Aalto University

Nummenmaa, Lauri; Smirnov, Dmitry; Lahnakoski, Juha M.; Glerean, Enrico; Jaaskelainen, Iiro P.; Sams, Mikko; Hari, Riitta

Mental Action Simulation Synchronizes Action-Observation Circuits across Individuals

Behavioral/Cognitive

Mental Action Simulation Synchronizes Action–Observation Circuits across Individuals

Lauri Nummenmaa,[1,2,3,4] **Dmitry Smirnov,**[1] **Juha M. Lahnakoski,**[1] **Enrico Glerean,**[1] **Iiro P. Jääskeläinen,**[1] **Mikko Sams,**[1,4] **and Riitta Hari**[2,3]

[1]Department of Biomedical Engineering and Computational Science, [2]Brain Research Unit, O.V. Lounasmaa Laboratory, and [3]Advanced Magnetic Imaging Centre, Aalto NeuroImaging, Aalto University, FI-0076 Espoo, Finland, and [4]Turku PET Centre, FI-20520 Turku, Finland

A frontoparietal action– observation network (AON) has been proposed to support understanding others' actions and goals. We show that the AON "ticks together" in human subjects who are sharing a third person's feelings. During functional magnetic resonance imaging, 20 volunteers watched movies depicting boxing matches passively or while simulating a prespecified boxer's feelings. Instantaneous intersubject phase synchronization (ISPS) was computed to derive multisubject voxelwise similarity of hemodynamic activity and inter-area functional connectivity. During passive viewing, subjects' brain activity was synchronized in sensory projection and posterior temporal cortices. Simulation induced widespread increase of ISPS in the AON (premotor, posterior parietal, and superior temporal cortices), primary and secondary somatosensory cortices, and the dorsal attention circuits (frontal eye fields, intraparietal sulcus). Moreover, interconnectivity of these regions strengthened during simulation. We propose that sharing a third person's feelings synchronizes the observer's own brain mechanisms supporting sensations and motor planning, thereby likely promoting mutual understanding.

Key words: action; emotion; mirroring; simulation; somatosensation; synchronization

Introduction

Perception and action are tightly linked in the brain. While watching an exciting hockey match, we can feel the powerful checks in our bodies, and when the players raise their hands to celebrate a goal, we may have an irresistible urge to follow them. Apparently the observer automatically mimics or "mirrors" some sensorimotor information, thereby tracking and in part replicating the mental and bodily states of others (Gallese and Goldman, 1998; Hari and Kujala, 2009). Support for this common-coding hypothesis comes from neuroimaging studies that have revealed overlapping brain activation for perception and execution of motor actions in a parietofrontal network spanning the inferior parietal and frontal as well as precentral motor and somatosensory cortices (Hari et al., 1998; Buccino et al., 2001; Rizzolatti and Craighero, 2004). Together with the posterior superior temporal sulcus (pSTS) encoding intentionality of an agent's actions (Nummenmaa and Calder, 2009) and providing input to the

mirroring circuits, these regions constitute the action–observation network (AON) subserving action understanding (Caspers et al., 2010; Kilner, 2011). To support social perception, the AON may also interact with the somatosensory cortices, as limbic circuit involved in emotional behavior (Nummenmaa et al., 2008; Keysers et al., 2010).

Automatic remapping of others' bodily and mental states may provide the observers with a sensorimotor framework that facilitates understanding and prediction of others' intentions, and that also allows synchronized and coordinated behavior and thinking across individuals (Hatfield et al., 1994; Rizzolatti and Fabbri-Destro, 2008). Recent evidence suggests that during naturalistic stimulation, individuals' brain activity can be synchronized at time scales of a few seconds. Viewing a movie or listening to a narrative result in time-locked and functionally selective hemodynamic responses in early sensory cortices and in areas involved in higher cognitive functions, suggested to reflect intersubject similarity of information processing (Hasson et al., 2004; Malinen et al., 2007; Wilson et al., 2008). Because simulating others' mental and bodily states likely helps individuals to understand and view the external world in a similar fashion (Hari and Kujala, 2009; Nummenmaa et al., 2012), synchronized neural activity across individuals could be the basic mechanism supporting mutual understanding. Indeed, speaker–listener neural synchronization is associated with successful comprehension of a verbal message (Stephens et al., 2010), and communication by hand gestures (Schippers et al., 2010) enhances synchronization of specific regions between the communicating persons' brains. Similar brain states could thus form a prerequisite for similar mind states.
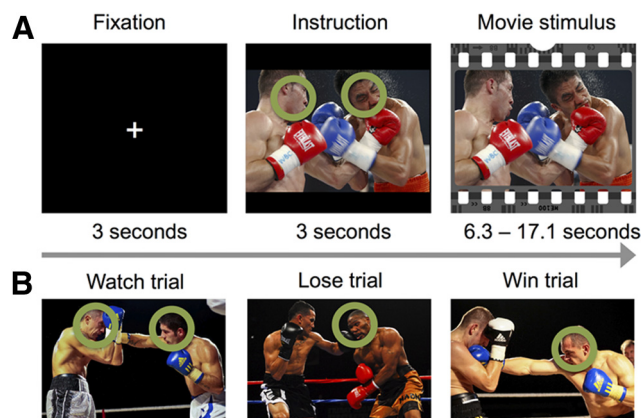
Figure 1. Overview of the trial structure and illustrations of the stimuli.

Here we used our recently developed instantaneous intersubject phase synchronization (ISPS; Glerean et al., 2012) analysis method for testing whether explicitly simulating feelings and actions of other individuals seen in a movie increases intersubject synchrony as well as dynamic interconnectivity of the key nodes of the AON, emotion, and somatosensory circuits. Instead of studying how brain activity becomes synchronized across individuals performing and observing actions, we thus quantified the tendency for brain responses to synchronize across the members of a group who are explicitly simulating similar actions.

Materials and Methods

Twenty right-handed healthy adults (11 females, 9 males; 21–47 years, mean 26 years) with normal or corrected to normal vision volunteered for the study. Individuals with a history of neurological or psychiatric disease or current medication affecting the CNS were excluded. All subjects were compensated for their time and travel costs, and they signed informed consent forms. The Ethics Committee of the Helsinki and Uusimaa Hospital District approved the study protocol, and the study was conducted in accordance with the Declaration of Helsinki.

Stimuli. Figure 1 summarizes the stimuli and experimental design. The stimuli comprised 18, on average 9.81 ± 2.83-s (mean ± SD)-long segments cut from videos of professional boxing matches. The clips depicted typical highlights of the matches, in which one boxer was clearly winning and causing pain to the losing boxer with his punches. Such scenes were chosen because they contain plenty of intentional motor actions that are expected to stimulate the AON (Caspers et al., 2010) and because aggressive scenes are known to evoke particularly reliable activation in the sensorimotor nodes of the AON (Bradley et al., 2003; Nummenmaa et al., 2008). All movies were presented with sound. In a behavioral pilot study, 10 participants (eight males) watched the stimulus videos either passively or trying to explicitly simulate the actions and feelings of one of the boxers—the winner or the loser of the match—thus mimicking the fMRI experimental design described below. We want to emphasize, however, that the movies were very involving so that even during the "passive" conditions when no targets were given, the subjects viewed the boxing matches attentively. After watching each clip, participants provided ratings for subjectively experienced pain, valence (pleasantness–unpleasantness), and arousal. Furthermore, after the simulation trials the participants estimated how much pain the boxer they simulated experienced and how likely he would be to win the match. All responses were provided with a computer keyboard using a scale ranging from 1 to 9.

