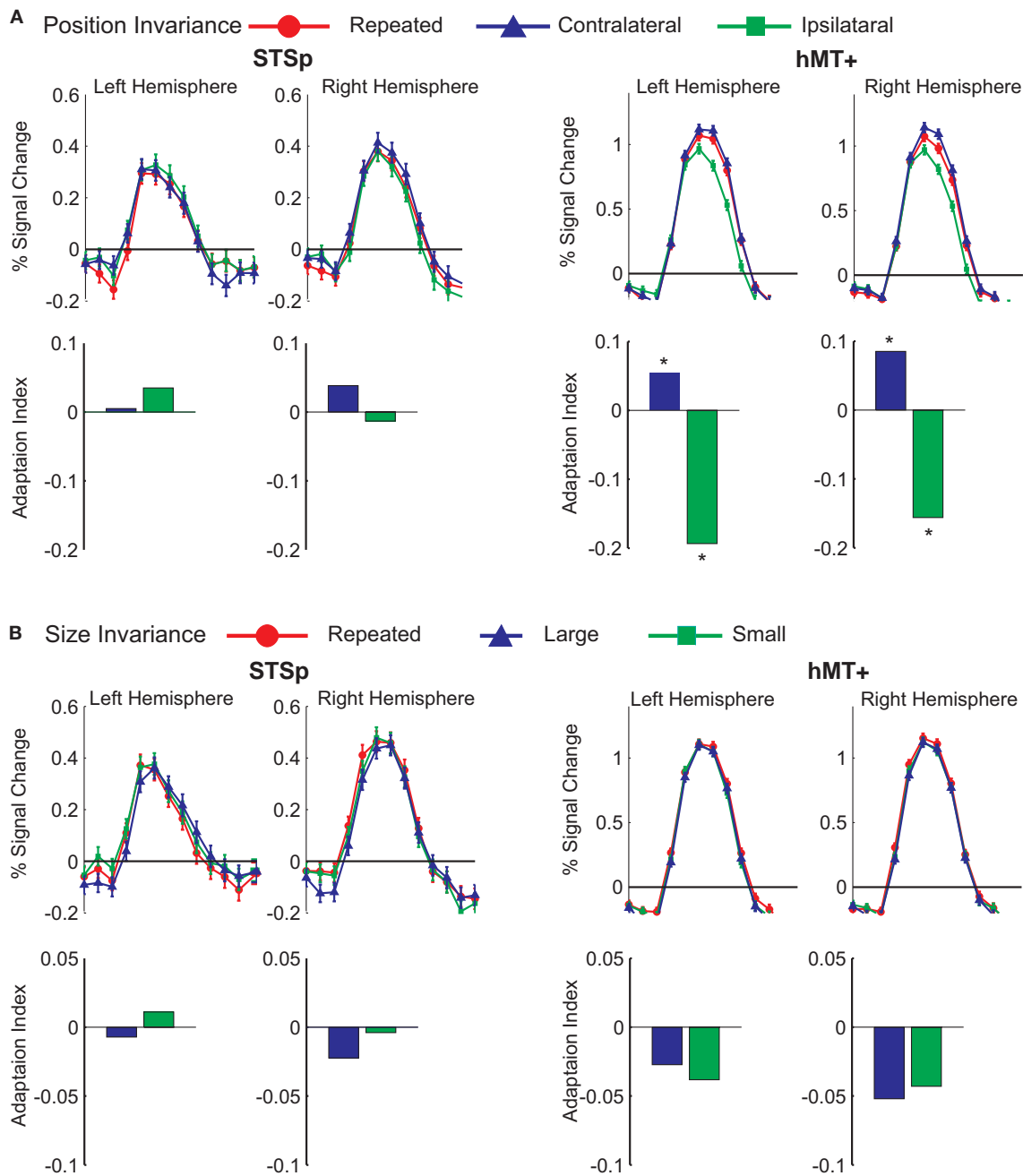
We found no effects of position shift on the BOLD response in the STSp, with the same BOLD responses for Repeated trials as for Ipsilateral and Contralateral shifts. This is the first report of position invariance in the human STS. In contrast, the hMT+ yielded responses that would be expected from a retinotopically mapped brain area, with the largest responses on Contralateral trials, the weakest for Ipsilateral trials, and intermediate responses for foveally Repeated trials.

#### EXPERIMENT 4: SIZE INVARIANCE

We know from psychophysical studies that action recognition is quite stable across changes in size of the figure, from very small (approximately  $2^\circ$ ) up to very large ( $16^\circ$ ) (Ikeda et al., 2005).

We therefore tested for size invariance by measuring the BOLD response for pairs of actions, scaled in size within a range expected to result in approximately equivalent psychophysical performance ( $3\text{--}9^\circ$ ). Subjects viewed repeated presentations of the same action, shown as pairs of animations of the same size (Repeated, medium size), or a medium size animation paired with the same sequence scaled smaller (Small) or larger (Large). Results from these measurements are shown in **Figure 5B** with the corresponding statistical analyses in **Table 2**.

We found no effects of the size changes on the BOLD response from both the STSp and hMT+. Neither of these ROIs had BOLD signals that differentiated the same size conditions (Repeated) from those trials in which the size of the depicted action changed, although the hMT+ response was marginally weaker for the size change conditions as compared to the Repeated trials. The STSp results dovetail with previous single-unit findings of action tuned neurons on the STS that are invariant to changes size perspective (Jellema et al., 2004), and extend them to related cortical areas.