The results of these evaluations (Table 1) confirmed that participants could unambiguously detect which boxer was going to win or lose the match (intraclass correlation = 0.91). Repeated-measures ANOVAs with Bonferroni corrections for multiple comparisons revealed that the likelihood of the simulated boxer's victory was lower on the lose rather than on the win trials, $F_{(1,9)} = 99.61$, $p < 0.001$, $\eta_p^2 = 0.92$ ($M_{lose} = 2.41$, $M_{win} = 7.09$), and that the losing boxers were evaluated to experience more pain than the winners $F_{(1,9)} = 132.47$, $p < 0.001$, $\eta_p^2 = 0.94$ ($M_{lose} = 6.07$, $M_{win} = 2.06$). Furthermore, simulating the losing versus winning boxer increased the experience of pain ($F_{(1,9)} = 10.52$, $p = 0.03$, $\eta_p^2 = 0.54$) and decreased the experience of pleasure ($F_{(1,9)} = 12.88$, $p = 0.02$, $\eta_p^2 = 0.59$), whereas no differences in experienced pain and pleasure were found between passive viewing versus simulate winner and passive viewing versus simulate loser conditions, F values <4.87, p values > 0.65 (Pain: $M_{watch} = 2.81$, $M_{lose} = 3.36$, $M_{win} = 0.2.04$; Pleasure: $M_{watch} = 5.02$, $M_{lose} = 3.90$, $M_{win} = 4.78$). Arousal scores did not differ across task conditions after correcting for multiple comparisons, although they were numerically larger for the simulation versus passive viewing conditions F values <4.01, p values >0.15 ($M_{watch} = 3.57$, $M_{lose} = 4.59$, $M_{win} = 4.56$)

Experimental design for fMRI. During fMRI, the stimuli were delivered using Presentation software (Neurobehavioral Systems), and they were back-projected on a semitransparent screen using a 3-micromirror data projector (Christie X3, Christie Digital Systems Ltd.) and from there via a mirror to the subject. The viewing distance was 34 cm, and the width of the projected image was 28 cm. The audio track of the movie was played to the subjects with a UNIDES ADU2a audio system (Unides Design) via plastic tubes through porous EAR-tip (Etymotic Research, ER3) earplugs. Sounds were adjusted to be loud enough to be heard over the scanner noise and the loudness was individually fine-tuned to a comfortable level.

Before the experiment, the participants were told that they are going to see a set of boxing videos and that before each video begins, the head of either one or both of the boxers would be tagged with a circle for a few seconds. If both boxers were tagged, the subjects should simply watch the movie similar to how they would be watching TV ("watch" trials), but if only one boxer was tagged, the task would be to explicitly mentally simulate, as accurately as possible, the thoughts, actions, and feelings of the tagged individual. It was stressed that no actual motor actions should be performed during the experiment. On half of these trials, the target boxer would lose the match ("lose" trials) and on half he would win the match ("win" trials). However, the instruction screen did not yet reveal which boxer would actually win the upcoming match.

Each trial began with a 3 s fixation screen. Next, the first frame of the upcoming movie was shown for 3 s with the cue circle(s) that informed the participant how they should behave during the upcoming trial. Finally, the videos were presented. Each stimulus video was shown once with each instruction ("watch," "lose," or "win"), resulting in 18 trials per type and a total of 54 trials, with total task duration of 14 min 23 s. All participants watched the films in a fixed, pseudorandom order; full randomization was not possible as fixed stimulus presentation order was required for intersubject phase synchronization analyses (see below). However, we had two different counterbalanced orders of the instruction screens to control for possible order effects. In these counterbalancings, the "simulate winner" and "simulate loser" instructions for each clip were reversed across counterbalancings. Eye movements were allowed and recorded during the task, because human high-level scene perception strongly depends on saccadic eye movements (Henderson, 2003; Hayhoe and Ballard, 2005). Consequently, restricting eye movements during the study would have biased the participants to perform the explicit simulation and passive tasks in a very unnatural way, as it is known that simulating a specific persons' thoughts influences eye movements during real-world vision (Kaakinen et al., 2011).

fMRI acquisition and analysis. MR imaging was performed with a General Electric Signa 3.0T MRI scanner with Excite upgrade at the Advanced Magnetic Imaging Centre of the Aalto University. Whole-brain data were acquired with T2*-weighted echoplanar imaging (EPI), sensitive to blood oxygen level-dependent (BOLD) signal contrast with the following parameters: 36 axial slices, 4 mm slice thickness, TR = 1800 ms, TE = 30 ms, flip angle = 75°, FOV = 240 mm, voxel size 3 × 3 × 4 mm³, ascending interleaved acquisition with no gaps between slices. A total of 485 volumes were acquired, and the first 4 volumes were discarded to allow for equilibration effects. T1-weighted structural images were acquired at a resolution of 1 × 1 × 1 mm³. The data were preprocessed using FSL software (www.fmrib.ox.ac.uk/fsl/). The EPI images were sinc-interpolated in time to correct for slice time differences and realigned to

Table 1. Self-report data for the stimuli

	Watch		Simulate loser		Simulate winner	
	M	SEM	M	SEM	M	SEM
Likelihood of simulated boxer winning			26.85^a	3.10	78.82^b	2.82
Simulated boxer's pain			67.40^a	4.25	22.83^b	2.42
Subjective experience of pain	31.23^{ab}	5.02	37.34^a	5.95	22.71^b	3.17
Subjective experience of pleasure	55.80^{ab}	3.68	43.33^a	3.27	52.96^b	3.48
Subjective experience of arousal	39.69	5.14	50.98	4.85	50.67	4.89

Mean (M) scores with a different letter (horizontally) are significantly different after Bonferroni correction; means sharing a letter are equivalent.
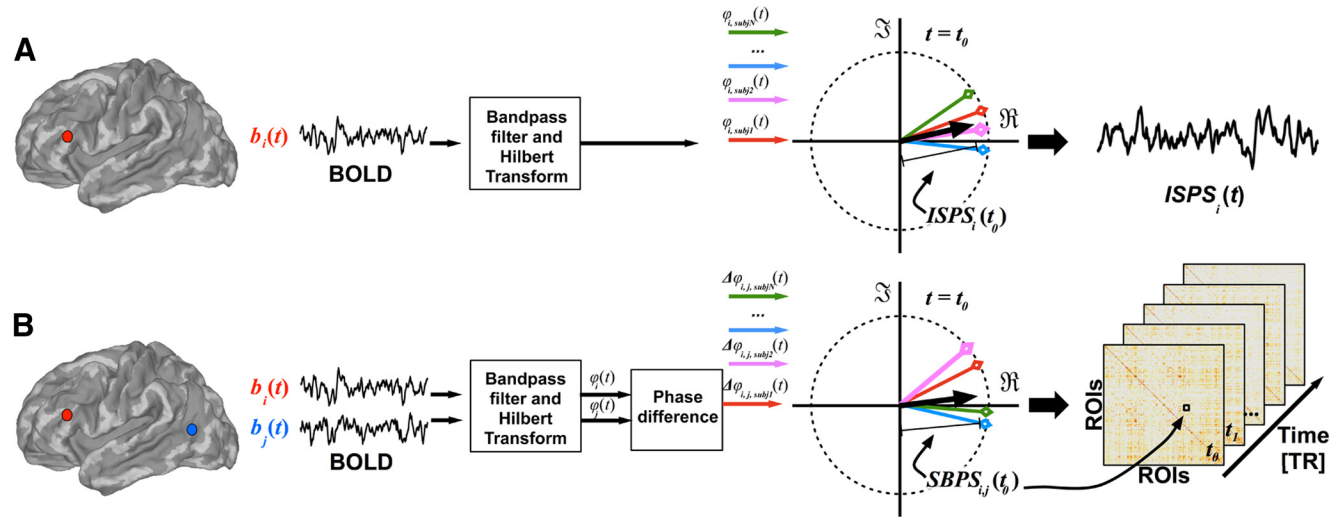

Figure 2. Overview of the instantaneous ISPS (*A*) and instantaneous SBPS (*B*) analysis. *t*, Time, φ, phase; PS, phase synchrony; *i,j*, seed regions; *b*, BOLD signal.

the first scan by rigid body transformations to correct for head movements. EPI and structural images were coregistered and normalized to the T1 standard template in MNI space (Evans et al., 1994) using linear and nonlinear transformations, and smoothed with a Gaussian kernel of FWHM 8 mm.

Intersubject phase synchronization. The data were analyzed using a fMRI Phase Synchronization toolbox introduced recently (Glerean et al., 2012; https://code.google.com/p/funpsy/). ISPS is a measure similar to moment-to-moment intersubject correlation (ISC) computed with a sliding temporal window (Kauppi et al., 2010; Nummenmaa et al., 2012), but it has significantly higher temporal resolution (1 TR of fMRI acquisition, which is also the theoretically maximum resolution). It can thus be used to estimate instantaneous synchronization of brain activity across individuals. Figure 2*A* presents the overall framework. Briefly, we first performed head motion quality control to determine the subjectwise framewise displacement indices (Power et al., 2012). However, given that on average <0.5% of volumes were affected by motion (thus resulting in maximally 0.25% noise to average ISPS values), motion was simply covaried out from the BOLD signal. Next the data were bandpass filtered (0.30 – 0.95 Hz) to remove noise and because the concept of phase synchronization is meaningful only when a narrowband signal is considered. After Hilbert transform, ISPS time series was calculated for each voxel and EPI image in the time series. Given that subject movement may significantly confound the connectivity estimates, movement parameters were regressed out when estimating ISPS and seed-based phase synchronization (SBPS; see below). The voxelwise ISPS time series may be modeled with experimental regressors to estimate the effect of experimental manipulation on regional phase synchrony. As ISPS measures intersubject similarity in phase rather than by computing covariance (cf. ISC), it is temporally more accurate than ISC and also better suited for quantifying intersubject synchronization in blocked designs, where sliding-window ISC would smear signals coming from different blocks. Importantly, phase difference information between voxel pairs can be further used for estimating dynamic functional connectivity (see below).

Time series of the experimental conditions (lose, win, watch) were downsampled to 1 TR and convolved with a gamma function ($\theta = 1, k = 6$) to account for the hemodynamic lag. A simple gamma function rather than the canonical double gamma HRF was used, because ISPS reflects increased similarity rather than the amplitude of hemodynamic activation. Consequently, ISPS signal has only a positive stimulus-driven (or manipulation-driven) deflection without the following undershoot of BOLD signal, and thus a single gamma function will serve as an effective filter (convolution function) for increasing the signal-to-noise ratio in the analysis as well as for compensating for the hemodynamic delay.

The experimental-condition regressors were used to predict voxelwise ISPS time courses in the general linear model (GLM). The resulting correlation coefficients were stored in synchronization maps, where voxel intensities reflected the degree to which ISPS depended on the current experimental condition. The data were modeled separately for the two counterbalanced conditions, and finally the correlation coefficient maps were averaged to obtain an index of mean task-driven ISPS changes in the experiment. To compare the ISPS technique with the prevailing ISC method, we also analyzed the data using the sliding-window ISC described by Nummenmaa et al. (2012).

Dynamic functional connectivity analysis with instantaneous seed-based phase synchronization. To assess whether sharing of the boxer's feelings would modify the connectivity within the AON and emotion circuits, we estimated dynamic functional connectivity of regional time courses using instantaneous SBPS (Glerean et al., 2012). Regions of interest (ROIs) were selected from a recent meta-analysis of the brain basis of action observation (Caspers et al., 2010; see Table 3). To obtain a symmetrical set of ROIs across hemispheres, flipping the x-coordinates was used to mirror unilateral activation clusters. As explicit simulation of a specific action could be expected to have a profound impact on eye movements and visual attention (as when tracking that person's movements), we also derived coordinates for the frontal eye field (FEF) from another meta-analysis (Paus, 1996). To address the role of emotion circuits in the simulation process (Nummenmaa et al., 2008), atlas-based masks for amygdala, insula, thalamus, and anterior cingulate cortex were derived

Table 2. Seed regions and their coordinates for the seed-based phase synchronization analyses

Region	Left-hemispheric coordinates			Right-hemispheric coordinates		
	X	Y	Z	X	Y	Z
Amygdala	−22	−4	−18	24	−4	−18
Insula	−38	2	−2	38	4	−2
Thalamus	−10	−18	6	12	−18	6
Anterior cingulate cortex	−2	20	24	4	20	24
Caudate nucleus	−12	10	10	14	10	10
Nucleus accumbens	−10	12	−8	10	12	−6
Inferior frontal gyrus	56	30	−2	56	30	−2
Primary somatosensory cortex (SI)	−60	−20	40	60	−20	40
Superior parietal lobule	−22	−62	64	22	−62	64
Intraparietal sulcus	−30	−54	48	30	−54	48
Medial premotor cortex (SMA)	−2	18	50	4	12	58
Inferior frontal gyrus	−50	9	30	52	12	26
Dorsal premotor cortex	−26	−4	56	34	−2	54
Inferior parietal lobule	−60	−24	36	44	−34	44
Superior temporal sulcus/Middle temporal gyrus	−54	−50	8	56	−40	4
Lateral occipital complex	−46	−72	2	52	−64	0
Fusiform gyrus	−44	−56	−18	44	−54	−18
Frontal eye field	−33	4	47	35	4	47

Table 3. Seed regions and their coordinates for seed-voxel correlation analysis, and correlations, *p*-values, and percentages of significantly correlated voxels within each network when testing for the effect of explicit simulation (simulate loser and simulate winner pooled) versus passive observation for spatially averaged ISPS time series in the six intrinsic networks at *p* < 0.05 (FDR corrected)

Network	Seed region	x	y	z	r	p	% voxels
Auditory	Superior temporal gyrus	56	−22	0	0.34	< 0.05	51
Default mode	Posterior cingulate cortex	0	−50	30	0.01	NS	0
Dorsal attention	Intraparietal sulcus	32	−46	56	0.42	< 0.05	77
Executive	Superior frontal gyrus	0	44	28	0.02	NS	0
Sensorimotor	Postcentral gyrus	40	0	48	0.36	< 0.05	53
Visual	Calcarine sulcus	−10	−96	0	0.38	< 0.05	60

NS, Not significant.

from the AAL (Automated Anatomical Labeling) atlas (Tzourio-Mazoyer et al., 2002). Table 2 lists the ROIs and their coordinates.

Spheres of 6 mm diameter were drawn around these coordinates, and instantaneous SBPS was used as a group-level time-varying connectivity measure between each pair of regions, resulting in 481 (number of EPI volumes), 36 × 36 connectivity matrices (between 36 regions, i.e., a time-varying functional network of 36 nodes; Fig. 2B). Finally, the gamma-convolved experimental-condition regressors were used to predict each connection's time series in the GLM to assess the effects of simulation on AON connectivity. The resulting connectivity matrices thus revealed connections that were strengthened when participants shared the boxer's feelings versus when they watched the movies without any predetermined target.

To test the statistical significance of the ISPS maps and SBPS connections, we performed a fully nonparametric voxelwise permutation test for the r statistic (Wilson et al., 2008; Kauppi et al., 2010). Because calculation of all possible time-shift combinations would be computationally prohibitive, we approximated the full permutation distribution with A = 1,000,000 realizations. Sampling was randomized over every brain voxel and shifting point without any restrictions. A nonparametric test was used because, this way, one does not have to make assumptions regarding the null distribution. We corrected the resulting p-values using false discovery rate (FDR) multiple-comparisons correction with independence (or positive dependence) assumption.