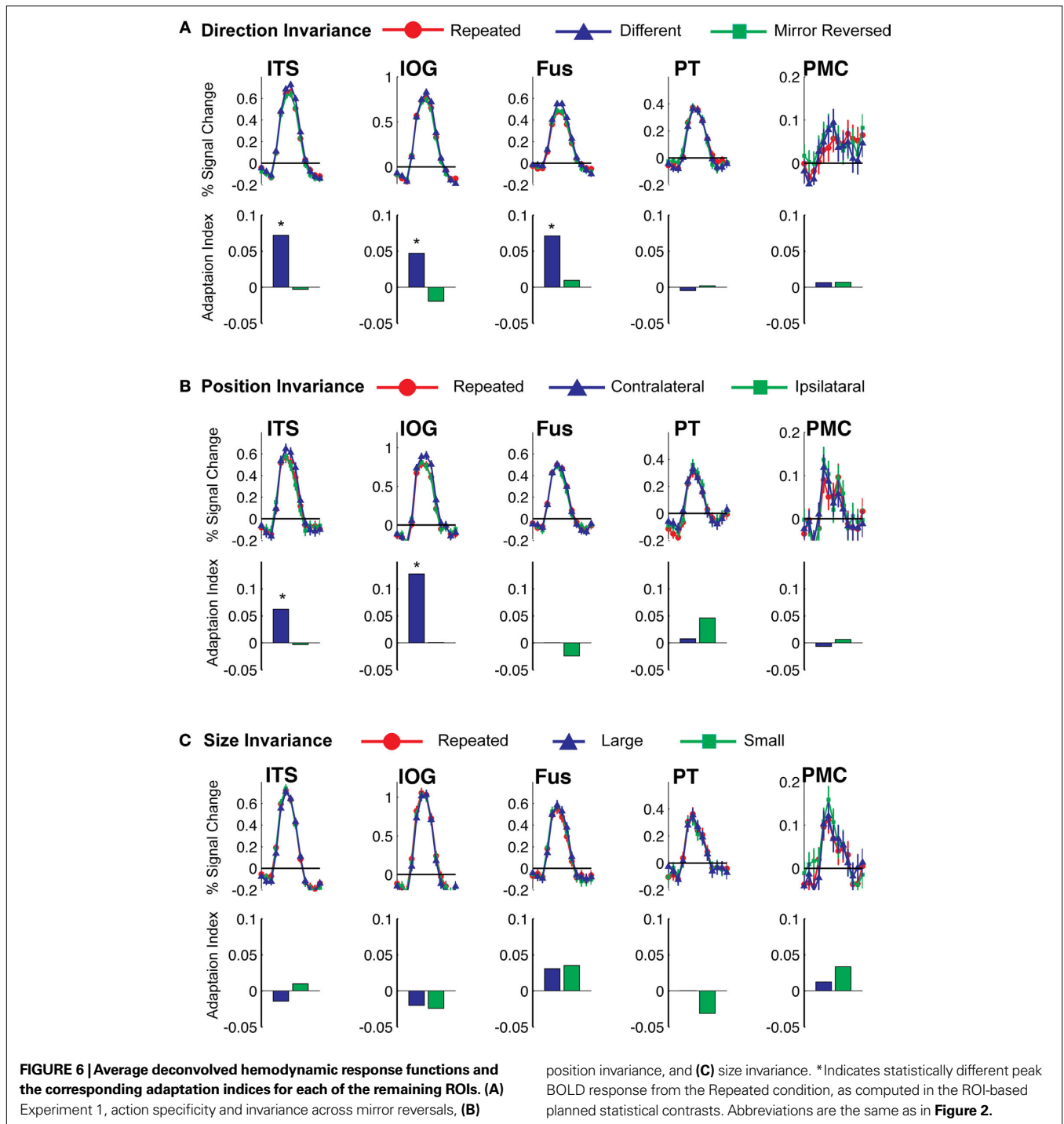
**FIGURE 5 | Average deconvolved hemodynamic response functions and the corresponding adaptation indices for each hemisphere of the STSp and hMT+ for (A) position changes and (B) size changes. \*Indicates statistically significant peak response from the Repeated (foveal or medium) condition, as computed from the deconvolved hemodynamic response functions.**

### OTHER BIOLOGICAL MOTION SELECTIVE ROIs

In a further analysis, we considered the pattern of responses in the additional biological motion selective ROIs that comprise the action recognition network. These include a region on the ITS that may correspond to the EBA, a more ventral region on the inferior occipital sulcus (IOG), a region in ventral temporal cortex (Fus) that may correspond to the FBA, PMC, and a region on the posterior

extent of the PT that is known to have multisensory properties. The results from these ROIs are shown in **Figure 6**, with the corresponding statistical analyses in **Table 3**.

We found evidence for action specificity in the posterior regions of interest (ITS, IOG and Fus), but not in the more anterior PT or frontal PMC. The occipital and temporal ROIs all had BOLD responses that adapted for Repeated trials as compared to Different



action trials. These posterior areas also adapted across the mirror reversal manipulation, suggesting a similar high-level representation as we had found in the STSp.

In a second (non-independent) analysis, we conducted a whole-brain deconvolution analysis probing for brain areas expressing fMR-adaptation BOLD effects (**Figure 7A**). The group GLM analysis revealed patches of adaptation for Repeated actions across anterior occipital and posterior parietal cortex, with small patches in the ventral temporal lobe. These patches include the

STSp (right and left hemispheres), the ITS and IOG in the right hemisphere, and a small region on the fusiform gyrus in the right hemisphere.

In a second whole-brain group analysis, we tested for action-specific, but not exemplar-specific, adaptation by comparing peak BOLD response for the Mirror Reversal trials against the Different trials (**Figure 7B**). This analysis revealed significant patches of adaptation in the right STS and IOG, and bilaterally in the ITS and FG.

**Table 3 | Statistical analyses for the remaining biological motion selective ROIs for Experiments 1, 3 and 4.** Planned statistical contrasts were computed on from the timepoints surrounding the peak response, as deconvolved using a general linear analysis (see Section “Materials and Methods”).

EXPERIMENT 1. DIRECTION INVARIANCE			
ROI	Different–Repeated	Different–Mirror	Mirror–Repeated
ITS	$t = 4.84^{***}$	$t = 5.87^{***}$	$t = -0.96, p = 0.34$
IOG	$t = 3.29^{**}$	$t = 4.64^{***}$	$t = -1.35, p = 0.18$
Fus	$t = 5.57^{***}$	$t = 4.82^{***}$	$t = 0.74, p = 0.46$
PT	$t = 0.29, p = 0.77$	$t = -0.42, p = 0.67$	$t = 0.13, p = 0.90$
PMC	$t = 1.30, p = 0.19$	$t = -0.36, p = 0.72$	$t = 1.67, p = 0.10$
EXPERIMENT 3. POSITION INVARIANCE			
ROI	Contralateral	Ipsilateral	
ITS	$t = 3.51^{**}$	$t = -1.20, p = 0.23$	
IOG	$t = 4.89^{***}$	$t = 0.02, p = 0.98$	
Fus	$t = 0.01, p = 1.0$	$t = -1.55, p = 0.12$	
PT	$t = 0.54, p = 0.59$	$t = 1.47, p = 0.14$	
PMC	$t = 0.96, p = 0.34$	$t = 1.39, p = 0.17$	
EXPERIMENT 4. SIZE INVARIANCE			
ROI	Large	Small	
ITS	$t = -0.90, p = 0.37$	$t = 0.72, p = 0.47$	
IOG	$t = -1.38, p = 0.17$	$t = -0.96, p = 0.34$	
Fus	$t = -0.24, p = 0.81$	$t = 0.05, p = 0.96$	
PT	$t = -1.0, p = 0.31$	$t = -1.74, p = 0.08$	
PMC	$t = 0.30, p = 0.76$	$t = 1.93, p = 0.05$	

Abbreviations are the same as in **Table 1**.

Statistical significance at  $*p < 0.05$ ,  $**p < 0.01$ ,  $***p < 0.0001$ .

When we tested for invariance across shifts in position, we found the neural signals in the ITS and IOG that were stronger for the Contralateral trials as compared to the Repeated trials. This pattern of responses is more consistent with space-based encoding than adaptation. Given the proximity of these regions to retinotopically mapped areas (e.g., Wandell et al., 2007), this is not surprising.

When we tested for invariance across shifts in size of the action target, we found no statistically significant differences in any of our ROIs (although the PMC reached marginal significance for changes to the smaller size). Together, these findings dovetail with previous reports of position invariance in face, body and object processing in the ventral temporal cortex (Fang et al., 2007; Taylor et al., 2009), and extend them to action recognition.