Seed-voxel correlation analysis. To assess whether explicit simulation would enhance intersubject synchronization of large-scale intrinsic networks rather than of distinct brain regions, we used seed-voxel correlation analysis for delineating the six intrinsic brain networks (Raichle, 2010) to be used as ROIs in ISPS analysis. The seed-voxel correlation technique enables characterization of task-independent patterns of functional connectivity as well as mapping of the functional organization of large-scale brain networks (Fox et al., 2006; Margulies et al., 2007). Seeds

were anatomical foci routinely used in seed-voxel correlation analysis (visual, sensorimotor, auditory, default-mode, dorsal attention, and executive control networks; Table 3). Spherical ROIs with 6 mm radius were generated around these coordinates, and mean time-series were extracted for each ROI and participant. The data were filtered through 0.01–0.08 Hz. Means of each participant's regionwise time series were subsequently used to identify individual correlation maps for each of the six networks by correlating seed region time series with time series of all other voxels in brain. Subject motion, as well as mean signal in whole volume, white matter, and ventricles, was regressed out. The resulting subjectwise network maps were finally Fisher-transformed and averaged across subjects. A fully nonparametric voxelwise permutation test was applied to determine the final population-level statistical threshold (p < 0.01, FDR corrected) for the maps.

To analyze whether the regional synchronization in each network would be modulated by the simulation task, a spatially averaged ISPS time series was extracted for each statistically thresholded network ROI. Finally, these time series were correlated with the experimental time series of explicit simulation versus passive viewing.

Task-evoked BOLD responses. As a complementary approach, we also analyzed regional responses during simulating winning and simulating losing, as well as during passive viewing. A random-effects model was implemented using a two-stage process (first and second level). For each participant, we used the classical GLM with boxcar regressors to assess regional effects of task parameters on BOLD indices of activation. The model included three experimental conditions (watch, lose, win) and effects of no interest (realignment parameters) to account for motion-related variance. Low-frequency signal drift was removed using a high-pass filter (cutoff 256 s), and AR(1) modeling of temporal autocorrelations was applied. Individual contrast images were generated for the following contrasts: simulate winning > watch, simulate losing > watch, simulate > watch, simulate losing > simulate winning, and simulate win-
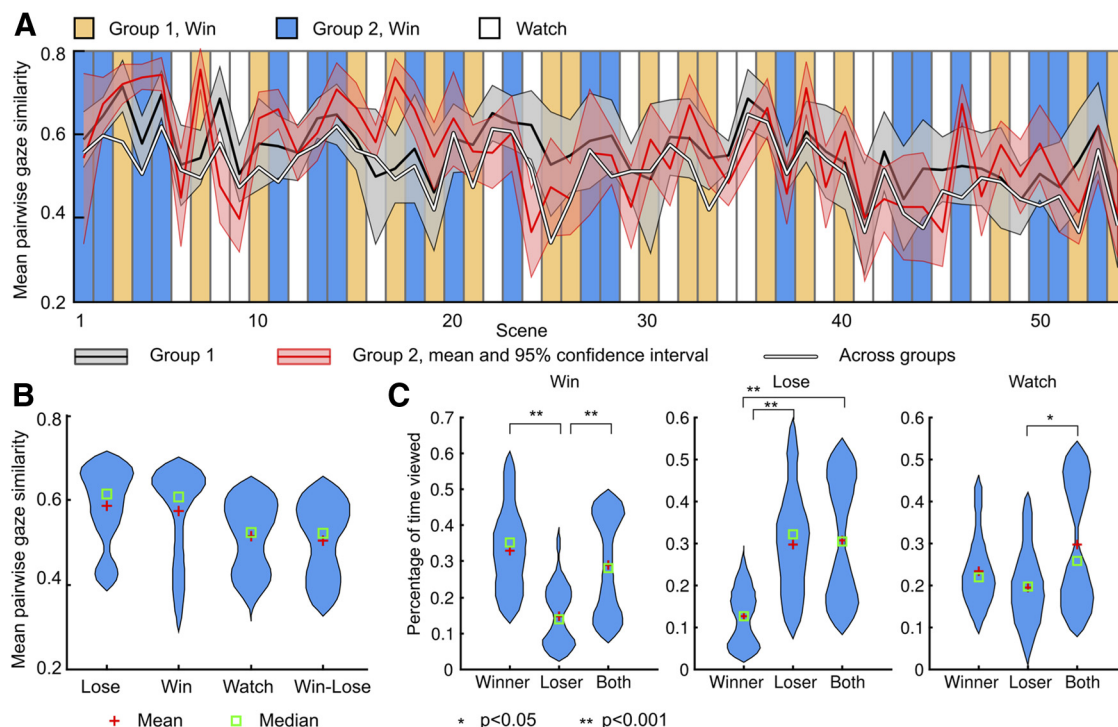
Figure 3. Eye-movement results. **A**, Mean ± 95% confidence intervals for momentary e-ISC for Groups 1 and 2 and across groups averaged in time-domain for each trial scene. Vertical panels denote the different experimental conditions. Time series are plotted separately for the two counterbalanced conditions; thus, win trials for Group 1 were lose trials for Group 2 and vice versa. **B**, Distribution of e-ISC scores across the experimental conditions (lose, win, and watch). **C**, Distributions of the percentage of time the participant fixated on the loser, winner, or both boxers in the different experimental conditions.

ning > simulate losing. The second-level analysis used these contrast images in a new GLM and generated statistical images, that is, SPM-t maps. With balanced designs at first level (i.e., similar events for each subject, in similar numbers), this second-level analysis closely approximates a true mixed effects design, with both within- and between-subject variance. Statistical threshold was set at $T > 3.0$ and $p < 0.05$, FDR corrected at cluster level.

Finally, ROI analysis was conducted for the BOLD-GLM and ISPS data to test which somatosensory regions showed (1) increased signal amplitude and (2) increased intersubject synchronization during explicit simulation of observed actions. To that end, probabilistic cytoarchitectonic regions of interest were generated for the primary somatosensory cortex (Brodmann areas 1, 2, 3a, and 3b) and for the opercular areas (OP1–OP4) in the region of the second somatosensory cortex using the SPM Anatomy toolbox (Eickhoff et al., 2005). Mean BOLD contrast estimates for simulation versus passive viewing, and mean association (Pearson's r) between the simulation versus passive viewing conditions time series and ISPS time series, were subsequently extracted for each ROI.

Eye movement recordings and questionnaire analyses. Eye movements were successfully recorded from 10 participants during the fMRI experiment with SMI Eye Track long-range eye tracking system (Sensomotoric Instruments GmbH), based on video-oculography and the dark pupil-corneal reflection method. Originally, eye movement recordings were attempted for 15 participants, but the camera could not be set up at all for two of those (due to large abdominal circumference) and could not be calibrated for an additional three. The sampling rate of the eye tracker was 60 Hz. A five-point calibration and validation was completed before the experiment. Fixation data were drift corrected using the mean gaze position during the fixation period.

Eye movement data acquired during movie presentation were analyzed with a Matlab toolbox developed for the purposes of this study. The degree of spatiotemporal intersubject synchronization of eye movements (e-ISC) was measured by computing subjectwise heatmaps for each trial. For each participant, each fixation was modeled as a Gaussian function with mean of fixation's Cartesian coordinate and SD of 1°—based on the

assumption that the foveal field of view is ~2.5°—and multiplied with fixation duration in the tens of milliseconds. A mean intersubject correlation index was computed for a sliding window (length 1 s, step size 100 ms), and average intersubject similarity scores were computed for the different experimental conditions (lose, win, watch). This synchronization time series was also used to predict cerebral ISPS to test whether synchronization of visual attention across subjects would be associated with enhanced intersubject synchrony of brain activation.

In a complementary methodological approach, we manually annotated the locations of the boxers separately by drawing a polygonal ROI around each person. Whenever one boxer was occluded by another, the eye fixations were considered to land on both boxers. We then updated the ROI each time the boxer was no longer within the region. Finally, we calculated the percentage of time the subjects were fixating within each region of interest or the overlapping region, averaged the viewing times across subjects, and used a two-sample t test assuming unequal variances to test whether the viewing times were significantly different.

Results

Eye movements
During movie viewing, eye movements were strongly synchronized across participants ($M_{e-ISC} = 0.53$, $p < 0.05$; Fig. 3A), and the synchronization was significantly stronger for participants engaged in same versus different explicit simulation conditions (e-ISC = 0.58 vs 0.50; $p < 0.005$, permutation test), as well as in explicit simulation versus passive condition ($p < 0.01$). These results confirmed that the participants were following the task instructions (Fig. 3B). ROI-based analysis also confirmed these findings by showing that participants spent more time looking at the task-relevant versus task-irrelevant boxer in the simulate-loser and simulate-winner conditions (p values < 0.05 in paired t test; Fig. 3C). However, synchronization of eye movements was
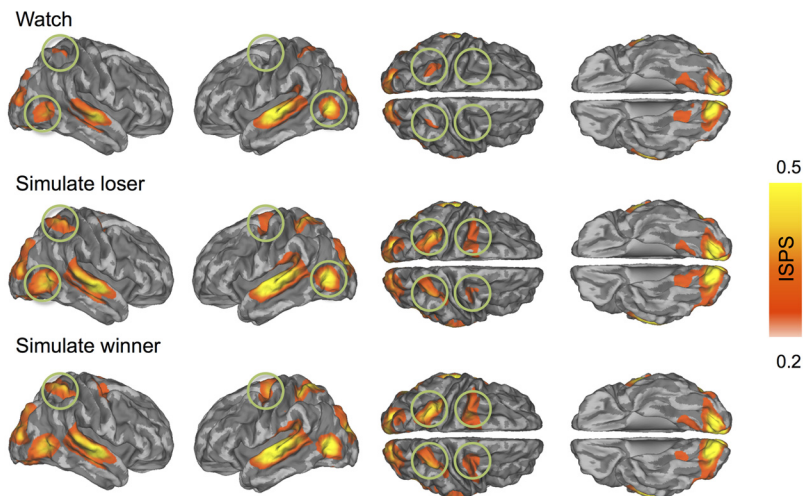
Figure 4. Brain regions showing statistically significant ISPS during passive viewing (top), simulation of a losing (middle), and simulation of a winning (bottom) boxer. The maps are thresholded at $p < 0.05$ (FDR corrected). Circles point to regions where ISPS was larger during explicit simulation versus passive observation conditions. Color bar denotes the ISPS statistic range.
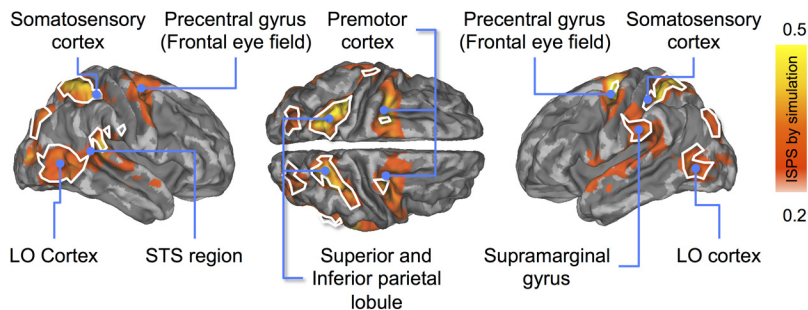

Figure 5. Brain regions showing statistically significantly stronger ISPS during simulation (averaged across simulate winner and simulate loser) than passive viewing (red to yellow). White borders show regions with overlapping ISPS and ISC. The data are thresholded at $p < 0.05$ (FDR corrected).

not associated with intersubject synchronization of brain activity in any region.

Instantaneous intersubject phase synchronization

Participants' brain activity was time-locked in several brain regions, but the degree of synchronization was contingent on how the participants were viewing the movie clips. During passive watching, the ISPS differed significantly from zero (i.e., ISPS > 0.18 at $p < 0.05$ FDR corrected; Fig. 4, top row) in the superior temporal cortex (auditory areas), early visual cortices and fusiform and lingual gyri, the lateral occipital cortex (MT/V5 region), and the parietal cortex (superior and inferior parietal lobule; intraparietal sulcus). Simulation of either the winner or the loser elicited much more widespread synchronization of cortical and subcortical brain regions: strong ISPS occurred now also in sensorimotor cortices, superior temporal gyrus, intraparietal sulcus, and frontal eye fields (Fig. 4, middle and bottom rows).

Figure 5 compares statistically the simulation (loser- and winner-simulation data pooled) versus passive watching. Simulation increased ISPS in sensory [visual, auditory, and primary (BA 2) and secondary (OP2) somatosensory cortices], but also in nodes of the AON. Specifically, this effect was observed in the precentral gyrus (premotor cortex) and parietal cortex (superior and inferior parietal lobules), superior and middle temporal cortices, as well as in regions involved in voluntary attentional control (intraparietal sulci and frontal eye fields). Synchronization

was weaker during simulation only in bilateral calcarine and lingual gyri and in the middle cingulate gyrus. Conventional sliding-window ISC analyses yielded similar parietal and posterior temporal effects but with lesser power than ISPS because the sliding window had to be longer than the trial duration; critically, they completely missed the frontal and anterior-temporal-lobe effects. However, neither ISPS nor ISC displayed differences between the simulate-winner and simulate-loser conditions.

Synchronization in intrinsic networks defined by seed-voxel correlation analysis

The ISPS data in Figure 5 and the connectivity matrix in Figure 6 are suggestive of condition-related ISPS differences in the well known dorsal attention, sensorimotor, auditory, and visual networks. This interpretation was supported by correlating the mean ISPS time courses of all voxels in each of the six intrinsic networks defined by seed-voxel correlation analysis (auditory, default-mode, dorsal attention, executive control, sensorimotor, and visual networks), with the regressor representing the experimental conditions (i.e., simulation vs passive viewing; Table 3). This result confirmed network-specific effects of simulation: compared with passive viewing, simulation resulted in stronger mean ISPS in auditory, dorsal attention, sensorimotor, and visual networks (r values >0.34, p values <0.05), but not in default-mode or in executive control networks (r values <0, p values >0.05).

Seed-based phase synchronization connectivity

The SBPS analysis (Fig. 6) revealed increased large-scale dynamic connectivity of the left-hemispheric components of the AON— particularly of dorsal premotor cortex (dPMC), inferior frontal gyrus (IFG), superior parietal lobule (SPL), and somatosensory cortex (SSC; SI)— during explicit simulation compared with passive viewing. From these areas, dPMC showed increased functional connectivity with caudate, nucleus accumbens (nACC), intraparietal sulcus (IPS), and supplementary motor area (SMA). For SI cortex, connectivity changes occurred with SPL, IPS, SMA, and IFG. For IFG, connectivity changes were mainly found for SPL, IPS, lateral occipital cortex, and fusiform gyrus (FG). This pattern was markedly absent in the right hemisphere. These analyses also revealed that the limbic emotions circuitry interacted with the left-hemispheric AON: during simulation, right caudate showed enhanced connectivity between IFG and dPMC, whereas right amygdala's connectivity was increased with left-hemispheric components of the AON (IFG, SPL, and dPMC). Finally, connectivity between right insula and left SI cortex was also enhanced during simulation. No connectivity changes were observed when the two simulation conditions were contrasted directly with each other.