## DISCUSSION

Action recognition is an essential component of daily interpersonal interactions. Neurophysiological studies of biological motion perception have implicated the human STS as the region most critically involved (for review, see Allison et al., 2000). More broadly, there appears to be a network of brain areas across occipital and frontal cortex that supports action recognition, putatively linking recognition to motor planning. To better understand the response properties of the neurons underlying this network, we conducted four experiments to measure the specificity and invariance of neural action encoding.

Using an fMR-adaptation paradigm, we found evidence for repetition suppression across cortex when observers viewed pairs of identical actions as compared to when they viewed pairs of two different actions. This basic adaptation finding in the STSp is consistent with the hypothesis that unique neural networks code different actions, with each of the component brain areas exhibiting stimulus specific adaptation. We also found repetition suppression in these brain areas when subjects viewed the same actions mirror-reversed, suggesting some invariance across changes in viewing perspective. To determine whether this adaptation effect is due to low-level encoding of velocity features (only some of which are altered in mirror reversals) or high-level action encoding, we tested for cross-adaptation between biological and motion-matched scrambled sequences. We found evidence for the existence of action-tuned neurons in the STSp, with likely only low-level velocity coding in the hMT+. Overall, these findings in STSp are consistent with the hypothesis that the STS is organized with subpopulations of neurons tuned to different actions, a proposal initially based on earlier single-unit findings in monkeys (Oram and Perrett, 1994).

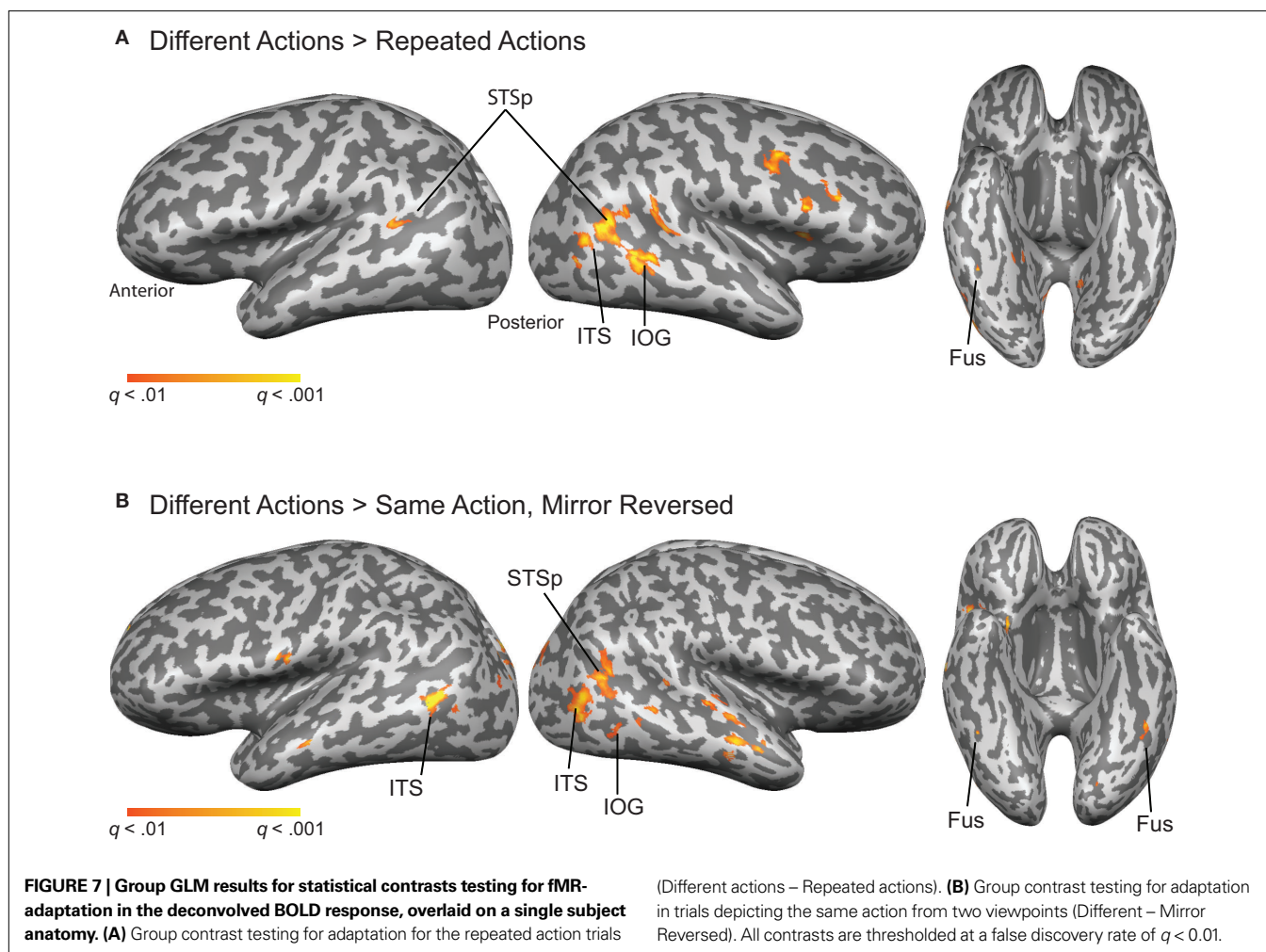
In two additional experiments, we measured for neural invariance across common perspective changes, namely shifts in position and changes in size. The STSp was invariant to changes across both of these conditions, with the same peak BOLD response as when there was no perspective shift. This was mirrored in the fusiform, and is in contrast to the hMT+, ITS and IOG that all coded for position.

These findings are evidence for high-level representations of actions in subpopulations of neurons in the STSp, are consistent with previous single-unit reports of functional specialization within the monkey STSa, and are further evidence for functional homologies across the STS of the two species. Because we found similar results in the fusiform, these findings also suggest that much of the visual processing engaged during action recognition is abstracted away from retinal coordinates.

## ACTION RECOGNITION AS A REPRESENTATIONAL CHALLENGE

Action recognition could be considered as analogous to object recognition, with all the associated computational difficulties connected to properly assigning unique and stable identities to objects through the range of changes in viewing perspective that commonly occur in natural vision. There are caveats to this analogy, however, because human actions are dynamic, defined by a sequence of changing form over time, and are constrained by the inherent structure and connectivity of the human body. Thus one might infer that a range of action representations may exist in the brain, from stationary form-based “snapshots” of specific postures (as have been proposed by a number of computational and experimental studies, Giese and Poggio, 2003; Lange et al., 2006; Vangeneugden et al., 2009), to the encoding of action specific dynamic patterns (e.g., sprites or space-time fragments, Cavanagh, 1992; Davies and Hoffman, 2003), to object-oriented representations that encode underlying body structure (Oram and Perrett, 1996; Jellema and Perrett, 2006).