Task-evoked BOLD responses

Neither ISPS nor SBPS connectivity differed between simulating of losing versus winning boxers. However, conventional GLM analysis of the BOLD responses revealed stronger responses during simulating losing versus winning boxers in the right anterior insula, Rolandic somatosensory cortex, temporoparietal junction, and bilateral fusiform gyri (Fig. 7), whereas the opposite contrast revealed no significant clusters. Finally, when compared against the passive viewing condition, both explicit simulation conditions (simulate loser and simulate winner) resulted in widespread activation in the emotion-related circuits (including thalamus, anterior insula, amygdala) in superior and posterior parietal cortices as well as in bilateral precentral gyri.

Next we used ROI analysis in the somatosensory cortical circuitry comprising the SI and SII cortices to compare whether regionally enhanced between-subjects reliability (ISPS) of brain responses during explicit simulation would also be associated with enhanced response amplitudes (i.e., β values from BOLD-GLM). This analysis (Fig. 8) revealed regionally specific dissociation between response similarity and amplitude in the SI and SII cortices. Whereas simulation induced significant ISPS mainly in SI (area 2) and OP1, BOLD-GLM responses were significant bilaterally in all the tested regions except OP2–OP3 and right-hemispheric OP4.

Discussion

We show, for the first time, that when a group of individuals engages in explicit simulation of another person's actions and feelings, the degree of their intersubject phase synchronization of hemodynamic brain activity increases in selected brain areas. During passive viewing, brain activity was significantly time-locked across participants only in the lateral occipital cortex (MT/V5), visual (lingual and calcarine gyri), auditory (STG), and superior parietal (SPL) cortex, whereas during mental simulation of a boxer's bodily and mental feelings, the intersubject synchronization also extended to the nodes of the AON proposed to encode action intentionality (pSTS), motor planning (premotor cortices), and elaboration of motor plans before action execution (posterior parietal cortices; Kilner, 2011). Additionally, enhanced synchronization was observed in the primary (SI) and secondary (SII) somatosensory cortices and in the IPS, precentral gyrus/sulcus (FEF), and STG regions proposed to serve as nodes of the dorsal and ventral frontoparietal attention circuits (Corbetta and Shulman, 2002). Altogether, these findings suggest that mentally simulating other individuals' feelings and actions involves "resonance" between the observers, both in their sensorimotor and attentional systems.

Mentally simulating others extends intersubject synchronization beyond sensory cortices

When humans view complex dynamic scenes, such as movies, their sensory projection cortices and attention-controlling networks become temporally synchronized (Hasson et al., 2004;
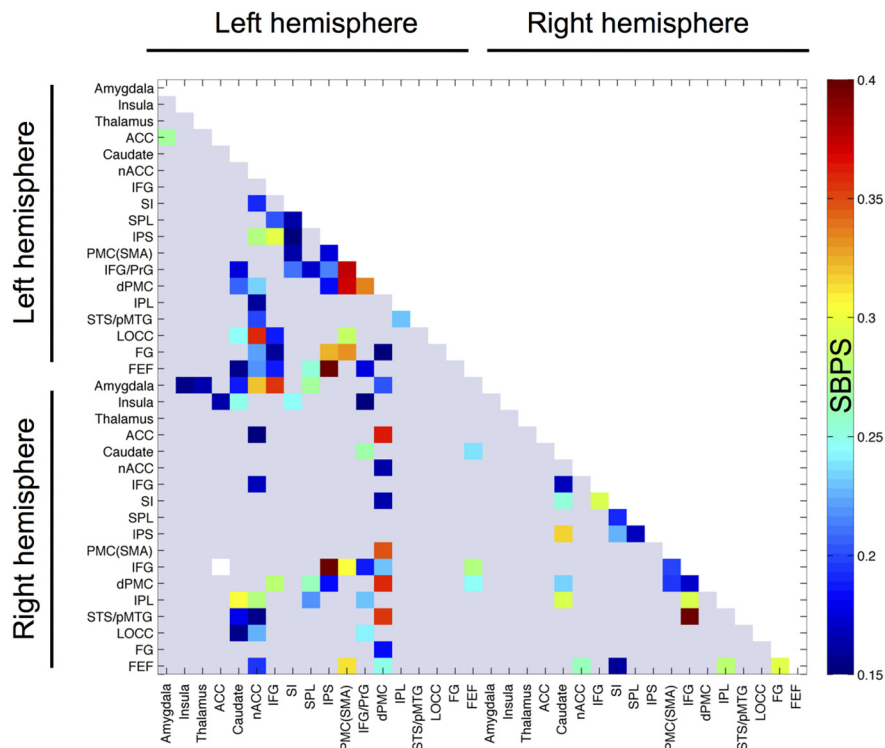

Figure 6. Connectivity matrix for brain regions showing stronger instantaneous SBPS during simulation than passive viewing. The data are thresholded at $p < 0.05$ (FDR corrected). ACC, Anterior cingulate cortex.

Malinen et al., 2007). We extend these findings by showing that such synchronization of brain activity is stronger and more widespread when observers are explicitly simulating others' mental and bodily states. This finding accords with prior studies showing that similarity in attentional processing amplifies the degree of ISC: unlike feature films where cutting and camera runs guide observers' attention exogenously, unedited "segments of reality" such as surveillance camera shots with fixed point of view do not trigger large-scale cortical synchronization across individuals (Hasson et al., 2010). In a similar vein, an edited movie—compared with a video displaying a talking face with a relatively interesting narrative—synchronizes strongly the dorsal attention network across subjects (Malinen and Hari, 2011).

Our data reveal that even when movies triggered relatively weak ISPS during passive viewing, ISPS for exactly the same movies was significantly amplified when participants were explicitly simulating the actions and feelings of a specific person shown in the movie. This postulation accords with studies showing that speaker–listener neural synchronization is associated with successful comprehension of verbal (Stephens et al., 2010) and nonverbal (Schippers et al., 2010) messages, thus highlighting the role of time-locked brain activity across individuals as the basic mechanism supporting interpersonal understanding of the world. Well known nodes of the dorsal attention system involved in attention and eye movement control (Corbetta et al., 2008) showed increased ISPS during explicit simulation, in accord with higher e-ISC in the explicit versus passive simulation conditions. However, because e-ISC did not predict cerebral ISPS, it is likely that the degree of similarity in participant's brain activity related also to perception and simulation of the external world, rather than mere visual sampling of the environment with eye fixations (Hasson et al., 2004; Nummenmaa et al., 2012).
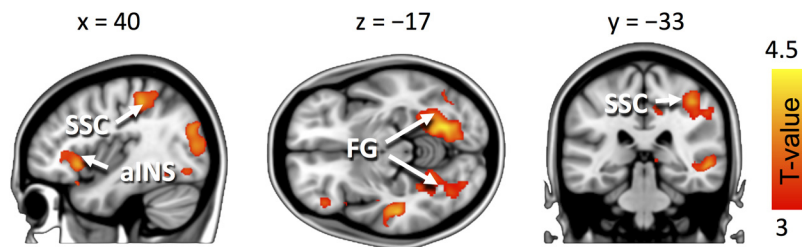
Figure 7. Brain regions showing stronger BOLD responses during simulation of a losing than a winning boxer. The data are thresholded at $p < 0.05$ (FDR corrected). Color bar denotes the T-statistic range. aINS, Anterior insula.
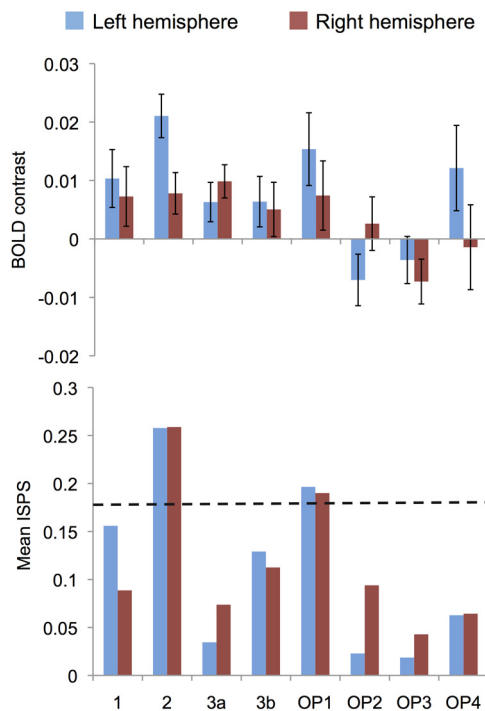