We know from neurophysiological studies of object vision that both viewpoint-specific and viewpoint-invariant neural representations are likely to exist. For example, single-unit studies have revealed neurons in monkey inferotemporal cortex that are sharply



tuned to complex objects, with nearby neurons tuned more broadly to a variety of complex shapes (e.g., Tanaka, 1996). Together these neurons form a sort of robust population code that supports generalization of object recognition across viewpoint-specific neural representations (DiCarlo and Cox, 2007). Our findings are evidence for the same type of heterogeneous population coding in visual cortex during action recognition.

#### ACTION ENCODING IN THE MONKEY STS

The hypotheses for these experiments were drawn primarily from previous single-unit physiology studies of action recognition targeting the anterior portions of the monkey STS (STSa). The STSa, sometimes called the anterior STPa or the temporal–parietal–occipital area (TPO, Bruce et al., 1981; Pandya and Seltzer, 1982; Cusick et al., 1995) is a heterogeneous, higher-order brain area that is a convergence zone for the visual, auditory and somatosensory systems (Jones and Powell, 1970; Seltzer and Pandya, 1978; Cusick, 1996). Accordingly, neurons in this area have complex response properties including the tendency to be multi-modal, with vision being the dominant modality (Bruce et al., 1981). The visually responsive cells tend to have large receptive fields, many of which are best driven by specific moving patterns, such as the movements of faces, heads, arms and bodies (Perrett et al., 1985a). More recent

studies have found these biological motion selective neurons to be intermixed with neurons that respond equally well to stationary images of body postures, perhaps with some patchy organization (Perrett et al., 1985a; Oram and Perrett, 1994; Nelissen et al., 2006; Vangeneugden et al., 2009). Together these single-unit studies have documented the response features necessary for shape and motion analysis on the monkey STSa during action recognition (for review, see Puce and Perrett, 2003).

Of the biological motion tuned neurons, the vast majority is highly specific (Oram and Perrett, 1996; Jellema et al., 2004). These neurons require particular combinations of body movements and postures to generate the strongest neuronal firing. For example, a neuron may fire for a leftward facing actor walking forward, but not when that actor bends forward, walks backwards or faces the opposite direction. Nearly 90% of the biological motion tuned neurons have this property, which renders them effectively action and viewpoint-specific.

Even with action specificity, many of these STS neurons generalize over a range of changes in the visual scene that accompany shifts in viewing perspective, such as changes in the actor size (which is associated with near versus far viewing) and in position across the visual field (Jellema and Perrett, 2006). Computationally, response invariance over these types of properties can be built from

pooling inputs across a range of more narrowly tuned subunits (e.g., Riesenhuber and Poggio, 1999; Giese and Poggio, 2003), and is supported by the anatomical evidence for highly refined input to the monkey STS, and by single-unit evidence for large receptive fields in these neurons.

A smaller minority of the STSa neurons has response properties that generalize over a much larger range of views. For example, some STS neurons will generalize across depictions of actions as shown in live action viewing, in movies, as stick figures or in point-light form (Bruce et al., 1981; Oram and Perrett, 1994). That these neurons generalize across many of these low-level changes is evidence for some abstracted representation of actions in the tuning of STS neurons, sometimes discussed as object-centered encoding (as opposed to strictly viewpoint specific), or as more abstracted goal-centered coding. In both interpretations the emphasis is on the ability of these STSa neurons to generalize across relatively large changes in the visual depictions of the actions, without loss of specificity for the action itself.

Our measurements provide evidence for these same types of computations (action specificity and invariance to mirror reversals, size, and position) on the human STS. It is currently not clear from our measurements the relative proportions of viewpoint specific and object-oriented (e.g., size and position invariant) sub-populations of neurons, although we do not have any reason to believe them to be different from that measured in the monkey. From single unit measures, researchers have estimated the more viewpoint-invariant neurons to represent a minority. Our population BOLD measures found equivalent fMR-adaptation in the STSp for many of our perspective changes, which would seem to imply at least a substantive number so as to generate the same response suppression. However, because there is not yet a quantitative link between the magnitude of the adaptation effect and relative numbers of neurons underlying the BOLD response, it remains an open question of whether the human STS is more or less densely populated with neurons coding these viewpoint invariant responses.

### **ACTION ENCODING OUTSIDE OF THE STS**

We also found evidence for viewing perspective invariances in ROIs outside of the STS, including the more ventral and lateral temporal lobe (ITS, IOG), and in ventral temporal cortex (Fus). Unfortunately, we have relatively little knowledge of the monkey homologies to these areas, and even within human fMRI studies, only a handful of studies have targeted these regions. As a result, we have little physiological data on which to build hypotheses as the underlying neural tuning and the relative contributions to action perception as a whole. However, we can make predictions on the basis of computational models, which have suggested that more ventral brain areas serve to extract key body postures the are embedded within action sequences (Giese and Poggio, 2003), with some theories suggesting that these key postures alone could be sufficient to support action recognition (Lange et al., 2006). Under such a scheme, one could then hypothesize that neurons in these areas should be largely action specific (to the extent that each action relies on a unique set of key body postures), and may also be narrowly tuned across mirror reversals. One would expect, however, that this tuning is position and size invariant, within some suitable range, as many ventral temporal neurons tuned to complex objects are.

We found little evidence for action specificity in the more anterior PMC, although this may largely be attributed to the low amplitude of the BOLD response and the relatively high variance in that area. PMC is believed to be involved in action planning, and as such would be anticipated to have a relatively high-level representation of actions (Rizzolatti and Craighero, 2004), however the existence of neurons that generalize across perceptual and motor representations of actions in human PMC is currently a very contentious scientific debate (Hickok, 2009).

### **NEUROIMAGING STUDIES OF VIEWPOINT SPECIFICITY ON THE STS**

It is important to note that our study is not the first to investigate neural tuning on the human STS, particularly in the context of action encoding. A previous investigation using animated clips of actors performing various activities (such as playing basketball or doing cartwheels) found repetition suppression on the STS for previously viewed action sequences, even when those actions were executed by a new actor and seen several minutes later (Kable and Chatterjee, 2006). Interestingly, this study also found these repetition effects in the hMT+ and the EBA, two brain areas that appear to be involved in motion analysis and shape recognition during action recognition (Peuskens et al., 2005; Peelen et al., 2006).

More recently, Lestou et al. (2008) investigated the role of low-level visual features and inferred action goals in the encoding of hand movements. They found the STS to discriminate between trials depicting sequential pairs of the same hand movements as compared to different movements, the basic fMR-adaptation effect for repeated action sequences. Using morphed sequences of the same actions, they also measured adaptation for action sequences depicting the same overall goal while controlling for low-level kinematic patterns. They found fMR-adaptation for actions that had the same implied goal whether played forwards or backwards (e.g., waving), but a release from adaptation when playback implied different implied goals (e.g., lifting up versus lowering down). This study thus argues for goal-oriented encoding on the STS that is abstracted away from the low-level kinematics of the action sequences.