Figure 8. Mean ($\pm$ SE) BOLD contrast (top) and ISPS (bottom) estimates for the contrasts explicit simulation versus passive viewing in left- and right-hemispheric ROIs in the SI and SII somatosensory cortex. Dashed horizontal line indicates the critical ISPS value for $p < 0.05$ obtained in permutation testing. 1, 2, 3b, Cytoarchitectonic areas 1, 2, and 3b of the SI cortex; OP1–OP4, parietal operculum (SII) cytoarchitectonic areas 1–4 according to the Anatomy Toolbox by Eickhoff et al. (2005).

Synchronization as the general supporting mechanism for mutual action understanding

Numerous studies have revealed overlapping neural activation for observing and executing motor actions (for review, see Rizzolatti et al., 2001; Rizzolatti and Fabbri-Destro, 2008; Hari and Kujala, 2009), but also for perception and experience of emotional states, such as pain (Singer et al., 2004; Saarela et al., 2007), disgust (Wicker et al., 2003), and pleasure (Jabbi et al., 2007), suggesting that humans use in part the same neural mechanisms for both producing their own and understanding others' action goals and internal states. Our findings show that brain activity fluctuates in a temporally synchronized manner across a group of individuals who simulate others' actions and feelings. At a behavioral level, such "keeping in sync" with others may enable initiating rapid motor responses to others' actions in the dynamic, constantly changing environment. This view is supported by results from a range of experiments that have confirmed that intersubject synchronization occurs at the level of behavior (Dimberg and Thunberg, 1998; Chartrand and Bargh, 1999), as well as of

psychophysiological (Marci et al., 2007; Konvalinka et al., 2011) and mental states (Hatfield et al., 1994), and that the similarity of action time courses across individuals supports social interaction (Lakin and Chartrand, 2003; Marci et al., 2007). The current results further suggest that large-scale functional networks could support action understanding across individuals during highly naturalistic, dynamic contexts.

ISPS did not differ between the simulating-losing versus simulating-winning conditions at our a priori threshold in SI or SII cortices, or elsewhere. However, conventional GLM analysis confirmed that simulating losing did indeed activate brain areas supporting affective and sensory components of pain (Davis, 2000; Singer et al., 2004; Saarela et al., 2007): simulating the boxer losing the match (and thus experiencing pain) enhanced activity in the right anterior insula and in SI and SII cortices. Additional activations were observed in the face- and body-sensitive fusiform cortices, consistent with the notion that observing pain in others engages the ventral face and body processing systems (Singer et al., 2004). We thus propose that engaging in explicit simulation involves increased large-scale interindividual synchronization of the AON. In contrast, the GLM results of BOLD signals suggest that information specific to others' mental and bodily states (such as winning and losing) is reflected in response amplitudes of this circuitry.

It must nevertheless be stressed that the findings are strictly related to synchronous activation in a group of observers explicitly simulating the same actions, rather than to cortical synchrony between an actor and an observer (cf. Schippers et al., 2010). This difference may explain why the synchronization effects were to some extent unspecific, extending also to the frontoparietal attention systems. Future studies need to establish how actors' and observers' brain activity becomes synchronized during action performance and observation.

Somatosensory mechanisms of mutual understanding

Compared with passive observation, explicit simulation enhanced ISPS as well as response amplitudes in both SI (area 2) and SII (OPII) cortices. Recent work indicates that both these areas can be recruited vicariously, just by viewing another person to be touched (Keysers et al., 2010). Our current results further indicate that both SI and SII cortices are critical nodes in transforming observed third-person actions to the first-person experience in a task-dependent manner, even so that the SI and SII cortices become temporally synchronized across individuals who simulate others' actions. This time-locking suggests that the SI and SII cortices can subserve mutual action understanding via temporally shared neural response profiles across individuals.

Our results on dynamic connectivity demonstrated that the coupling between the key regions of the AON (IPC, PMC, and SI) was enhanced during explicit simulation. With further input from the limbic emotion systems (amygdala, caudate), these regions may operate together as a dynamic network for supporting understanding of other persons' actions and feelings. Future studies using dynamic connectivity measures encompassing the whole brain (Glerean et al., 2012) could elucidate how these networks couple with, for example, cortical midline systems involved in self-referential information processing (Northoff and Bermpohl, 2004) during action observation.

In prior fMRI studies using conventional analysis of BOLD response amplitudes, the action observation system has often been bilateral (Aziz-Zadeh et al., 2006; Caspers et al., 2010). The current SBPS analysis of dynamic brain connectivity in right-handed subjects revealed strong left-hemispheric lateralization in the dynamic connectivity of the AON extending from the somatosensory and motion-sensitive visual cortices to posterior parietal and superior temporal cortices. Moreover, right-hemispheric limbic regions (amygdala, insula) also increased connectivity with the aforementioned left-hemispheric circuit during explicit simulation. Thus, although the AON is activated bilaterally during action simulation, the activation of the individual nodes covaries more strongly within the left hemisphere, as well as across hemispheres. Consequently, coordinated activity of, particularly, the left-hemisphere network may ultimately support mutual understanding between people exposed to the same external world.

Conclusions

To navigate successfully in the complex and continuously changing social world, it is essential to understand what others are doing, and aiming at why they are behaving in the way that they do and how they are feeling in a particular social situation. We conclude that sharing other persons' feelings during naturalistic action observation triggers synchronized brain activity across the observers' attentional and action–observation networks, as well as in the cortical somatosensory circuitry. Such consistent mapping of others' mental and bodily states into the observers' sensorimotor system may ultimately support social interaction: actions resonating and replicated in the observers' brains provide rapid means for understanding others' action goals, and support mutual understanding of other agents' actions and feelings.