### **SPACE-BASED ENCODING ON THE STS**

Although retinotopic organization is a hallmark of visual brain areas, relatively little is known about any space-based organization outside of occipital cortex. Single-unit findings have demonstrated monkey STSa receptive fields to be large, across all sensory modalities. For example, the STSa neurons that fire for somatotopic input tend to respond to touch on any part of the body, not just for a single limb (Mistlin and Perrett, 1990). Visual receptive fields likewise are large (on average, greater than 30° of visual angle) and often cross into the ipsilateral visual field. There is some evidence that even with classical receptive fields covering such a large region of space, there is a tendency for neurons to fire stronger for visual patterns in the contralateral field or in foveal vision, as compared to stimuli restricted to the ipsilateral field (Bruce et al., 1981).

There are few studies of retinotopy on the human STS, with the findings beginning to converge on a similar space-based representation as that found in monkey. Retinotopic maps generated with “traveling waves” of point-light biological motion reveal a crude map of the contralateral field, particularly in the fundus and lateral

surface of the STS (Saygin and Sereno, 2008). BOLD responses were “virtually nonexistent” for ipsilateral viewing, and, like much of visual cortex, the strength of the responses was strongly modulated by attention. We found weak, but not significant, evidence for space-based encoding in the right hemisphere STSp, which had slightly stronger BOLD responses for trials in which the action was shifted to the contralateral hemifield.

A recent study by Michels et al. (2009) using parafoveally viewed point-light walkers found strong BOLD signals in the right hemisphere STS for biological motion viewed on the right or left, while BOLD signals in the left hemisphere STS were highest only for the contralateral trials. Together these findings would seem to imply some space-based encoding (and in particular, an ipsilateral hemifield representation) in the right hemisphere that does not exist in the left STS. The behavioral consequence of this inequitable retinotopic organization is not entirely clear, as observers can discriminate biological motion in the near periphery with relatively good accuracy, and with no clear bias for viewing in the right hemifield as compared to the left (Gibson et al., 2005; Ikeda et al., 2005; Thompson et al., 2007). An exception is an apparent facing effect, which appears to have some basis in the neural signals in ventral temporal cortex, but not the STS (de Lussanet et al., 2008).

### SIZE INVARIANCE

To the best of our knowledge, ours is the first neuroimaging study to target biological motion size invariance in human cortex. Psychophysical studies have demonstrated that point-light animations ranging from approximately 2–8° are all discriminated with the same apparent ease, with sharp declines in performance for smaller stimuli and slower deterioration at larger sizes (Ikeda et al., 2005). Thus biological motion recognition is robust across a wide range of distance-based viewing perspectives.

Scientists have used a wide range of stimulus sizes in neurophysiological measurements, with little attempt to control this parameter directly. An exception is a single-unit study that found approximately half the biological motion selective STS neurons to tolerate approximately two-fold increases in viewing distances (Jellema et al., 2004). When actions are held stationary, however, this tolerance appears to be minimized, with invariance across relatively small changes in size (e.g., within 6°), and only a small population of STS neurons tolerating larger size changes (approximately 15°, Ashbridge et al., 2000).

In human neurophysiological studies there have been no published reports that we have found, that directly test for size specificity in the biological motion response. Researchers across studies have used a wide range of stimulus sizes, from 3° of visual angle (Garcia and Grossman, 2008) up to 11° (Vaina et al., 2001a). Often the stimulus size is not even reported.

A weakness of our findings was the lack of size differentiation in any of the ROIs that we measured, with all having equivalent neural responses for the size change conditions as compared to the Repeated size trials. We have interpreted this as evidence for response invariance, however one could argue that this interpretation would require inclusion of conditions that result in a release from adaptation (much like the Different Action condition in Experiment 1). It is therefore possible that our selection of stimulus

sizes, chosen within the range associated with equivalent behavioral performance, contributed to similarity in the BOLD response across all three conditions.

We consider, however, that size and position invariance would be expected from brain areas that encode actions as unique, even when viewed from different vantage points. Computationally, size and position invariance have been modeled in feed-forward networks that pool across subunits with smaller receptive fields, taking the maximum response across the entire pool of units (Giese and Poggio, 2003). This winner-take-all computation benefits from the narrow tuning properties inherited from the subunits, without loss of generality across the entire range of tuning features in the underlying subunits. Interestingly, we found evidence for size invariance across most of the ROIs we measured, suggesting this feature may be passed up along the entire network of brain areas supporting action recognition.

### IMPLICATIONS FOR MODELS OF SOCIAL PERCEPTION

The human STS is a large expanse of cortex that extends from the anterior pole of the temporal lobe to the posterior aspects of posterior parietal cortex. As should be expected from such a large portion of neural real estate, the STS appears to be engaged in many functions specializations, including some that appear to be uniquely human. These include speech perception on the more anterior aspects (Hickok and Poeppel, 2000), face and body perception on the more central aspects (Haxby et al., 2000; Campbell et al., 2001; Materna et al., 2008), and brain systems supporting social awareness on the most posterior and dorsal extent (Martin and Weisberg, 2003; Saxe et al., 2004; Gobbini et al., 2007; Mitchell, 2008).

There has been little agreement on how to synthesize the human experimental findings related to STS functional specialization, to the extent that some researchers to wonder whether it is appropriate to assume that the STS is organized by functional specificity (Hein and Knight, 2008). Others have argued that the common thread though the wide-range of reported functional specialization is in successfully navigating social interactions, which has led to the dubious label of the STS as the “social brain” (Brothers, 2002).

Attempts to resolve these issues of functional localization using meta-analyses and even within-subject comparisons have largely been inconclusive (e.g., Pelphrey et al., 2005; Gobbini et al., 2007; Hein and Knight, 2008; Mitchell, 2008). This suggests that the underlying structure to the functional organization on the STS exists at the subvoxel level, and thus demands the application of newer, more sophisticated experimental techniques.

There is basis for this idea in the single-unit physiology, which finds neurons responsive to biological motion intermixed with those that respond to a much wider range of moving and stationary patterns. Even within the action-tuned neurons, there appears to be much overlap between those that are strictly viewpoint specific and those that have viewpoint invariant tuning (e.g., discriminate more clearly between actions than between different viewpoints of the same action). From this evidence there is every reason to anticipate the STS to have heterogeneous functional specialization within close anatomical proximity. In one unifying proposal, researchers have argued that together these neurons form a hierarchical network of action recognition, with viewpoint-tuned units pooling inputs

into units that code specific actions, which then feed into more abstracted goal-oriented computations (e.g., Jellema and Perrett, 2006). Thus, the multiplexed network serves multiple functional specializations, including action recognition and social cognition, depending on the level of analysis. Under this view, studying the voxel-wide tunings for biological motion representations reveals more about the properties of the region as a whole.

This study, using the fMR-adaptation technique to identify sub-populations of neurons underlying the BOLD response, was designed to measure some of the properties of this network. Our results show that the biological motion-sensitive STS is largely invariant to changes in direction, position, and size. Moreover, the STSp had neural signals consistent with a mixture of interdependent high- and low-level neural populations, as evidenced by the asymmetry in our cross-adaptation experiments. Thus our findings are compatible with theories of social cognition that propose hierarchically abstracted action representations.

## REFERENCES

- Adolphs, R. (2009). The social brain: neural basis of social knowledge. *Annu. Rev. Psychol.* 60, 693–716.
- Allison, T., Puce, A., and McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends Cogn. Sci.* 4, 267–278.
- Andresen, D. R., Vinberg, J., and Grill-Spector, K. (2009). The representation of object viewpoint in human visual cortex. *Neuroimage* 45, 522–536.
- Andrews, T. J., and Ewbank, M. P. (2004). Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *Neuroimage* 23, 905–913.
- Ashbridge, E., Perrett, D. I., Oram, M. W., and Jellema, T. (2000). Effect of image orientation and size on object recognition: responses of single units in the macaque monkey temporal cortex. *Cogn. Neuropsychol.* 17, 13–34.
- Barraclough, N. E., Xiao, D., Oram, M. W., and Perrett, D. I. (2006). The sensitivity of primate STS neurons to walking sequences and to the degree of articulation in static images. *Prog. Brain Res.* 154, 135–148.
- Battelli, L., Cavanagh, P., and Thornton, I. M. (2003). Perception of biological motion in parietal patients. *Neuropsychologia* 41, 1808–1816.
- Baylis, G. C., and Rolls, E. T. (1987). Responses of neurons in the inferior temporal cortex in short term and serial recognition memory tasks. *Exp. Brain Res.* 65, 614–622.
- Beauchamp, M. S., Lee, K. E., Haxby, J. V., and Martin, A. (2003). fMRI responses to video and point-light displays of moving humans and manipulable objects. *J. Cogn. Neurosci.* 15, 991–1001.
- Bedny, M., Caramazza, A., Grossman, E., Pascual-Leone, A., and Saxe, R. (2008). Concepts are more than percepts: the case of action verbs. *J. Neurosci.* 28, 11347–11353.
- Blake, R., and Shiffrar, M. (2007). Perception of human motion. *Annu. Rev. Psychol.* 58, 47–73.
- Bonda, E., Petrides, M., Ostry, D., and Evans, A. (1996). Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *J. Neurosci.* 16, 3737–3744.
- Boynton, G. M., Demb, J. B., Glover, G. H., and Heeger, D. J. (1999). Neuronal basis of contrast discrimination. *Vision Res.* 39, 257–269.
- Boynton, G. M., Engel, S. A., Glover, G. H., and Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. *J. Neurosci.* 16, 4207–4221.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vis.* 10, 437–442.
- Brothers, L. (2002). The social brain: a project for integrating primate behavior and neurophysiology in a new domain. In *Foundations in Social Neuroscience*, J. T. Cacioppo, G. G. Berntson, S. E. Taylor and D. L. Schacter, eds (Cambridge, MIT Press), 1357 p.
- Bruce, C., Desimone, R., and Gross, C. G. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *J. Neurophysiol.* 46, 369–384.
- Buckner, R. L., Goodman, J., Burock, M., Rotte, M., Koutstaal, W., Schacter, D., Rosen, B., and Dale, A. M. (1998). Functional-anatomic correlates of object priming in humans revealed by rapid presentation event-related fMRI. *Neuron* 20, 285–296.
- Campbell, R., MacSweeney, M., Surguladze, S., Calvert, G., McGuire, P., Suckling, J., Brammer, M. J., and David, A. S. (2001). Cortical substrates for the perception of face actions: an fMRI study of the specificity of activation for seen speech and for meaningless lower-face acts (gurning). *Brain Res. Cogn. Brain Res.* 12, 233–243.
- Casile, A., and Giese, M. A. (2005). Critical features for the recognition of biological motion. *J. Vis.* 5, 348–360.
- Cavanagh, P. (1992). Attention-based motion perception. *Science* 257, 1563–1565.
- Cusick, C. G. (1996). The superior temporal polysensory region in monkeys. In *Extrastriate Cortex in Primates*, Vol. 14, K. S. Rockland, J. H. Kaas and A. Peters, eds (New York, Plenum Press), pp. 435–468.
- Cusick, C. G., Seltzer, B., Cola, M., and Griggs, E. (1995). Chemoarchitectonics and corticocortical terminations within the superior temporal sulcus of the rhesus monkey: evidence for subdivisions of superior temporal polysensory cortex. *J. Comp. Neurol.* 360, 513–535.
- Dale, A. M. (1999). Optimal experimental design for event-related fMRI. *Hum. Brain Mapp.* 8, 109–114.
- Dale, A. M., and Buckner, R. L. (1997). Selective averaging of rapidly presented individual trials using fMRI. *Hum. Brain Mapp.* 5, 329–340.
- Davies, T. N., and Hoffman, D. D. (2003). Facial attention and spacetime fragments. *Axiomathes* 13, 303–327.
- de Lussanet, M. H., Fadiga, L., Michels, L., Seitz, R. J., Kleiser, R., and Lappe, M. (2008). Interaction of visual hemifield and body view in biological motion perception. *Eur. J. Neurosci.* 27, 514–522.
- DiCarlo, J. J., and Cox, D. D. (2007). Untangling invariant object recognition. *Trends Cogn. Sci.* 11, 333–341.
- Fahy, F. L., Riches, I. P., and Brown, M. W. (1993). Neuronal activity related to visual recognition memory: long-term memory and the encoding of recency and familiarity information in the primate anterior and medial inferior temporal and rhinal cortex. *Exp. Brain Res.* 96, 457–472.
- Fang, F., Murray, S. O., and He, S. (2007). Duration-dependent fMRI adaptation and distributed viewer-centered face representation in human visual cortex. *Cereb. Cortex* 17, 1402–1411.
- Ganel, T., Gonzalez, C. L., Valyear, K. F., Culham, J. C., Goodale, M. A., and Kohler, S. (2006). The relationship between fMRI adaptation and repetition priming. *Neuroimage* 32, 1432–1440.
- Garcia, J. O., and Grossman, E. D. (2008). Necessary but not sufficient: motion perception is required for perceiving biological motion. *Vision Res.* 48, 1144–1149.
- Genovese, C. R., Lazar, N. A., and Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage* 15, 870–878.
- Gibson, L. A., Sadr, J., Troje, N. F., and Nakayama, K. (2005). Perception of biological motion at varying eccentricity. *J. Vis.* 5, 16a.
- Giese, M. A., and Poggio, T. (2003). Neural mechanisms for the recognition of biological movements. *Nat. Rev. Neurosci.* 4, 179–192.
- Gobbini, M. I., Koralek, A. C., Bryan, R. E., Montgomery, K. J., and Haxby, J. V. (2007). Two takes on the social brain: a

## CONCLUSION

To the extent that the fMR-adaptation technique can identify unique populations of neurons underlying the voxel BOLD response, our findings report on the tuning properties of neurons on the human STS. Our findings support the existence of unique representations for different actions on the STS that are largely position and size invariant. We should note that, much like object-tuned responses from individual neurons in ventral temporal cortex, it is presumed that these highly tuned individual neurons represent small units within a much larger network. It is this larger network that is likely to be best detected by the fMRI BOLD responses reported in this study.

## ACKNOWLEDGMENTS

This work was supported in part by NSF BCS0748314 to Emily D. Grossman, and in part by a UROP grant from University of California Irvine to Nicole L. Jardine.