References

Aziz-Zadeh L, Koski L, Zaidel E, Mazziotta J, Iacoboni M (2006) Lateralization of the human mirror neuron system. J Neurosci 26:2964–2970. CrossRef Medline

Bradley MM, Sabatinelli D, Lang PJ, Fitzsimmons JR, King W, Desai P (2003) Activation of the visual cortex in motivated attention. Behav Neurosci 117:369–380. CrossRef Medline

Buccino G, Binkofski F, Fink GR, Fadiga L, Fogassi L, Gallese V, Seitz RJ, Zilles K, Rizzolatti G, Freund HJ (2001) Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. Eur J Neurosci 13:400–404. CrossRef Medline

Caspers S, Zilles K, Laird AR, Eickhoff SB (2010) ALE meta-analysis of action observation and imitation in the human brain. Neuroimage 50:1148–1167. CrossRef Medline

Chartrand TL, Bargh JA (1999) The chameleon effect: the perception-behavior link and social interaction. J Pers Soc Psychol 76:893–910. CrossRef Medline

Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci 3:201–215. Medline

Corbetta M, Patel G, Shulman GL (2008) The reorienting system of the human brain: from environment to theory of mind. Neuron 58:306–324. CrossRef Medline

Davis KD (2000) The neural circuitry of pain as explored with functional MRI. Neurol Res 22:313–317. Medline

Dimberg U, Thunberg M (1998) Rapid facial reactions to emotional facial expressions. Scand J Psychol 39:39–45. CrossRef Medline

Eickhoff S, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K (2005) A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage 25:1325–1335. CrossRef Medline

Evans AC, Collins DL, Mills SR, Brown ED, Kelly RL, Peters TM (1994) 3D statistical neuroanatomical models from 305 MRI volumes. In: IEEE Nuclear Science Symposium and Medical Imaging Conference 1993, San Francisco, Oct 31-Nov 6, 1993, pp 1813–1817. CrossRef

Fox MD, Corbetta M, Snyder AZ, Vincent JL, Raichle ME (2006) Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. Proc Natl Acad Sci U S A 103:10046–10051. CrossRef Medline

Gallese V, Goldman A (1998) Mirror neurons and the simulation theory of mind-reading. Trends Cogn Sci 2:493–501. CrossRef Medline

Glerean E, Salmi J, Lahnakoski JM, Jääskeläinen IP, Sams M (2012) Functional magnetic resonance imaging phase synchronization as a measure of dynamic functional connectivity. Brain Connectivity 2:91–101. CrossRef Medline

Hari R, Kujala MV (2009) Brain basis of human social interaction: from concepts to brain imaging. Physiol Rev 89:453–479. CrossRef Medline

Hari R, Forss N, Avikainen S, Kirveskari E, Salenius S, Rizzolatti G (1998) Activation of human primary motor cortex during action observation: a neuromagnetic study. Proc Natl Acad Sci U S A 95:15061–15065. CrossRef Medline

Hasson U, Nir Y, Levy I, Fuhrmann G, Malach R (2004) Intersubject synchronization of cortical activity during natural vision. Science 303:1634–1640. CrossRef Medline

Hasson U, Malach R, Heeger DJ (2010) Reliability of cortical activity during natural stimulation. Trends Cogn Sci 14:40–48. CrossRef Medline

Hatfield E, Cacioppo J, Rapson RL (1994) Emotional contagion. New York: Cambridge UP.

Hayhoe M, Ballard D (2005) Eye movements in natural behavior. Trends Cogn Sci 9:188–194. CrossRef Medline

Henderson JM (2003) Human gaze control during real-world scene perception. Trends Cogn Sci 7:498–504. CrossRef Medline

Jabbi M, Swart M, Keysers C (2007) Empathy for positive and negative emotions in the gustatory cortex. Neuroimage 34:1744–1753. CrossRef Medline

Kaakinen JK, Hyönä J, Viljanen M (2011) Influence of a psychological perspective on scene viewing and memory for scenes. Q J Exp Psychol 64:1372–1387. CrossRef

Kauppi JP, Jääskeläinen IP, Sams M, Tohka J (2010) Inter-subject correlation of brain hemodynamic responses during watching a movie: localization in space and frequency. Front Neuroinform 4:12.

Keysers C, Kaas JH, Gazzola V (2010) Somatosensation in social perception. Nat Rev Neurosci 11:417–428. CrossRef Medline

Kilner JM (2011) More than one pathway to action understanding. Trends Cogn Sci 15:352–357. CrossRef Medline

Konvalinka I, Xygalatas D, Bulbulia J, Schjødt U, Jegindø EM, Wallot S, Van Orden G, Roepstorff A (2011) Synchronized arousal between performers and related spectators in a fire-walking ritual. Proc Natl Acad Sci U S A 108:8514–8519. CrossRef Medline

Lakin JL, Chartrand TL (2003) Using nonconscious behavioral mimicry to create affiliation and rapport. Psychol Sci 14:334–339. CrossRef Medline

Malinen S, Hari R (2011) Data-based functional template for sorting independent components of fMRI activity. Neurosci Res 71:369–376. CrossRef Medline

Malinen S, Hlushchuk Y, Hari R (2007) Towards natural stimulation in MRI—issues of data analysis. Neuroimage 35:131–139. CrossRef Medline

Marci CD, Ham J, Moran E, Orr SP (2007) Physiologic correlates of perceived therapist empathy and social-emotional process during psychotherapy. J Nerv Ment Dis 195:103–111. CrossRef Medline

Margulies DS, Kelly AMC, Uddin LQ, Biswal BB, Castellanos FX, Milham MP (2007) Mapping the functional connectivity of anterior cingulate cortex. Neuroimage 37:579–588. CrossRef Medline

Northoff G, Bermpohl F (2004) Cortical midline structures and the self. Trends Cogn Sci 8:102–107. CrossRef Medline

Nummenmaa L, Calder AJ (2009) Neural mechanisms of social attention. Trends Cogn Sci 13:135–143. CrossRef Medline

Nummenmaa L, Hirvonen J, Parkkola R, Hietanen JK (2008) Is emotional contagion special? An fMRI study on neural systems for affective and cognitive empathy. Neuroimage 43:571–580. CrossRef Medline

Nummenmaa L, Glerean E, Viinikainen M, Jääskeläinen IP, Hari R, Sams M (2012) Emotions promote social interaction by synchronizing brain activity across individuals. Proc Natl Acad Sci U S A 109:9599–9604. CrossRef Medline

Paus T (1996) Location and function of the human frontal eye-field: a selective review. Neuropsychologia 34:475–483. CrossRef Medline

Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE (2012) Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage 59:2142–2154. CrossRef Medline

Raichle ME (2010) Two views of brain function. Trends Cogn Sci 14:180–190. CrossRef Medline

Rizzolatti G, Craighero L (2004) The mirror-neuron system. Annu Rev Neurosci 27:169–192. CrossRef Medline

Rizzolatti G, Fabbri-Destro M (2008) The mirror system and its role in social cognition. Curr Opin Neurobiol 18:179–184. CrossRef Medline

Rizzolatti G, Fogassi L, Gallese V (2001) Neurophysiological mechanisms underlying the understanding and imitation of action. Nat Rev Neurosci 2:661–670. CrossRef Medline

Saarela MV, Hlushchuk Y, Williams ACD, Schürmann M, Kalso E, Hari R (2007) The compassionate brain: humans detect intensity of pain from another's face. Cereb Cortex 17:230–237. Medline

Schippers MB, Roebroeck A, Renken R, Nanetti L, Keysers C (2010) Mapping the information flow from one brain to another during gestural communication. Proc Natl Acad Sci U S A 107:9388–9393. CrossRef Medline

Singer T, Seymour B, O'Doherty J, Kaube H, Dolan RJ, Frith CD (2004) Empathy for pain involves the affective but not sensory components of pain. Science 303:1157–1162. CrossRef Medline

Stephens GJ, Silbert LJ, Hasson U (2010) Speaker-listener neural coupling underlies successful communication. Proc Natl Acad Sci U S A 107: 14425–14430. CrossRef Medline

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automatic anatomical labelling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15:273–289. CrossRef Medline

Wicker B, Keysers C, Plailly J, Royet JP, Gallese V, Rizzolatti G (2003) Both of us disgusted in my insula: the common neural basis of seeing and feeling disgust. Neuron 40:655–664. CrossRef Medline

Wilson SM, Molnar-Szakacs I, Iacoboni M (2008) Beyond superior temporal cortex: intersubject correlations in narrative speech comprehension. Cereb Cortex 18:230–242. CrossRef Medline