- comparison of theory of mind tasks. *J. Cogn. Neurosci.* 19, 1803–1814.
- Grill-Spector, K. (2006). Selectivity of adaptation in single units: implications for fMRI experiments. *Neuron* 49, 170–171.
- Grill-Spector, K., Henson, R., and Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn. Sci.* 10, 14–23.
- Grill-Spector, K., and Malach, R. (2001). fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychol. (Amst)* 107, 293–321.
- Grossman, E., Battelli, L., and Pascual-Leone, A. (2005). TMS over STSp disrupts perception of biological motion. *Vision Res.* 45, 2847–2853.
- Grossman, E., and Blake, R. (2002). Brain areas active during visual perception of biological motion. *Neuron* 35, 1157–1165.
- Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., and Blake, R. (2000). Brain areas involved in perception of biological motion. *J. Cogn. Neurosci.* 12, 711–720.
- Grossman, E. D., and Blake, R. (2001). Brain activity evoked by inverted and imagined biological motion. *Vision Res.* 41, 1475–1482.
- Grossman, E. D., Kim, C. Y., and Blake, R. (2004). Learning to see biological motion: Brain activity parallels behavior. *J. Cogn. Neurosci.* 16, 1–11.
- Haxby, J. V., Hoffman, E. A., and Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends Cogn. Sci.* 4, 223–233.
- Hein, G., and Knight, R. T. (2008). Superior temporal sulcus – it's my area: or is it? *J. Cogn. Neurosci.* 20, 2125–2136.
- Henson, R. N., Cansino, S., Herron, J. E., Robb, W. G., and Rugg, M. D. (2003). A familiarity signal in human anterior medial temporal cortex? *Hippocampus* 13, 301–304.
- Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *J. Cogn. Neurosci.* 21, 1229–1243.
- Hickok, G., and Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends Cogn. Sci.* 4, 131–138.
- Hickok, G., and Saberi, K. (in press). Redefining the functional organization of the planum temporale region: space, objects and sensory-motor integration. In Springer Handbook on Auditory Research: Human Auditory Cortex, D. Poeppel, T. Overath, A. N. Popper and R. R. Fay, eds (Springer).
- Huk, A. C., Dougherty, R. F., and Heeger, D. J. (2002). Retinotopy and functional subdivision of human areas MT and MST. *J. Neurosci.* 22, 7195–7205.
- Huk, A. C., Ress, D., and Heeger, D. J. (2001). Neuronal basis of the motion aftereffect reconsidered. *Neuron* 32, 161–172.
- Ikeda, H., Blake, R., and Watanabe, K. (2005). Eccentric perception of biological motion is unscalably poor. *Vision Res.* 45, 1935–1943.
- Jellema, T., Baker, C. I., Wicker, B., and Perrett, D. I. (2000). Neural representation for the perception of the intentionality of actions. *Brain Cogn.* 44, 280–302.
- Jellema, T., Maassen, G., and Perrett, D. I. (2004). Single cell integration of animate form, motion and location in the superior temporal cortex of the macaque monkey. *Cereb. Cortex* 14, 781–790.
- Jellema, T., and Perrett, D. I. (2006). Neural representations of perceived bodily actions using a categorical frame of reference. *Neuropsychologia* 44, 1535–1546.
- Jones, E. G., and Powell, T. P. (1970). An anatomical study of converging sensory pathways within the cerebral cortex of the monkey. *Brain* 93, 793–820.
- Kable, J. W., and Chatterjee, A. (2006). Specificity of action representations in the lateral occipitotemporal cortex. *J. Cogn. Neurosci.* 18, 1498–1517.
- Kourtzi, Z., and Kanwisher, N. (2000). Activation in human MT/MST by static images with implied motion. *J. Cogn. Neurosci.* 12, 48–55.
- Kourtzi, Z., and Kanwisher, N. (2001). Representation of perceived object shape by the human lateral occipital complex. *Science* 293, 1506–1509.
- Krekelberg, B., Boynton, G. M., and van Wezel, R. J. (2006). Adaptation: from single cells to BOLD signals. *Trends Neurosci.* 29, 250–256.
- Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., and Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nat. Neurosci.* 12, 535–540.
- Lange, J., Georg, K., and Lappe, M. (2006). Visual perception of biological motion by form: a template-matching analysis. *J. Vis.* 6, 836–849.
- Lestou, V., Pollick, E. E., and Kourtzi, Z. (2008). Neural substrates for action understanding at different description levels in the human brain. *J. Cogn. Neurosci.* 20, 324–341.
- Li, L., Miller, E. K., and Desimone, R. (1993). The representation of stimulus familiarity in anterior inferior temporal cortex. *J. Neurophysiol.* 69, 1918–1929.
- Martin, A., and Weisberg, J. (2003). Neural foundations for understanding social and mechanical concepts. *Cogn. Neuropsychol.* 20, 575–587.
- Materna, S., Dicke, P. W., and Thier, P. (2008). Dissociable roles of the superior temporal sulcus and the intraparietal sulcus in joint attention: a functional magnetic resonance imaging study. *J. Cogn. Neurosci.* 20, 108–119.
- Meyer, M., Baumann, S., Marchina, S., and Jancke, L. (2007). Hemodynamic responses in human multisensory and auditory association cortex to purely visual stimulation. *BMC Neurosci.* 8, 14.
- Michels, L., Kleiser, R., de Lussanet, M. H., Seitz, R. J., and Lappe, M. (2009). Brain activity for peripheral biological motion in the posterior superior temporal gyrus and the fusiform gyrus: dependence on visual hemifield and view orientation. *Neuroimage* 45, 151–159.
- Michels, L., Lappe, M., and Vaina, L. M. (2005). Visual areas involved in the perception of human movement from dynamic form analysis. *Neuroreport* 16, 1037–1041.
- Mistlin, A. J., and Perrett, D. I. (1990). Visual and somatosensory processing in the macaque temporal cortex: the role of 'expectation'. *Exp. Brain Res.* 82, 437–450.
- Mitchell, J. P. (2008). Activity in right temporo-parietal junction is not selective for theory-of-mind. *Cereb. Cortex* 18, 262–271.
- Muller, J. R., Metha, A. B., Krauskopf, J., and Lennie, P. (1999). Rapid adaptation in visual cortex to the structure of images. *Science* 285, 1405–1408.
- Musselwhite, M. J., and Jeffreys, D. A. (1983). Visual evoked potentials to double-pulse pattern presentation. *Vision Res.* 23, 135–143.
- Nelissen, K., Vanduffel, W., and Orban, G. A. (2006). Charting the lower superior temporal region, a new motion-sensitive region in monkey superior temporal sulcus. *J. Neurosci.* 26, 5929–5947.
- Ono, M., Kubie, S., and Abernathy, C. D. (1991). Atlas of the Cerebral Sulci. Stuttgart, Georg Thieme Verlag.
- Oram, M. W., and Perrett, D. I. (1994). Responses of anterior superior temporal polysensory (STPa) neurons to "biological motion" stimuli. *J. Cogn. Neurosci.* 6, 99–116.
- Oram, M. W., and Perrett, D. I. (1996). Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. *J. Neurophysiol.* 76, 109–129.
- Pandya, D. N., and Seltzer, B. (1982). Intrinsic connections and architectonics of posterior parietal cortex in the rhesus monkey. *J. Comp. Neurol.* 204, 196–210.
- Pavlova, M., Staudt, M., Sokolov, A., Birbaumer, N., and Krageloh-Mann, I. (2003). Perception and production of biological movements in patients with early periventricular brain lesions. *Brain* 126(Pt. 3), 692–701.
- Peelen, M. V., Glaser, B., Vuilleumier, P., and Eliez, S. (2009). Differential development of selectivity for faces and bodies in the fusiform gyrus. *Dev. Sci.* 12, F16–F25.
- Peelen, M. V., Wiggett, A. J., and Downing, P. E. (2006). Patterns of fMRI activity dissociate overlapping functional brain areas that respond to biological motion. *Neuron* 49, 815–822.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vis.* 10, 437–442.
- Pelphrey, K. A., Morris, J. P., Michelich, C. R., Allison, T., and McCarthy, G. (2005). Functional anatomy of biological motion perception in posterior temporal cortex: an fMRI study of eye, mouth and hand movements. *Cereb. Cortex* 15, 1866–1876.
- Perrett, D. I., Smith, P. A., Mistlin, A. J., Chitty, A. J., Head, A. S., Potter, D. D., Broennimann, R., Milner, A. D., and Jeeves, M. A. (1985a). Visual analysis of body movements by neurones in the temporal cortex of the macaque monkey: a preliminary report. *Behav. Brain Res.* 16, 153–170.
- Perrett, D. I., Smith, P. A., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., and Jeeves, M. A. (1985b). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proc. R. Soc. Lond., B, Biol. Sci.* 223, 293–317.
- Peuskens, H., Vanrie, J., Verfaillie, K., and Orban, G. A. (2005). Specificity of regions processing biological motion. *Eur. J. Neurosci.* 21, 2864–2875.
- Puce, A., and Perrett, D. (2003). Electrophysiology and brain imaging of biological motion. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 358, 435–445.
- Pyles, J. A., Garcia, J. O., Hoffman, D. D., and Grossman, E. D. (2007). Visual perception and neural correlates of novel 'biological motion'. *Vision Res.* 47, 2786–2797.
- Pyles, J. A., and Grossman, E. D. (2009). Neural adaptation for novel objects during dynamic articulation. *Neuropsychologia* 47, 1261–1268.
- Riesenhuber, M., and Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nat. Neurosci.* 2, 1019–1025.
- Riesenhuber, M., and Poggio, T. (2002). Neural mechanisms of object recognition. *Curr. Opin. Neurobiol.* 12, 162–168.

- Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192.
- Saxe, R., Xiao, D. K., Kovacs, G., Perrett, D. I., and Kanwisher, N. (2004). A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia* 42, 1435–1446.
- Saygin, A. P., and Sereno, M. I. (2008). Retinotopy and attention in human occipital, temporal, parietal, and frontal cortex. *Cereb. Cortex* 18, 2158–2168.
- Saygin, A. P., Wilson, S. M., Hagler, D. J., Bates, E., and Sereno, M. I. (2004). Point-light biological motion perception activates human premotor cortex. *J. Neurosci.* 24, 6181–6188.
- Self, M. W., and Zeki, S. (2005). The integration of colour and motion by the human visual brain. *Cereb. Cortex* 15, 1270–1279.
- Seltzer, B., and Pandya, D. N. (1978). Afferent cortical connections and architectonics of the superior temporal sulcus and surrounding cortex in the rhesus monkey. *Brain Res.* 149, 1–24.
- Serences, J. T. (2004). A comparison of methods for characterizing the event-related BOLD timeseries in rapid fMRI. *Neuroimage* 21, 1690–1700.
- Sobotka, S., and Ringo, J. L. (1994). Stimulus specific adaptation in excited but not in inhibited cells in inferotemporal cortex of macaque. *Brain Res.* 646, 95–99.
- Talairach, J., and Tournoux, P. (1988). Co-planar Stereotaxic Atlas of the Human Brain. New York, Thieme Medical.
- Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annu. Rev. Neurosci.* 19, 109–139.
- Taylor, J. C., Wiggett, A. J., and Downing, P. E. (2007). Functional MRI analysis of body and body part representations in the extrastriate and fusiform body areas. *J. Neurophysiol.* 98, 1626–1633.
- Taylor, J. C., Wiggett, A. J., and Downing, P. E. (2009). fMRI-adaptation studies of viewpoint tuning in the extrastriate and fusiform body areas. *J. Neurophysiol.* 98, 1626–1633.
- Thompson, B., Hansen, B. C., Hess, R. E., and Troje, N. F. (2007). Peripheral vision: good for biological motion, bad for signal noise segregation? *J. Vis.* 7, 12: 11–17.
- Thompson, J. C., Clarke, M., Stewart, T., and Puce, A. (2005). Configural processing of biological motion in human superior temporal sulcus. *J. Neurosci.* 25, 9059–9066.
- Thurman, S. M., and Grossman, E. D. (2008). Temporal “Bubbles” reveal key features in point-light biological motion perception. *J. Vis.* 8, 28: 1–11.
- Vaina, L. M., Cowey, A., Eskew, R. T. Jr., LeMay, M., and Kemper, T. (2001a). Regional cerebral correlates of global motion perception: evidence from unilateral cerebral brain damage. *Brain* 124(Pt. 2), 310–321.
- Vaina, L. M., Solomon, J., Chowdhury, S., Sinha, P., and Belliveau, J. W. (2001b). Functional neuroanatomy of biological motion perception in humans. *Proc. Natl. Acad. Sci. U.S.A.* 98, 11656–11661.
- Vangeneugden, J., Pollick, F., and Vogels, R. (2009). Functional differentiation of macaque visual temporal cortical neurons using a parametric action space. *Cereb. Cortex* 19, 593–611.
- Wachsmuth, E., Oram, M. W., and Perrett, D. I. (1994). Recognition of objects and their component parts: responses of single units in the temporal cortex of the macaque. *Cereb. Cortex* 5, 509–522.
- Wandell, B. A., Dumoulin, S. O., and Brewer, A. A. (2007). Visual field maps in human cortex. *Neuron* 56, 366–383.
- Wheaton, K. J., Thompson, J. C., Syngienotis, A., Abbott, D. F., and Puce, A. (2004). Viewing the motion of human body parts activates different regions of premotor, temporal, and parietal cortex. *Neuroimage* 22, 277–288.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 20 August 2009; paper pending published: 19 October 2009; accepted: 04 February 2010; published online: 23 March 2010.

Citation: Grossman ED, Jardine NL and Pyles JA (2010) fMR-adaptation reveals invariant coding of biological motion on the human STS. *Front. Hum. Neurosci.* 4:15. doi: 10.3389/neuro.09.015.2010

Copyright © 2010 Grossman, Jardine and Pyles. This is an open-access article subject to an exclusive license agreement between the authors and the Frontiers Research Foundation, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.